

# Towards Improved Semiconductor Defect Inspection for high-NA EUVL based on SEMI-SuperYOLO-NAS

Ying-Lin Chen<sup>a,b,\*</sup>, Jacob Deforce<sup>b,c,\*</sup>, Vic De Ridder<sup>b,c,\*</sup>, Bappaditya Dey<sup>b,\*</sup>, Victor Blanco<sup>b</sup>, Sandip Halder<sup>d, †</sup>, and Philippe Leray<sup>b</sup>

<sup>a</sup>Faculty of Science, KU Leuven, 3001 Leuven, Belgium

<sup>b</sup>Interuniversity Microelectronics Centre (imec), Kapeldreef 75, 3001 Leuven, Belgium

<sup>c</sup>Faculty of Engineering, Ghent University, 9000 Ghent, Belgium

<sup>d</sup>SCREEN SPE Germany GmbH, Ismaning, 85737, Germany

<sup>\*</sup>Equal Contribution

<sup>†</sup>This research was conducted during Sandip Halder's tenure at imec

## ABSTRACT

Continuous progression of Moore's law brings new challenges in metrology and defect inspection. As the semiconductor industry embraces High-Numerical Aperture Extreme Ultraviolet Lithography (High-NA EUVL), there is a current industry-wide evaluation of this technology for potential pitch reduction in future nodes. One of the primary hurdles in implementing High-NA EUVL in High Volume Manufacturing (HVM) is its low depth of focus. Consequently, suppliers of resist materials are compelled to opt for thin resist and/or new underlayers/hardmask's. Experimental combinations of thin resist materials with novel underlayers and hardmask's seem to pose signal detection challenges due to poor Signal-to-Noise Ratio (SNR). In such a scenario, manual classification of these nano-scale defects faces limitations in terms of required time and workforce, and the robustness and generalizability of outcomes are also questionable. In recent years, vision-based machine learning (ML) algorithms have emerged as an effective solution for image-based semiconductor defect inspection applications. However, developing a robust ML model across various image resolutions without explicit training remains a challenge for nano-scale defect inspection. The goal of this research is to propose a scale-invariant Automated Defect Classification and Detection (ADCD) framework capable to upscale images, addressing this issue. We propose an improvised ADCD framework as **SEMI-SuperYOLO-NAS**, which builds upon the baseline YOLO-NAS architecture. This framework integrates a Super-Resolution (SR) assisted branch to aid in learning high-resolution (HR) features by the defect detection backbone, particularly for detecting nano-scale defect instances from low-resolution (LR) images. Additionally, the SR-assisted branch can recursively generate or reconstruct upscaled images ( $\sim \times 2 / \times 4 / \times 8 \dots$ ) from their corresponding downscaled counterparts, enabling defect detection inference across various image resolutions without requiring explicit training. Moreover, we investigate improved data augmentation strategy aimed at generating diverse and realistic training datasets to enhance model performance. We have evaluated our proposed approach using two original FAB datasets obtained from two distinct processes and captured using two different imaging tools. Finally, we demonstrate zero-shot inference for our model on a new, originating from a process condition distinct from the training dataset and possessing different CD/Pitch characteristics. Our experimental validation demonstrates that our proposed ADCD framework aids in increasing the throughput of imaging tools ( $\sim \times 8$ ) for defect inspection by reducing the required image pixel resolutions.

**Keywords:** semiconductor manufacturing, scanning-electron-microscope, defect inspection, metrology, deep learning, machine learning, neural-architecture-search, super resolution

---

Further author information: (Send correspondence to Bappaditya Dey)  
Bappaditya Dey: E-mail: Bappaditya.Dey@imec.be

## 1. INTRODUCTION

At the advanced semiconductor nodes, the smallest possible pattern structures are targeted, accompanied by increased device complexity. To sustain this scaling trajectory, the industry is currently evaluating the utilization of High-Numerical Aperture Extreme Ultraviolet Lithography (High-NA EUVL).<sup>1</sup> However, EUVL necessitates an exceptionally long inspection time due to its propensity to induce stochastic defects.<sup>2</sup> To detect these nano-scale defects (if stochastic in nature) with a high level of certainty, Electron-beam (e-beam) tools typically need to inspect more areas on a wafer (generally an entire wafer),<sup>3</sup> commonly utilizing Scanning Electron Microscope (SEM) images. However, the e-beam inspection process itself suffers from low throughput and requires high-resolution SEM images. Conversely, SEM imaging throughput is contingent upon image resolution. This problem is depicted in Fig.1, where we observe that the CD-SEM imaging time (in seconds) per image is directly proportional to the image resolution. For instance, acquiring a SEM image at a resolution of (1024x1204) takes nearly twice as long as acquiring the same image at a resolution of (512x512). Another key challenge in integrating High-NA EUVL into High Volume Manufacturing (HVM) is the low depth-of-focus (DoF). This factor is driving resist material suppliers to adopt thin(ner) resists,<sup>4</sup> alongside innovative underlayers and hardmasks. However, the combination of thin resist materials with these novel underlayers and hardmasks often gives rise to signal detection challenges for inspection and metrology equipment, attributable to a low Signal-to-Noise Ratio (SNR), as illustrated in Fig.2.

Our proposed research concept is based on addressing several existing fundamental problems:

- (a) To improve e-beam inspection throughput: The potential to accelerate conventional e-beam inspection lies in the utilization of low-resolution (LR) images. However, this approach encounters challenges in effectively inspecting such images.
- (b) To improve SEM imaging throughput: Two primary factors contribute to increased imaging time:
  - < i > the necessity for higher frame numbers for averaging to enhance Signal-to-Noise Ratio (SNR). However, this requires additional time and can be generally destructive for resist materials, leading to pattern CD (Critical Dimension) shrinkage. Nevertheless, our previous research introduced an unsupervised deep learning denoiser<sup>5</sup> aimed at enhancing the SNR of raw images for low frame numbers while simultaneously maintaining reliable roughness measurements for thin resist,<sup>6</sup> thereby improving CD-SEM imaging throughput.
  - < ii > While a larger pixel size facilitates faster defect inspection within the same Field-Of-View (FOV), acquiring higher-resolution images demands more time due to the additional data and detail that need to be captured and processed. Conversely, lower pixel resolution leads to reduced detail, thus adversely impacting the detection of small defects. This trade-off is illustrated in Fig.1.

In the past, defect inspection was either conducted manually by human experts or relied on less effective rule/reference-based methods or algorithms. However, the challenges mentioned above in e-beam inspection of EUVL processed wafers have significantly degraded the performance of these rule-reference-based methods. With the emergence of computer vision-based advanced deep learning algorithms, particularly Convolutional Neural Networks (CNN)-based approaches, there is a growing popularity in semiconductor defect inspection applications, primarily because of their robustness and generalizability. Our main objective in this research is to accelerate e-beam inspection primarily by utilizing low-resolution SEM images (addressing challenge (a)), and to enhance SEM imaging throughput (providing a solution for challenge (b) < ii >), thereby improving nano-scale defect detection performance.

This research work contributes in the following ways:

1. We propose an improvised ADCD framework as SEMI-SuperYOLO-NAS, integrating a SR-assisted branch to facilitate high-resolution (HR) feature learning for detecting nano-scale defects in LR-images.
2. We demonstrate a scale-invariant ADCD framework capable of upscaling images from their low-resolution counterparts, enabling defect inspection across different image resolutions without explicit training. This facilitates increased image acquisition throughput (#Frames per second/#images per second) for imaging tools.

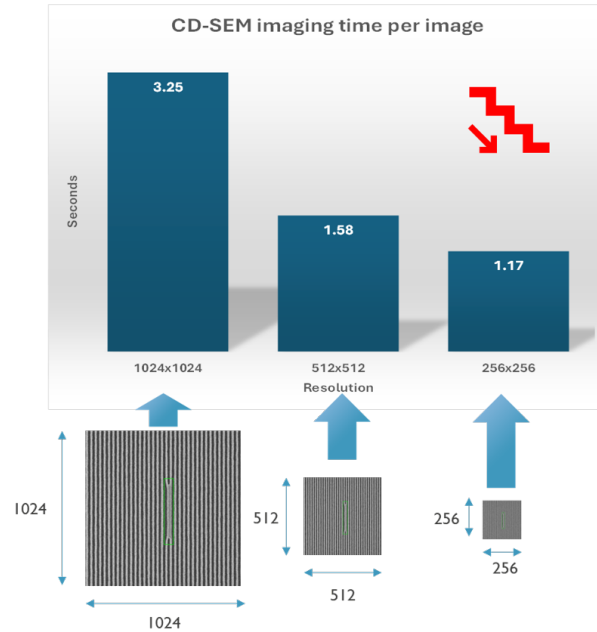


Figure 1. Depiction of CD-SEM imaging time (sec.) against image resolution (WXH).

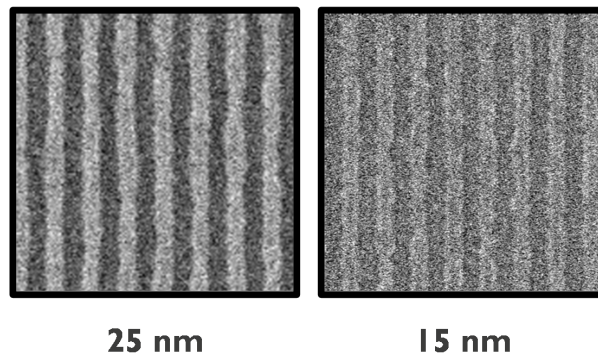


Figure 2. CD-SEM images comparing resist thicknesses of 25nm and 15nm. As the resist thickness decreases, the SEM Signal-to-Noise Ratio (SNR) deteriorates.

3. We investigate an improved data augmentation strategy to create diverse and realistic training datasets, leading to a more accurate and robust ADCD framework.
4. We benchmark the proposed framework on two original FAB datasets, as SEM-ADI and EDR-AEI. Furthermore, we demonstrate the "Zero-Shot" inference capability by making predictions on defect classes not seen during training, utilizing "(weakly) semantic embeddings."

In Section 2, we will review related research works on semiconductor defect inspection. Section 3 will delve into the details of the proposed methods and experiments conducted. Sections 4 and 5 will present our results and conclusions.

## 2. RELATED WORK

In this section, we review various works related to defect inspection in the semiconductor industry employing machine learning techniques. Chang et al.<sup>7</sup> proposed an unsupervised self-organizing neural network to address challenges in post-sawing wafer defect inspection, previously reliant on human inspection, thus requiring significant labor force and suffering from human fatigue. Dey et al.<sup>8</sup> utilized a ResNet architecture-based ensemble

model for nano-scale defect inspection and classification. An unsupervised machine learning model was applied to reduce false-positive defects arising from SEM noise. Dey et al. also benchmarked another ensemble defect detection framework based on the state-of-the-art one-stage object detection model YOLOv5,<sup>9</sup> which outperformed the previous ResNet-based ensemble framework.

A data-centric approach was also demonstrated,<sup>10</sup> showing a way to consensually ensemble labels from different anonymous labelers for the same dataset, requiring minimal intervention from experts for defect labeling expectations. The final combined label, along with expert-aware post-processing data, yielded the highest test mAP of 0.919 with the state-of-the-art YOLOv8 model. The concept of pseudo-label prediction and model ensembles from different labelling partitions to avoid duplicate labeling work was proposed. Cheon et al.<sup>11</sup> utilized a Convolutional Neural Network (CNN) along with a k-NN method to classify wafer surface defects from SEM images. The final activation layer outputs were used as features for k-Nearest Neighbors (k-NN), but with a threshold for clustering, allowing the algorithm to identify unknown defect classes. However, when a new class is discovered, the CNN needs retraining to ensure consistent high performance over time. To address this issue, the authors proposed the idea of incremental training. In previous study,<sup>12</sup> a noise characterization for future denoising algorithms demonstrated that as scanning speed increased, a shift in noise property from a Gaussian to a Gamma distribution was observed. In recent years, CD-SEM tool vendors have also incorporated machine learning to enhance tool efficiency.<sup>2,3</sup> The idea of ML-based super-resolution defect inspection was proposed,<sup>3</sup> where a CNN model takes the database of designed features and CD-SEM images as input and generates the pixel distribution for defect prediction. Their model utilized non-defect regions to train the algorithm, generating the pixel luminosity distribution for the given database. If the pixel brightness values of the original image significantly deviated from the predicted pixel distribution center, the pixel was classified as defective. This approach demonstrated the potential for identifying bridge defects and necking defects even in lower pixel resolution images of 4nm/pixel. However, in this method, knowledge of the design database (mask layout) is necessary, making automated defect type classification challenging. Nonetheless, their research highlights the advantages of massive metrology, showcasing how a FOV of 80×80 microns can facilitate automatic defect identification through machine learning. Thus, the development of a super-resolution ML method capable of both identifying and classifying defects without requiring knowledge of the mask design remains a crucial missing component.

### 3. METHODOLOGY

In this section, we provide a brief introduction to our training datasets, the proposed data augmentation strategy, an overview of the architecture of our proposed SEMI-SuperYOLO-NAS, and the strategy for preparing the dataset of High-resolution/Low-resolution (HR/LR) image pairs to train our SR-branch assisted defect detector model.

#### 3.1 Dataset

Two Line-Space (LS) pattern datasets were utilized. The first dataset, SEM-ADI, comprises 1324 images captured with the CD-SEM tool during After-Development-Inspection (ADI). The second dataset, EDR-AEI, comprises 527 images captured with the EDR tool during After-Etching-Inspection (AEI). Each image in these datasets contains at least one defect. SEM-ADI and EDR-AEI have original pixel resolutions of 1024×1024 and 480×480, respectively. SEM-ADI encompasses 5 different defect types, while EDR-AEI has 4 different defect types. We divided the two datasets into training and validation sets, as outlined in Table 1 and Table 2, where we also provide details regarding the original defect classes and their respective total number of instances. In Fig.3, we illustrate the various defect types in the SEM-ADI dataset. In Fig.4, we present all defect types in the EDR-AEI dataset. The imbalance in defect classes and the limited availability of data in both datasets prompted us to develop novel augmentation strategies tailored to SEM and EDR imaging characteristics and assess their impact on model performance.

#### 3.2 Proposed Data-augmentation strategy

We implemented different image augmentation<sup>13</sup> techniques to enhance the robustness of our model and address imbalances in the number of instances across defect classes. Within our image augmentation pipeline, utilizing the Albumentations library,<sup>14</sup> we integrated the following enhancements to introduce variability in illumination, noise, and atmospheric conditions:



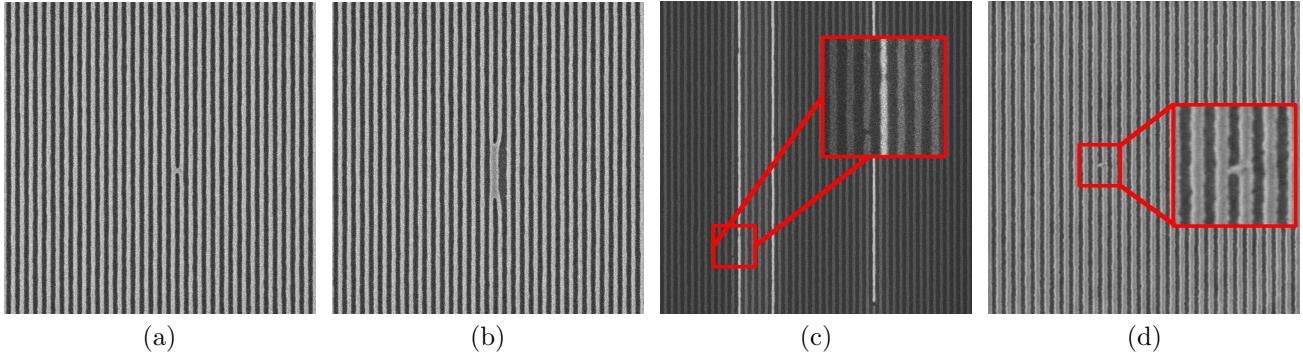


Figure 3. Defect class examples in SEM-ADI dataset. From left to right: (a) Bridge, (b) Line-collapse, (c) Gap & Probable-gap and (d) Micro-bridges.

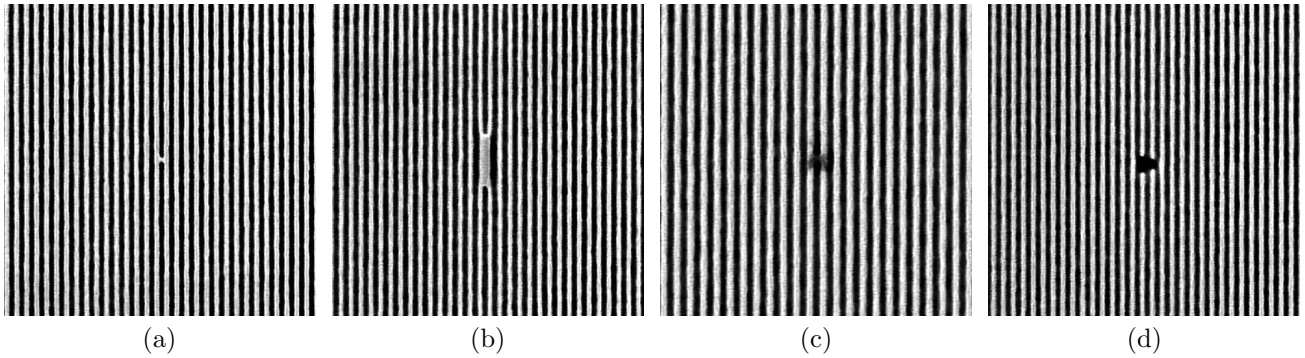


Figure 4. Defect class examples in EDR-AEI dataset. (a) bridge (b) pattern-collapse (c) dark-spot and (d) break.

- Random Shadow: Generates sharp shadows of a random polygon shapes across the image.
- GaussNoise: Applies per-pixel-gaussian-distributed grayscale noise.
- RandomFog: Introduces blurry spots randomly throughout the image.
- RandomBrightness: Randomly changes the brightness and contrast of the input image.
- RandomGamma: Modifies image luminance using gamma correction to simulate different lighting conditions.

Table 1. SEM-ADI dataset with corresponding defect class distribution (Train/Val)

Class Name	Train (1054 images)	Val (272 images)
micro-bridge	380	125
gap	1046	330
bridge	238	36
line-collapse	550	142
probable gap	315	103

Table 2. EDR-AEI dataset with corresponding defect class distribution (Train/Val)

Class Name	Train (425 images)	Val (104 images)
break	126	28
bridge	107	27
dark spot	255	60
pattern collapse	137	34

- RingingOvershoot: Convolves an image with a 2D sinc filter to create a typical ringing artefact obtained when processing a bandlimited signal. Subtle edges appear around sharp color changes.

Additionally, we integrated a Cutout augmentation,<sup>15</sup> which involves removing a portion of the image. Moreover, we introduced two novel augmentations specifically for Line-Space pattern SEM images. Fig.5 illustrates the above-mentioned augmentation methods on the EDR-AEI dataset.

- Copy defect: Replicates defect instances in different locations within the image. A rule-based algorithm aligns the pattern of the copied slice with the background, and its edges are faded to avoid any artefacts.
- ContrastChange: We sporadically decreased the horizontal darkness and contrast of the image. This mimics the effect observed in the SEM-ADI dataset.

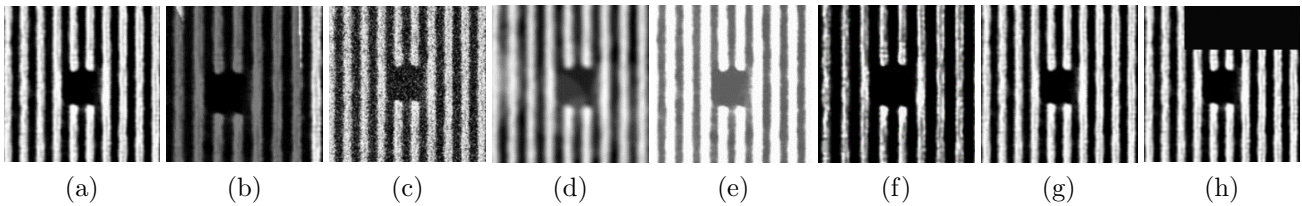


Figure 5. Sample of augmentation on EDR-AEI break defect. (a) original image, (b) RandomShadow, (c) GaussNoise, (d) RandomFog, (e) RandomBrightness, (f) RandomGamma, (g) RandomOvershoot, and (h) Cutout.

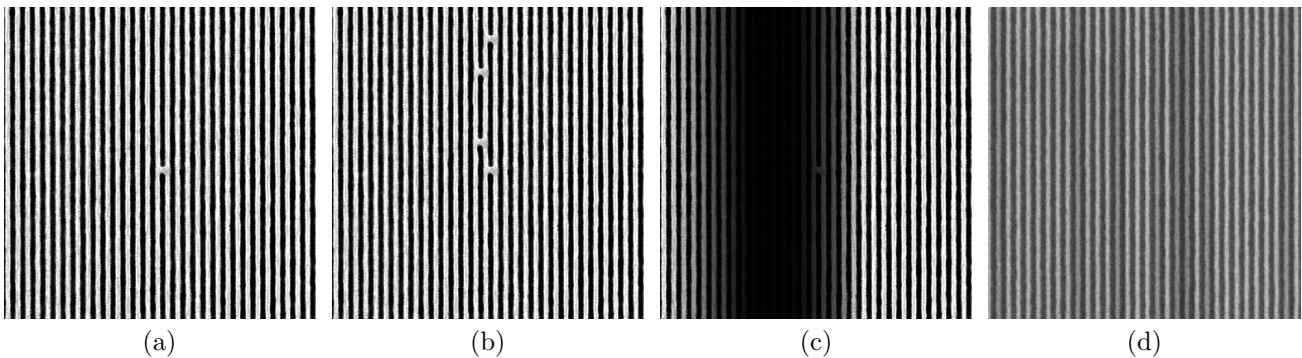


Figure 6. (a) original bridge defect (b) Copy Defect augmentation (c) ContrastChange augmentation and (d) original contrast change occurrence in SEM-ADI image.

We exclusively apply augmentation solely to the training dataset, aiming to further enhance the model's comprehension of semantic embeddings associated with defect types. During validation, we only use the original validation dataset, abstaining from any augmentation. The two novel augmentations tailored for LS defects are shown in Fig.6.

As discussed in the preceding section, we strategically apply augmentations on different defect types. For defect types that are easily discernible or already possess high instance numbers, we employ fewer augmentations. Conversely, for defect types that are sparse in the original data or pose greater challenges for ADCD, we deliberately apply more augmentations. This approach reallocates training efforts towards addressing challenging tasks, while not over-allocating resources to defect types that are already reliably detected. The original number of defect instances, augmented defect instances, and the total number of instances after augmentation are detailed in Table 3.

Table 3. Summary of Defect Instances in Original and Augmented Training Datasets.

Dataset	Defect type	# of instances in original images	# of instances in augmented images	total # of instances
SEM-ADI	micro-bridge	380	350	730
	gap	1046	44	1090
	bridge	238	117	355
	line-collapse	550	44	594
	probable gap	315	714	1029
EDR-AEI	break	126	451	577
	bridge	107	450	557
	dark spot	255	902	1157
	pattern-coll.	137	360	497

### 3.3 Dataset preparation for training SR-assisted branch

To facilitate training of SuperYOLOv5<sup>16</sup> and proposed SEMI-SuperYOLO-NAS, we prepared datasets comprising pairs of High-resolution (HR) and Low-resolution (LR) images. Images were initially downsampled using the Bilinear interpolation method before any augmentation, as discussed in Section 3.2. For the SEM-ADI dataset, we downsampled the original resolution of  $(1024 \times 1024)$  to consecutive low-resolutions of  $(512 \times 512)$ ,  $(256 \times 256)$ , and finally  $(128 \times 128)$ . Similarly, for the EDR-AEI dataset, we downsampled the original image resolution of  $(480 \times 480)$  to  $(240 \times 240)$  and finally  $(120 \times 120)$ . Fig.7 illustrates an example of downsampling for the EDR-AEI dataset. To train the SR-branch utilized by Super-YOLOv5<sup>16</sup> and SEMI-SuperYOLO-NAS, higher resolution (HR) images were paired with corresponding lower resolution (LR) images as [HR/LR  $\rightarrow$  1024/512, 512/256, 256/128]. Additionally, baseline YOLO-NAS models were trained individually.

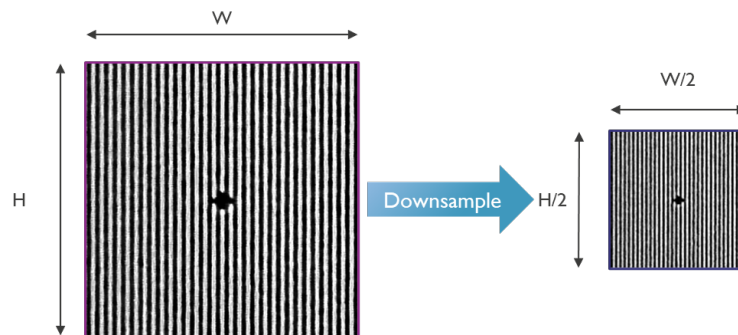


Figure 7. Downsampling example: Bilinear Interpolation method applied to EDR-AEI Image to generate LR  $(240 \times 240)$  from HR  $(480 \times 480)$ .

### 3.4 Proposed SEMI-Super-YOLO-NAS architecture

Our proposed SEMI-SuperYOLO-NAS architecture builds upon the recent baseline YOLO-NAS object detection architecture.<sup>17</sup> Neural Architecture Search (NAS)<sup>18</sup> enables automatic exploration of a search space comprising candidate neural network architectures tailored for specific applications, such as object detection, guided by a search algorithm. This process involves sampling numerous architectures from the search space and evaluating them to select the most optimal configuration, thus replacing conventional neural network design methods typically performed by human experts. Another advantage of the YOLO-NAS architecture is its quantization-aware nature, which allows for striking a balance between information loss and latency. This contrasts with conventional object detection models, which often sacrifice precision in model weights to achieve a similar balance. The baseline YOLO-NAS utilizes the QARepVGG<sup>19</sup> building block alongside the NAS algorithm to enhance accuracy preservation post-quantization. Fig.8 depicts the proposed SEMI-SuperYOLO-NAS architecture based ADCD

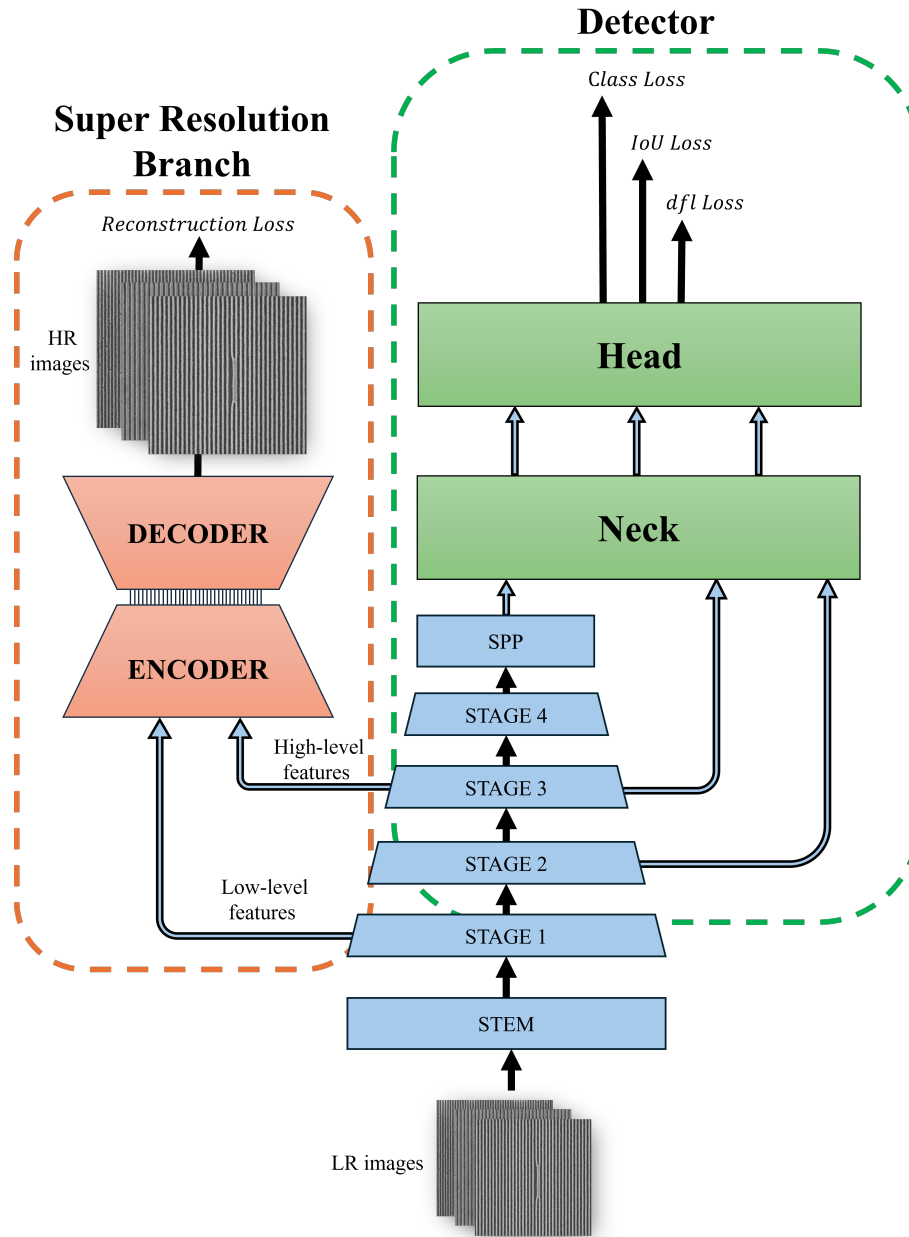


Figure 8. Proposed SEMI-Super-YOLO-NAS architecture.

framework. Zhang et al.<sup>16</sup> proposed adding a Super-Resolution (SR) branch to the customized YOLOv5 architecture.<sup>20</sup> They replaced the Focus module with a Multimodal Fusion module to prevent resolution degradation and spatial information loss, particularly for small object detection tasks. The SR structure consisted of a simple Encoder-Decoder model. This branch takes two backbone layers as input (one containing high-level features and the other containing low-level features) and outputs a high-resolution (HR) version of the original input image. Consequently, the loss function includes an additional term: the absolute difference between the reconstructed super-resolution image and the original HR image. This super-resolution loss function term facilitates learning HR feature extraction by the backbone, leading to improved detection performance, as demonstrated by Zhang et al.<sup>16</sup> in their original work.



Inspired by this approach, we incorporated a Super-Resolution branch into the baseline YOLO-NAS's architecture variants. We adopted Zhang et al.'s<sup>16</sup> design for the SR-CNN architecture, which was specifically designed to minimize the impact on training time while still facilitating learning by the backbone. As mentioned earlier, this SR branch should take input from two locations within the YOLO-NAS-s backbone:

- One where low-level features have been extracted. For this purpose, we selected the output of the first backbone stage.
- One where high-level features have been extracted. Here, we utilized the output of the third backbone stage.

Finally, we recursively train SEMI-SuperYolo-NAS on both SEM-ADI and EDR-AEI datasets, as :

1. SR-assisted branch with  $(x_{LR}^i, y_{HR}^i)$  image pairs to facilitate high-resolution features learning by the defect detection backbone,
  - $x_{LR}$   $\rightarrow$  input downsampled image (say, 512, 256, 128,...)
  - $y_{LR}$   $\rightarrow$  corresponding output super resolution/upscaled images (say, 1024, 512, 256,...)
2. Defect detection branch with  $(x_{LR}^i, a_{\text{bbox}}^i)$ ,
  - $x_{LR}$   $\rightarrow$  input downsampled image (say, 512, 256, 128,...)
  - $a_{\text{bbox}}$   $\rightarrow$  corresponding bounding-box GT annotation of the defect instance location.

### 3.5 Training methodology

All experiments were conducted on Nvidia A100 40GB GPUs. The baseline YOLO-NAS<sup>17</sup> model with three different architecture variants as large, medium, and small are trained. For comparison analysis with previous work, we trained Super-YOLOv5.<sup>16</sup> Additionally, we trained the proposed SEMI-SuperYOLO-NAS with integrated SR-branch, utilizing the default data augmentation strategy following the YOLO-NAS baseline training approach. Finally, we trained the proposed SEMI-SuperYOLO-NAS with integrated SR-branch using our suggested data augmentation strategy outlined in Section 3.2, replacing the previous default data augmentation strategy.

## 4. RESULTS AND DISCUSSION

The inference results of our proposed SEMI-SuperYOLO-NAS model are depicted in Fig.9(a)(b). Our discussion and analysis are divided into three sections:

1. Firstly, we benchmarked our proposed approach (SEMI-SuperYOLO-NAS with our proposed data augmentation strategy) against baseline YOLO-NAS, previous SuperYOLOv5,<sup>16</sup> and SEMI-SuperYOLO-NAS with the default data augmentation strategy. This benchmarking was conducted on the SEM-ADI dataset for image resolution pair (512  $\rightarrow$  1024) and on the EDR-AEI dataset for image resolution pair (240  $\rightarrow$  480).
2. We demonstrate zero-shot inference on a new CD-SEM test dataset, comprising image pairs captured at the exact same locations but with two different resolutions (1024 $\times$ 1024 and 512 $\times$ 512). This newly collected dataset, also a SEM-ADI dataset, originates from a process condition distinct from the training dataset and possesses different CD/Pitch characteristics. The observable defects in these images closely resemble line-collapse and bridge/micro-bridge defects observed in the training images, albeit not identical.
3. Finally, we assess the benefit of employing the SR-branch of our proposed SEMI-SuperYOLO-NAS model. This involves generating or reconstructing upscaled images from their corresponding downsampled counterparts (e.g., 512  $\rightarrow$  1024) and conducting detection inference by applying baseline YOLO-NAS on these artificially upscaled images. This approach aims to facilitate more precise nano-scale defect inspection across various image resolutions without the need for explicit training.

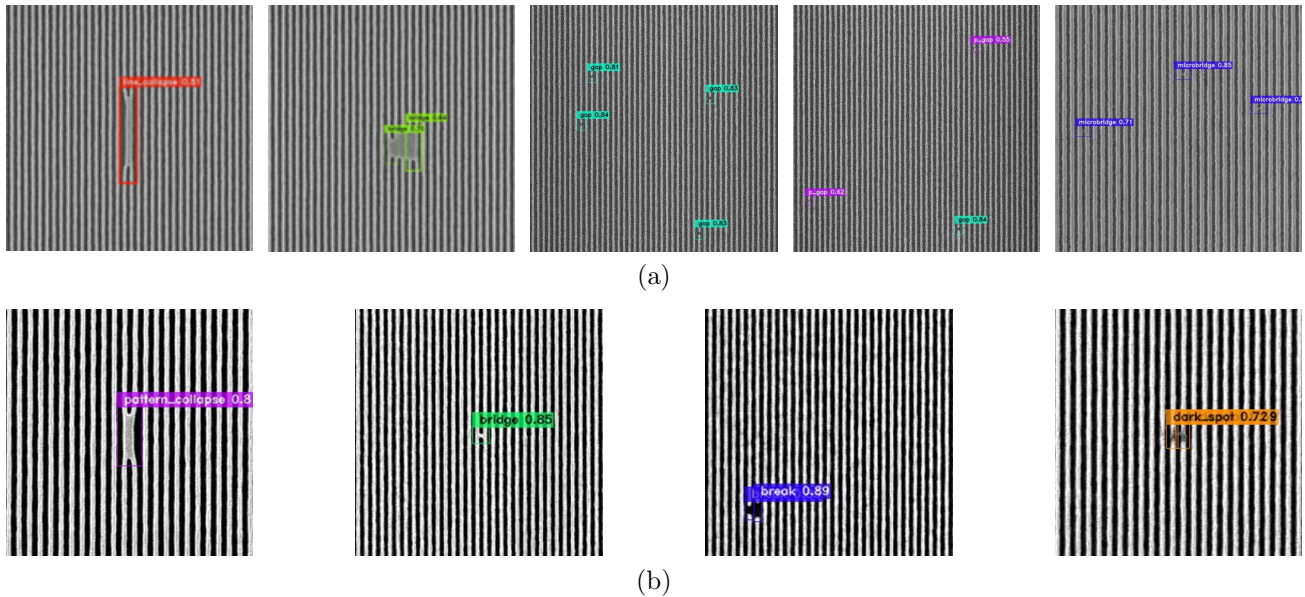


Figure 9. Inference results on (a) SEM-ADI test set [image resolution pair (512  $\rightarrow$  1024)] (upper) and on (b) EDR-AEI test set [image resolution pair (256  $\rightarrow$  512)] (bottom) of our proposed SEMI-SuperYOLO-NAS architecture.

#### 4.1 Benchmarking proposed SEMI-SuperYOLO-NAS ADCD framework

We assessed four models using different strategies on the SEM-ADI dataset for the resolution pair (LR/HR) [512 $\rightarrow$  1024] and on the EDR-AEI dataset for the resolution pair (LR/HR) [240  $\rightarrow$  480], and provided a comparative analysis for:

1. YOLO-NAS (baseline),<sup>17</sup>
2. Super-YOLOv5,<sup>16</sup>
3. Proposed SEMI-SuperYOLO-NAS with default YOLO-NAS data augmentation strategy,
4. Proposed SEMI-SuperYOLO-NAS with proposed data augmentation strategy outlined in Section 3.2

Table 4 illustrates that SuperYOLOv5<sup>16</sup> outperforms all other defect detector models on the SEM-ADI dataset. However, as we propose a novel ADCD framework based on the recent YOLO-NAS architecture, our primary objective is to enhance its detection capability to match the performance of SuperYOLOv5.<sup>16</sup> The baseline YOLO-NAS architecture performs the worst for all defect types, achieving a mAP of 0.47. While the performance for relatively prominent defect types like line-collapse and bridge is still comparable, it severely struggles with more challenging nano-scale defects such as gap, probable-gap, and micro-bridge. It is possible that the modules of the baseline YOLO-NAS architecture, such as the backbone, neck, and context modules, are not equipped to handle the intricacies of nano-scale defect detection, thus failing to meet the requirements in these aggressive pitches.

In contrast, our proposed SEMI-SuperYOLO-NAS outperforms the baseline YOLO-NAS<sup>17</sup> architecture in both per class AP and mAP metrics, achieving an mAP of 0.54 with the default YOLO-NAS data augmentation strategy and a slightly improved mAP of 0.58 with the proposed data augmentation strategy, which replaces the previous default strategy. For micro-bridge defects, the AP improves by 51.06% and 57.44% with the corresponding data augmentation strategies when applying the proposed SEMI-SuperYOLO-NAS architecture compared to the baseline YOLO-NAS. The probable gap defect was not detected at all by the baseline YOLO-NAS model, whereas the proposed SEMI-SuperYOLO-NAS architecture, with both data augmentation strategies, achieved a per class AP of up to 0.13. For challenging gap defects, the proposed SEMI-SuperYOLO-NAS with





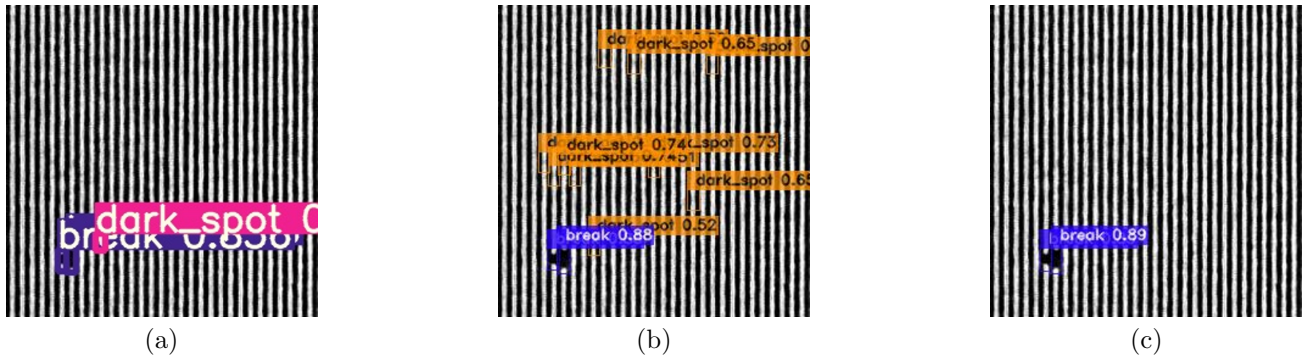


Figure 12. Detection results of break: (a) Baseline YOLO-NAS (b) Proposed SEMI-SuperYOLO-NAS + Default data augmentation (c) Proposed SEMI-SuperYOLO-NAS + proposed data augmentation. \*\*False-Positive detections have been removed/reduced with our proposed data augmentation strategy towards improved detection accuracy.

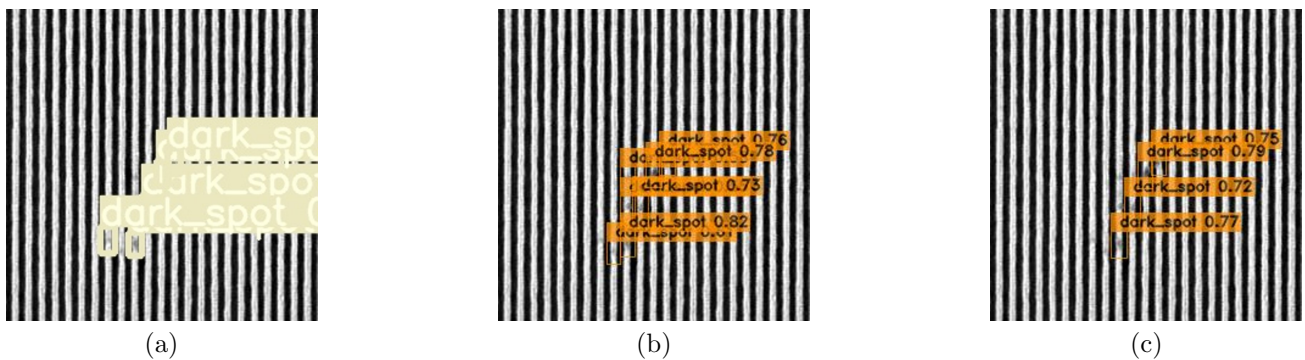


Figure 13. Detection results of dark-spot: (a) Baseline YOLO-NAS (b) Proposed SEMI-SuperYOLO-NAS + Default data augmentation (c) Proposed SEMI-SuperYOLO-NAS + proposed data augmentation. \*\*False-Positive detections have been removed/reduced with our proposed data augmentation strategy towards improved detection accuracy.

#### 4.2 Zero-shot inference with newly acquired High/Low resolution image pair (1024/512)

To demonstrate zero-shot inference across different resolutions and real SEM imaging conditions, we collected a new SEM-ADI dataset from a process condition distinct from the training dataset, as depicted in Fig.16. This novel dataset has a different CD/Pitch and was obtained from the same location at two different resolutions: 512 and 1024, respectively, as shown in Fig.16. The observable defects in these images closely resemble line-collapse and bridge/micro-bridge defects seen in the training images. Our primary aim was to ascertain whether our proposed ADCD framework could obviate the need for explicit training across diverse image conditions (process, resolutions, CD/Pitch) and acquire "weakly" semantic embeddings between closely matched defect categories,

Table 4. Per class AP and mean AP @IoU 0.5 on SEM-ADI dataset. Best, Second best and Third best colour codes: GREEN, RED, BLUE.

Defect Class	Defect Detector Models			
	YOLO-NAS (baseline)	SuperYOLOv5 <sup>16</sup>	Proposed SEMI-SuperYOLO-NAS (+Default Aug)	Proposed SEMI-SuperYOLO-NAS (+Proposed Aug)
micro-bridge	0.47	0.73	0.71	0.74
gap	0.15	0.95	0.15	0.19
bridge	0.73	0.76	0.79	0.84
line-collapse	0.99	1.00	1.00	1.00
probable gap	0.00	0.33	0.08	0.13
mAP	0.47	0.75	0.54	0.58

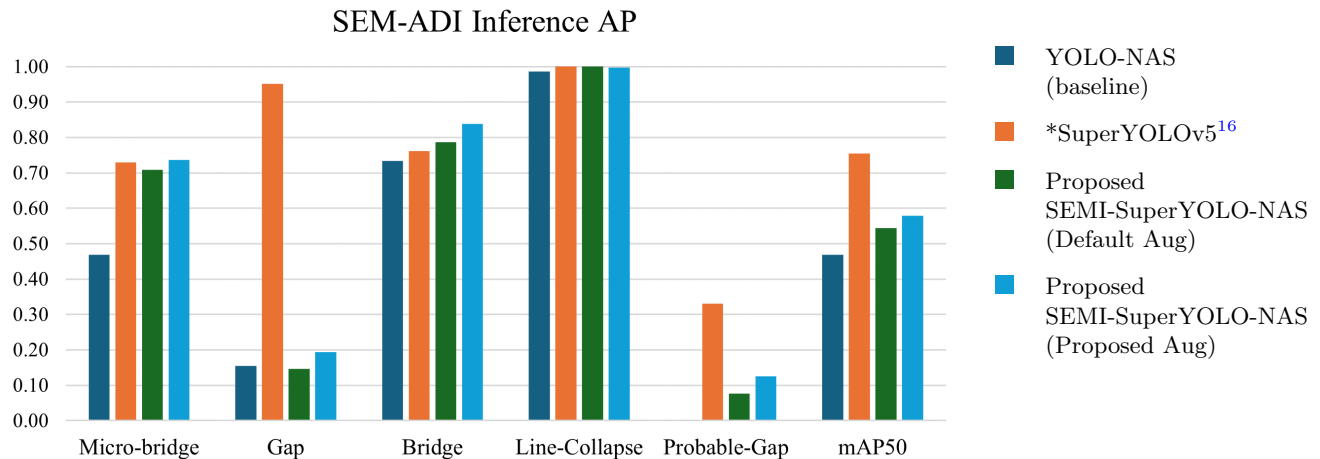


Figure 14. Comparison analysis of defect detection performance [per class AP and mean AP] on SEM-ADI dataset. [\*\*The original SEM image of resolution 1024 was downsampled to 512, and all detector architecture variants are Small].

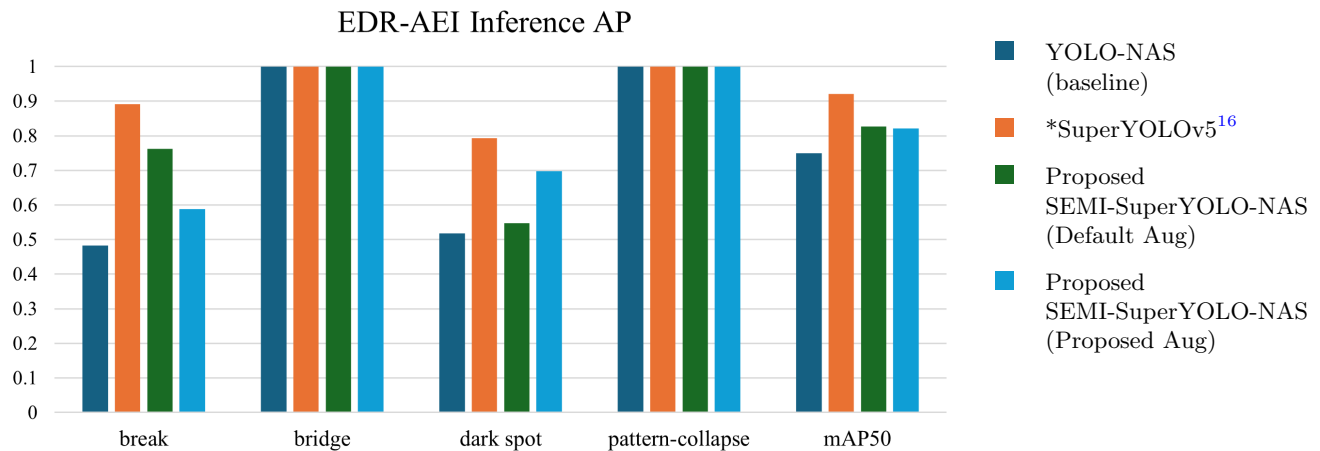


Figure 15. Comparison analysis of defect detection performance [per class AP and mean AP] on EDR-AEI dataset. [\*\*The original SEM image of resolution 512 was downsampled to 256, and all detector architecture variants are Small].

as demonstrated in Fig.17. Fig.18 depicts defect detection inference result (with confidence threshold  $\geq 0.5$ ) on this newly acquired ADI-SEM image of resolution 1024, using the baseline YOLO-NAS model trained on the SEM-ADI dataset of the same resolution (dataset shown in Fig.16). Similarly, Fig.19 illustrates the defect detection inference results (at different confidence thresholds: 0.1, 0.2, 0.3, and 0.5) on the identical dataset but at a lower resolution of 512. This was achieved using the same YOLO-NAS model trained on the previous SEM-ADI dataset of resolution 512 (dataset shown in Fig.16). These SEM images depicted in Fig.19 are taken at the exact same location as Fig.18 but with lower resolution using the imaging tool's settings. Notable observations include a decrease in the detection of defect instances for the image resolution of 512 compared to 1024, and an increase in the detection of defect instances with a lower confidence threshold (for example, the model detected more instances with a confidence threshold of 0.1 compared to when the confidence threshold was 0.5). Lowering the confidence threshold is expected as this dataset was never seen by the model. This highlights an existing bottleneck with supervised machine learning applications that should be addressed in the near future.

Table 5. Per class AP and mean AP @IoU 0.5 on EDR-AEI dataset. . Best, Second best and Third best colour codes: GREEN, RED, BLUE.

Defect Class	Defect Detector Models			
	YOLO-NAS (baseline)	SuperYOLOv5 <sup>16</sup>	Proposed SEMI-SuperYOLO-NAS (+Default Aug)	Proposed SEMI-SuperYOLO-NAS (+Proposed Aug)
break	0.48	0.89	0.76	0.59
bridge	1.00	1.00	1.00	1.00
dark spot	0.52	0.79	0.55	0.70
pattern coll.	1.00	1.00	1.00 </td <td>1.00</td>	1.00
mAP	0.75	0.92	0.83	0.82

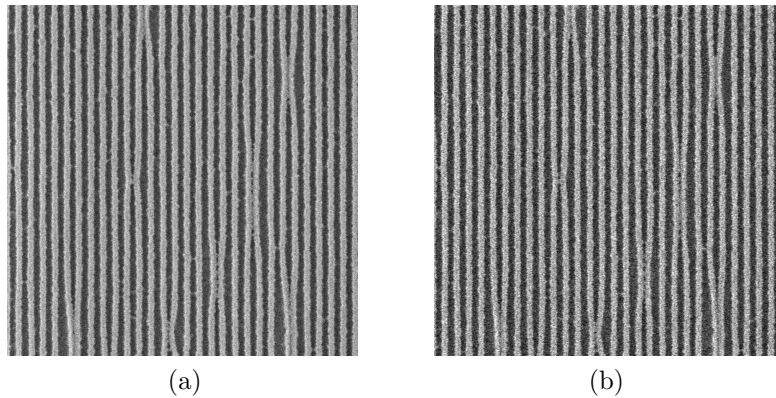


Figure 16. CD-SEM images of new collected defect at same location, (a) and (b) are 1024 and 512, respectively.

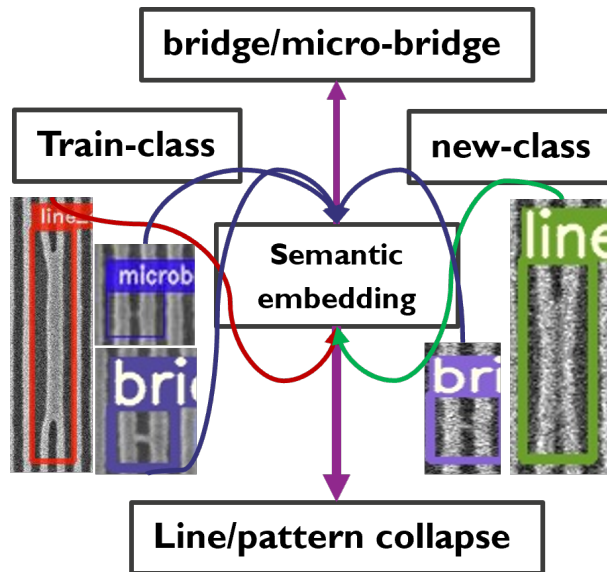


Figure 17. Illustration of zero-shot inference: predictions on defect classes not encountered during training. A "weakly" semantic embedding can be observed between the defect "training" class (left) and the "inference" class (right) of two different process conditions



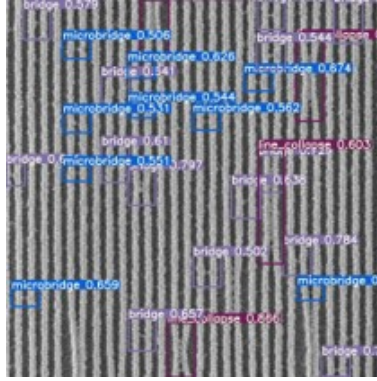


Figure 18. Illustration of zero-shot inference: Detection of defect instances using baseline YOLO-NAS on new SEM-ADI test dataset (Line-Space) with resolution 1024 @confidence threshold 0.5. This dataset was collected from a process distinct from training dataset, with different CD/Pitch and imaging conditions. \*\*YOLO-NAS model, trained with SEM-ADI dataset [as shown in Fig. 2] at a resolution of 1024, is employed for inference.

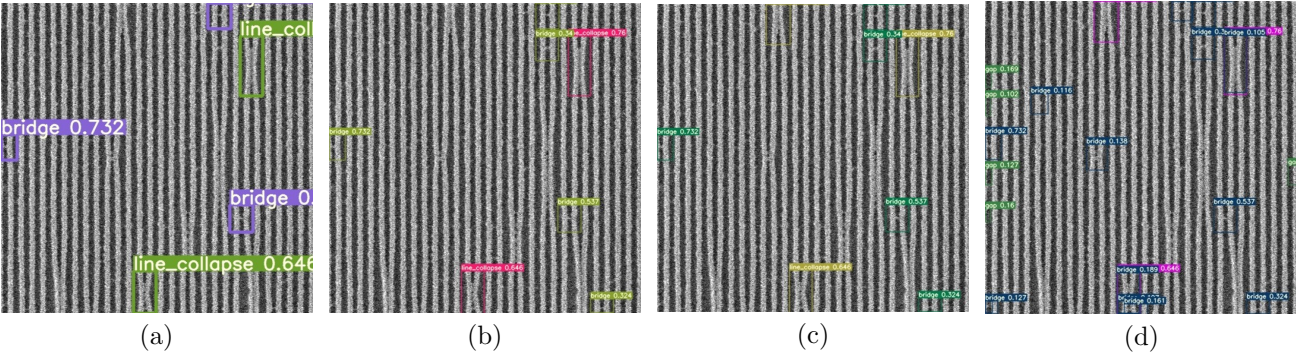


Figure 19. Illustration of zero-shot inference: Detection of defect instances using baseline YOLO-NAS on new SEM-ADI test dataset (Line-Space) with resolution 512 for different confidence thresholds. (a) 0.5, (b) 0.3, (c) 0.2, and (d) 0.1. This dataset was collected from a process distinct from training dataset, with different CD/Pitch and imaging conditions.. \*\*YOLO-NAS model, trained with the downsampled SEM-ADI dataset [as shown in Fig.3] at a resolution of 512, is employed for inference.

### 4.3 Inference with baseline YOLO-NAS on Super-Resolution images [Generated/Reconstructed with SR-assisted branch]

The primary objective of this research is to propose a scale-variant ADCD framework capable of upscaling images to enable more accurate nano-scale defect inspection across various image resolutions without explicit training. In this experiment, we initially trained the baseline YOLO-NAS<sup>21</sup> models on each image resolution, including 1024, 512, 256, 128, and 64, both with and without the proposed data augmentation strategy outlined in Section 3.2. We then reported the best mAP metric @IoU 0.5 for each resolution. The dataset preparation strategy for corresponding HR/LR image-pairs, as discussed in Section 3.3, was employed. We observed that the lowest resolution allowing for detection was 128×128 with our experimental setup. Beyond this resolution, detecting defects became challenging due to the limited number of pixels defining the line and space.

Subsequently, we incrementally upsampled each low-resolution (LR) image set (128, 256, and 512) by a factor of 2 (×2) and generated/reconstructed corresponding high-resolution (HR) image sets (256, 512, and 1024). Initially, we applied the SR-assisted branch of the previous SuperYOLOv5<sup>16</sup> and then the proposed SEMI-SuperYOLO-NAS. Our observations revealed that the detection performance on upsampled images (1024, 512, 256) using SuperYOLOv5<sup>16</sup> SR-branch was comparatively lower than that on corresponding low-resolution images (512, 256, 128), as illustrated in Fig.20 (b) against 20 (a). However, the detection performance on upsampled images (1024, 512, 256) using our proposed SEMI-SuperYOLO-NAS SR-branch was relatively higher

or on par with corresponding low-resolution images (512, 256, 128), as depicted in Fig.20(c) against 20 (a). The performance of defect detection is illustrated in Fig.21.

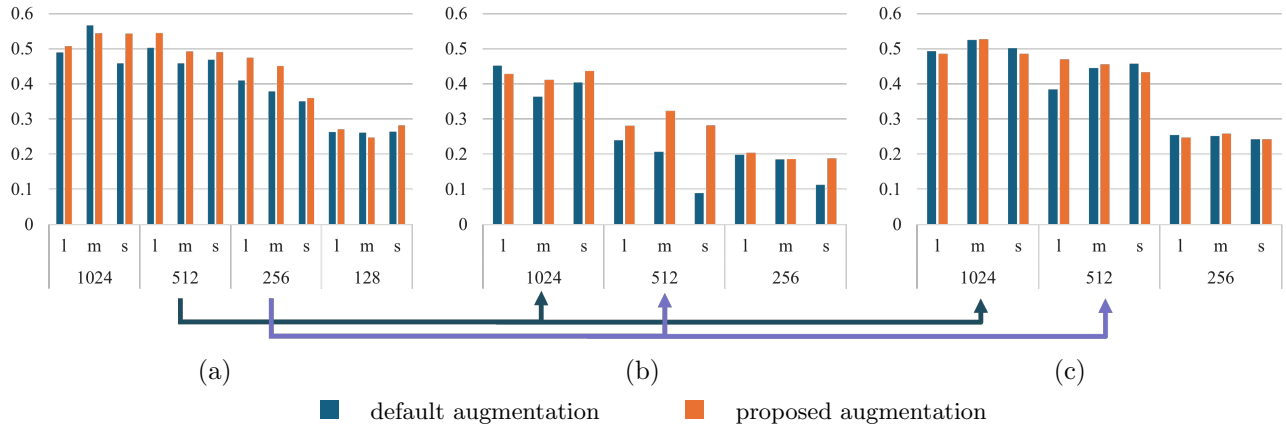


Figure 20. (a) Baseline YOLO-NAS inference results (mAP @IoU0.5) on ADI-SEM dataset for different image resolutions. [Ground-Truth image resolution: 1024, Downsampled (Bilinear interpolation) image resolutions: 128, 256, 512]. (b) Baseline YOLO-NAS inference results (mAP @IoU0.5) on ADI-SEM dataset for different image resolutions, upsampled by Super-YOLOv5<sup>16</sup> [512→1024, 256→512, 128→256]. (c) Baseline YOLO-NAS inference results (mAP @IoU0.5) on ADI-SEM dataset for different image resolutions, upsampled by proposed SEMI-SuperYOLO-NAS [512→1024, 256→512, 128→256]. \*\*Shown for baseline YOLO-NAS model sizes as Small (S), Medium (M) and Large (L) with and without proposed augmentation strategy.

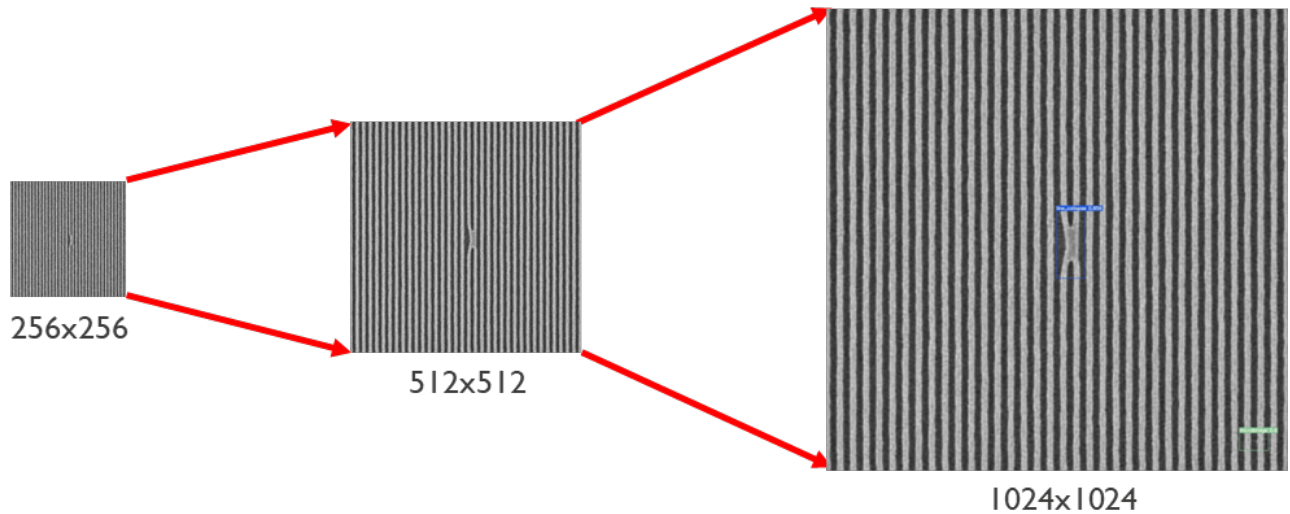


Figure 21. Detection of defect instances using baseline YOLO-NAS on upsampled image resolution [256 → 1024] with the proposed SEMI-SuperYOLO-NAS SR branch.

## 5. LIMITATION AND FUTURE WORK

The present study highlights several limitations and suggests avenues for future research based on the findings. Firstly, while the SR-branch employed improves intermediate high-resolution feature learning, it lacks optimization for high-quality SR image reconstruction or generation. One potential future direction involves training a novel SR-branch specifically tailored specifically for generating high-quality Super-Resolution SEM images. Moreover, the study reveals that the baseline YOLO-NAS struggles with detecting critically small defects, particularly



gap and probable-gap anomalies. Further investigation is necessary to determine if the NAS space is restricted by limited computational resources, potentially leading to suboptimal results for specific defect classes. This is particularly pertinent considering that NAS was conducted using the Object365 dataset and COCO Pseudo labeled dataset, which possess different characteristics compared to semiconductor image datasets. Additionally, the utilization of low-resolution/downscale images obtained through Bilinear interpolation poses challenges, as it fails to effectively handle high-frequency details or sharp transitions in the image, resulting in image blurring, loss of fine details, and aliasing artifacts. Future endeavors could focus on employing more sophisticated algorithms such as Bicubic interpolation, Lanczos resampling, or advanced deep learning methods to address these challenges. Additionally, the acquisition of original wafer SEM images at various resolution settings (such as 128, 256, 512, 1024,...) could offer valuable data for further research and development in this field.

## 6. CONCLUSION

In this research, we introduce a novel architecture of Super-YOLO-NAS for semiconductor defect inspection, aimed at enhancing SEM throughput by reducing required SEM image pixel resolutions by approximately a factor of 8 ( $\sim \times 8$ ). Compared to the baseline YOLO-NAS, our proposed methodology enhances detection precision by 23% for SEM-ADI and 9% for EDR-AEI dataset, at lower SEM resolutions. Furthermore, we illustrate zero-shot inference on a novel SEM-ADI test dataset, obtained under distinct process and imaging conditions, diverging from the distribution of the training dataset. The proposed architecture demonstrates robustness and generalizability, proficiently capturing semantic embeddings for specific defects such as line collapse, bridging, and micro-bridging, even when the train and test datasets pertain to different process conditions, varied CD/pitch, and imaging conditions. Finally, we demonstrated the performance of the baseline YOLO-NAS inference on upsampled images, which were generated from their respective low-resolution counterparts using the SR-assisted branch of the proposed Super-YOLO-NAS. This evaluation was conducted by comparing against upsampled images generated by the SR-assisted branch of Super-YOLOv5, a previous framework based on the YOLOv5 model.

## Author Contribution

- Bappaditya Dey conceived the idea, planned, and led the project.
- Bappaditya Dey proposed the methodology in discussion with Ying-Lin Chen, Vic De Ridder, Jacob Deforce, Victor Blanco, and Sandip Halder.
- Ying-Lin Chen, Vic De Ridder and Jacob Deforce performed experiments with the advice of Bappaditya Dey.
- Ying-Lin Chen, Vic De Ridder and Jacob Deforce performed data analysis with the advice of Bappaditya Dey.
- Ying-Lin Chen, Vic De Ridder and Jacob Deforce wrote the manuscript.
- Bappaditya Dey reviewed and edited the final manuscript.
- All the authors contributed to the discussions on the results and reviewed the manuscript.

## REFERENCES

- [1] Weiss, M., "Overlay challenges in the era of high-na," in [*Metrology, Inspection, and Process Control XXXVII*], **12496**, 1249603, SPIE (2023).
- [2] Ouchi, M., Ishikawa, M., Shinoda, S., Toyoda, Y., Yumiba, R., Shindo, H., and Izawa, M., "A trainable die-to-database for fast e-beam inspection: learning normal images to detect defects," in [*Metrology, inspection, and process control for microlithography XXXIV*], **11325**, 437–445, SPIE (2020).

- [3] Kondo, T., Ban, N., Ebizuka, Y., Toyoda, Y., Yamada, Y., Kashiwa, T., Koike, H., Shindo, H., Charley, A.-L., Saib, M., 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*], **11611**, 210–219, SPIE (2021).
- [4] Lorusso, G. F., Beral, C., Bogdanowicz, J., De Simone, D., Hasan, M., Jehoul, C., Moussa, A., Saib, M., Zidan, M., Severi, J., et al., “Metrology of thin resist for high na euvl,” in [*Metrology, Inspection, and Process Control XXXVI*], **12053**, 229–240, SPIE (2022).
- [5] Dey, B., Halder, S., Khalil, K., Lorusso, G., Severi, J., Leray, P., and Bayoumi, M. A., “Sem image denoising with unsupervised machine learning for better defect inspection and metrology,” in [*Metrology, Inspection, and Process Control for Semiconductor Manufacturing XXXV*], **11611**, 245–254, SPIE (2021).
- [6] Zidan, M., Dey, B., De Simone, D., Severi, J., Charley, A.-L., Halder, S., Leray, P., De Gendt, S., and Lorusso, G. F., “Extraction of roughness measurements from thin resists with low signal-to-noise-ratio (snr) sem images by applying deep learning denoiser,” in [*International Conference on Extreme Ultraviolet Lithography 2022*], **12292**, 137–149, SPIE (2022).
- [7] Chang, C.-Y., Chang, J.-W., and Der Jeng, M., “An unsupervised self-organizing neural network for automatic semiconductor wafer defect inspection,” in [*Proceedings of the 2005 IEEE International Conference on Robotics and Automation*], 3000–3005, IEEE (2005).
- [8] Dey, B., Goswami, D., Halder, S., Khalil, K., Leray, P., and Bayoumi, M. A., “Deep learning-based defect classification and detection in sem images,” in [*Metrology, Inspection, and Process Control XXXVI*], PC120530Y, SPIE (2022).
- [9] Dey, B., Dehaerne, E., and Halder, S., “Towards improving challenging stochastic defect detection in sem images based on improved yolov5,” in [*Photomask Technology 2022*], **12293**, 28–37, SPIE (2022).
- [10] Dehaerne, E., Dey, B., Esfandiari, H., Verstraete, L., Suh, H. S., Halder, S., and De Gendt, S., “Yolov8 for defect inspection of hexagonal directed self-assembly patterns: a data-centric approach,” in [*38th European Mask and Lithography Conference (EMLC 2023)*], **12802**, 286–299, SPIE (2023).
- [11] Cheon, S., Lee, H., Kim, C. O., and Lee, S. H., “Convolutional neural network for wafer surface defect classification and the detection of unknown defect class,” *IEEE Transactions on Semiconductor Manufacturing* **32**(2), 163–170 (2019).
- [12] Rahman, S. S. M. M., Salomon, M., and DembÉLÉ, S., “Noise analysis to guide denoising of scanning electron microscopy images,” in [*2023 9th International Conference on Control, Decision and Information Technologies (CoDIT)*], 1559–1564, IEEE (2023).
- [13] Tětková, L. and Hansen, L. K., “Robustness of visual explanations to common data augmentation methods,” in [*Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*], 3714–3719 (2023).
- [14] Buslaev, A., Iglovikov, V. I., Khvedchenya, E., Parinov, A., Druzhinin, M., and Kalinin, A. A., “Albumen-tations: fast and flexible image augmentations,” *Information* **11**(2), 125 (2020).
- [15] DeVries, T. and Taylor, G. W., “Improved regularization of convolutional neural networks with cutout,” *arXiv preprint arXiv:1708.04552* (2017).
- [16] Zhang, J., Lei, J., Xie, W., Fang, Z., Li, Y., and Du, Q., “Superyolo: Super resolution assisted object detection in multimodal remote sensing imagery,” *IEEE Transactions on Geoscience and Remote Sensing* **61**, 1–15 (2023).
- [17] team, R., “YOLO-NAS by Deci Achieves State-of-the-Art Performance on Object Detection Using Neural Architecture Search,” (2023). Accessed: May 12, 2023.
- [18] Elsken, T., Metzen, J. H., and Hutter, F., “Neural architecture search: A survey,” *Journal of Machine Learning Research* **20**(55), 1–21 (2019).
- [19] Chu, X., Li, L., and Zhang, B., “Make repvgg greater again: A quantization-aware approach,” in [*Proceedings of the AAAI Conference on Artificial Intelligence*], **38**(10), 11624–11632 (2024).
- [20] Terven, J., Córdova-Esparza, D.-M., and Romero-González, J.-A., “A comprehensive review of yolo architectures in computer vision: From yolov1 to yolov8 and yolo-nas,” *Machine Learning and Knowledge Extraction* **5**(4), 1680–1716 (2023).

- [21] team, R., “YOLO-NAS by Deci Achieves State-of-the-Art Performance on Object Detection Using Neural Architecture Search.” <https://deci.ai/blog/yolo-nas-object-detection-foundation-model/> (2023). Accessed: May 12, 2023.