ASME Journal of Manufacturing Science and Engineering
Online journal at:
https://asmedigitalcollection.asme.org/manufacturingscience

Changheon Han
School of Mechanical Engineering,
Purdue University,
585 Purdue Mall,
West Lafayette, IN 47907
e-mail: han711@purdue.edu

Heebum Chun
J. Mike Walker '66 Department of Mechanical Engineering,
Texas A&M University,
3123 TAMU,
College Station, TX 77843
e-mail: hchun@tamu.edu

Jiho Lee
School of Mechanical Engineering,
Purdue University,
585 Purdue Mall,
West Lafayette, IN 47907
e-mail: lee4503@purdue.edu

Fengfeng Zhou
School of Mechanical Engineering,
Purdue University,
585 Purdue Mall,
West Lafayette, IN 47907
e-mail: zhou966@purdue.edu

Huitaek Yun
Mem. ASME
Department of Mechanical Engineering,
Korea Advanced Institute of Science and Technology (KAIST),
291 Daehak-ro, Yuseong-gu,
Daejeon 34141, South Korea
e-mail: htyun@kaist.ac.kr

ChaBum Lee
Mem. ASME
J. Mike Walker '66 Department of Mechanical Engineering,
Texas A&M University,
3123 TAMU,
College Station, TX 77843
e-mail: cblee@tamu.edu

Martin B.G. Jun[1]
Mem. ASME
School of Mechanical Engineering,
Purdue University,
585 Purdue Mall,
West Lafayette, IN 47907
e-mail: mbgjun@purdue.edu

Hybrid Semiconductor Wafer Inspection Framework via Autonomous Data Annotation

In smart manufacturing, semiconductors play an indispensable role in collecting, processing, and analyzing data, ultimately enabling more agile and productive operations. Given the foundational importance of wafers, the purity of a wafer is essential to maintain the integrity of the overall semiconductor fabrication. This study proposes a novel automated visual inspection (AVI) framework for scrutinizing semiconductor wafers from scratch, capable of identifying defective wafers and pinpointing the location of defects through autonomous data annotation. Initially, this proposed methodology leveraged a texture analysis method known as gray-level co-occurrence matrix (GLCM) that categorized wafer images—captured via a stroboscopic imaging system—into distinct scenarios for high- and low-resolution wafer images. GLCM approaches further allowed for a complete separation of low-resolution wafer images into defective and normal wafer images, as well as the extraction of defect images from defective low-resolution wafer images, which were used for training a convolutional neural network (CNN) model. Consequently, the CNN model excelled in localizing defects on defective low-resolution wafer images, achieving an F1 score—the harmonic mean of precision and recall metrics—exceeding 90.1%. In high-resolution wafer images, a background subtraction technique represented defects as clusters of white points. The quantity of these white points determined the defectiveness and pinpointed locations of defects on high-resolution wafer images. Lastly, the CNN implementation further enhanced performance, robustness, and consistency irrespective of variations in the ratio of white point clusters. This technique demonstrated accuracy in localizing defects on high-resolution wafer images, yielding an F1 score greater than 99.3%. [DOI: 10.1115/1.4065276]

Keywords: semiconductor wafer inspection, automated visual inspection, autonomous data annotation, defect localization, metrology, semiconductor manufacturing, sensing, monitoring, and diagnostics

[1]Corresponding author.

Manuscript received September 24, 2023; final manuscript received March 29, 2024; published online May 9, 2024. Assoc. Editor: Cheryl Xu.

1 Introduction

In the landscape of smart manufacturing, fueled by the Industrial Internet of Things (IIoT) and cutting-edge data processing technologies, semiconductors are vital for leveraging data to enhance productivity, quality, and diversity of production [1]. Furthermore, a wide range of contemporary products—including smart televisions, smartphones, and automobiles—consist of semiconductors for data transfer, reception, and computation. The noteworthy growth in the semiconductor sector has a ripple effect on multiple other manufacturing domains. For example, a shortage in semiconductor supply has been observed to disrupt automobile production and inflate prices. Hence, securing high productivity and quality of semiconductor manufacturing is pivotal for proliferating smart manufacturing and sustaining various manufacturing sectors.

Serving as the foundation for semiconductor production, the purity of a wafer must be ensured to preserve the integrity of the overall manufacturing process. Traditional wafer inspections have often relied on manual visual examinations conducted by human experts [2,3]. While this approach is versatile, simple, efficient, high-throughput, and contactless, it carries inherent vulnerabilities on human factors such as inconsistencies due to an inspector's fluency, complicity and repeatability of a task, and product yield rate [4–7]. Given the extreme sensitivity of semiconductor production where even minor defects can compromise the entire process, automated visual inspection (AVI) has been implemented and aimed to provide a consistent, reliable, and cost-effective inspection method by reducing the dependence on human factors.

Although AVI has mitigated vulnerabilities by human factors, past AVI approaches cannot address current manufacturing requirements of diverse products in small batches led by varied customers' demands since they have focused on a certain part, task, and production line [8]. For example, current automobile manufacturers produce not only internal combustion vehicles but also electric vehicles. In this circumstance, the advancement of cutting-edge computer technology like computer vision and deep learning is propelling progress on new AVI methods. However, computer vision–based techniques necessitate feature engineering and intricate structures to cater to various inspection contexts, while deep learning offers autonomous feature extraction and adaptable predictions in complex and varied scenarios [9,10]. Nevertheless, supervised learning demands annotated data sets [11], and unsupervised learning also requires manual adjustment of thresholds and cannot provide satisfying performance [12,13]. For instance, a specific reconstruction error is required to distinguish two classes when using autoencoder (AE). Consequently, it is necessary to develop an AVI framework minimizing loads for feature engineering, data set generation, fine-tuning, and programming with acceptable performance for a new task.

This research explored a hybrid AVI framework tailored for two distinct wafer inspection scenarios—high- and low-resolution semiconductor wafer images—to identify defective wafers and localize the position of a defect without human intervention from scratch. These are the main contributions of this study.

- A classification method for distinguishing between high- and low-resolution wafer images was explored, leveraging a texture-awareness algorithm.
- For each type of wafer, a method for identifying defectiveness was developed using computer vision techniques.
- A computer vision–based technique for autonomously generating and annotating data sets was investigated for training a convolutional neural network (CNN) model.
- The CNN model was formulated to pinpoint defects in defective wafers.

Integrating the advantages of computer vision and deep learning techniques, the proposed hybrid AVI framework could classify the types of wafers and localize defective regions within a 32-pixel error range autonomously with unclassified and unannotated data.

2 Related Works

Raman spectroscopy is one of the methods for semiconductor inspection scenarios. However, it may lead to unexpected distortion of observation or deformation of material due to strong local heating tightly focused laser beam generates [14]. Among non-destructive testing (NDT) methods, the resolution of ultrasonic techniques (UT) relies on half the wavelength of a sound source. Therefore, the application of UT is limited to a case over a microscale [15]. In contrast, a microscopic system can observe a wide range of scales depending on a beam source and provides versatility, adaptability, and non-contact nature in vision inspection scenarios. While the semiconductor industry has embraced vision inspection methods for their simplicity, versatility, adaptability, cost-effectiveness, and non-contact nature [5,7], such methods are often limited by the expertise and inconsistencies of human inspectors [4–6]. To address these limitations, AVI systems have been implemented.

An AVI system consists of image acquisition and image analysis parts. In terms of image acquisition, a stroboscopic imaging system holds remarkable potential for applications in dynamically measuring rapidly moving objects [16]. This optical approach increases the imaging contrast of the featured area and reduces the image contrast of the background area. Its versatile application extends across diverse domains, including wafer inspection, aerospace, biomedicine, and even machining processes or monitoring procedures [17–21].

When it comes to image analysis, computer vision techniques have gained prominence. Huang and Pan [3] provide a comprehensive overview of semiconductor inspection algorithms and divide them into four types: projection methods, filter-based methods, learning-based methods, and hybrid approaches. First, projection methods focus on identifying relationships among samples and mapping them into a feature space that better represents their inherent characteristics. Techniques such as principal component analysis (PCA), linear discriminant analysis (LDA), and independent component analysis (ICA) are commonly used for optimizing information and reducing noise. Filter-based approaches, especially frequency-filtering techniques like Fourier transform [22] and wavelet transform [23], convert signals from a spatial domain into a frequency domain. These are useful for reducing noise, amplifying key features, and investigating spatial representation from images.

Although computer vision has significantly advanced AVI capabilities, real-world complexities often demand more sophisticated algorithms and computational resources. Learning-based approaches, including pattern recognition, support vector machine (SVM), and deep learning, circumvent those drawbacks. Specifically, deep learning extracts features of a data set and establishes a rule for inspection automatically. Therefore, deep learning algorithms have been applied in various AVI scenarios [24–26] and improved the adaptability, versatility, efficiency, and accuracy of inspection.

Kulkarni and Xu [27] introduced a supervised deep learning–based visual inspection method, implementing ResNet, to secure the integrity of a circuit board—a fundamental element of electric devices—with limited data. The semiconductor inspection domain also has implemented deep learning approaches such as a supervised learning method of CNN. Nakazawa and Kulkarni [28] generated wafer map defect patterns synthetically and classified them into 22 classes. Phua and Theng [29] captured defective wafer images by a scanning electron microscopic system and identified the defects automatically. While Wen et al. [30] identified two types of wafers with stain and scratch defects by deploying a CNN-based deep learning method, they pointed out the challenges in generating training image sets for CNN model training.

Although CNN is a useful approach for analyzing semiconductor wafer images, it necessitates annotated data sets to train a model [11]. In contrast, AE is an unsupervised learning method and does not need labeled data sets. It abstracts distinguished data

attributes by compressing and depressing data through encoder and decoder layers connected via a latent space [31]. Erhan et al. [32] demonstrated that unsupervised pre-training can help deep learning performance. Moreover, Feng et al. [33] introduced a self-taught learning method based on AE with unlabeled data and explained that it enhanced the performance of a classifier. For semiconductor wafer inspection, several studies also utilized AE techniques to compensate for the shortcomings of supervised learning. An investigation adopted a denoise autoencoder (DAE) and calculated reconstruction errors as thresholds to classify normal and defective wafers [34]. Nevertheless, this research indicated that fine-tuning or additional techniques were necessary to get acceptable performance, and its scope was limited to classifying normal and defective wafers, not detecting defects on a wafer. Yu et al. [35] proposed an approach that discriminated wafer map patterns using a stacked denoising autoencoder with convolutional layers and constructed an SVM classifier based on the recognized features of the patterns. They explained the performance of the models by experimental results from a simulation data set and real-world wafer map data set (WM-811K), then provided the comparison by different recognizers. Although the study demonstrated performance ranging from 81.25% to 97.79% of recognition rates, it used pre-classified data sets to build the models.

Hybrid methods have been explored due to the pros and cons of every algorithm and the multifaceted requirements of real-world inspections. Hybrid methods implement multiple feature extractors or combine them with a classifier like SVM. For instance, Wang [36] deployed spatial statistics, kernel-based eigendecomposition, and support vector clustering techniques to detect and divide defect clusters. Schlosser et al. [37] devised a deep learning–based hybrid multi-stage AVI for a semiconductor wafer. They combined computer vision and deep learning techniques and achieved up to an F1 score of 99.5% for detecting faults on wafers.

Although those hybrid methods address the shortcomings of each approach by integrating multiple techniques, they still possess limitations that require pre-organized data sets or extensive feature engineering processes to generate data sets or define a threshold for decision-making. Generally, curating field data requires a lot of manpower since real-world data is highly noisy and is generated in huge quantities. There are available prepared data, such as WM-811k, and many studies have utilized it. However, the semiconductor industry is rapidly developing new products that demand additional organized data sets for a physical inspection scenario. It results in a huge consumption of resources for preparing data sets for a new inspection case.

In response to the described challenges, establishing an inspection framework, including autonomous data annotation, can reduce costs and resources when inspecting a new product. Accordingly, this study proposes a wafer inspection method consisting of a multi-stage pipeline designed to automate entire inspection processes—generate and annotate data sets, train deep learning models, and execute inspection—by eliminating manual labeling, extensive feature engineering, and fine-tuning processes.

3 Experimental Setup

This study utilized a stroboscopic imaging system to address the need for expediting scanning times when capturing wafer images. This integration enables the real-time capture of expansive surface images of the wafer within a remarkably short period. Figure 1(*a*) illustrates the schematic and experimental setup of the stroboscopic imaging system incorporated into the wafer inspection process.

Unlike conventional raster scans that solely use a linear axis for wafer surface scanning, the inclusion of a rotary axis expedites the wafer scanning time. Images can be sequentially captured while the spindle is rotating. While the spindle is rotating, a stroboscopic effect allows for obtaining a sequence of images from distinct locations. This is achieved by adjusting the frequency of the light source or exposure time and the framerate of the imaging sensor, both of which can be synchronized with the rotational speed of the spindle using an encoder signal. The spindle speed and framerate of the imaging sensor were maintained at low levels—around 0.2 rpm and 2.0 fps, respectively.

Once the spindle starts to rotate, the encoder begins to generate binary signals that can be read using a pulse counter. The imaging sensor can then be triggered to capture a sequence of images under a certain framerate with a predetermined exposure time, utilizing the phase-shifted encoder signal. The microscope system can be moved in-line, either from the edge of the wafer toward the center of the rotational axis or in reverse. Therefore, the entire wafer can be scanned by moving the microscope system while the spindle is rotating, as schematically illustrated in Fig. 1(*a*).

As shown in the experimental setup, the wafer scanning configuration comprises one rotational axis and three linear axes. An optical microscope, mounted on the *Z*-axis stage, ensures precise focusing of the target object, while the *X* and *Y* linear stages scan and align the wafer. The wafer can be firmly held in place by a vacuum chuck affixed to the artifact of the spindle axis.

The inspection imaging system features a high-resolution microscopy encompassing a 20-megapixel color CMOS camera, interchangeable objective lenses (100× and 250× magnification), and a 4.5× microscope zoom lens. This high-resolution imaging system excels in capturing particles as small as a few micrometers on a ϕ4-inch wafer. The use of continuous white light illumination facilitated the acquisition of brightfield images of the surface of the wafer.

When preparing wafer samples, a blank wafer without any particles was assessed to ensure the integrity of the wafer. As a result, the tested blank wafer itself appeared clean without any presence of particles or defects. For the preparation of defective wafers, particles as small as 1.0 μm were distributed across the entire wafer surface. Subsequently, compressed air was applied to the surface, ensuring that only small particles securely adhered to the wafer surface. Additionally, the experiment was conducted under a laminar flow hood that provides clean air through a high-efficiency particulate air (HEPA) filter, capable of filtering dust or particles down to 0.5 μm, installed in the hood to prevent any unwanted particle or dust inclusion.

Two sets of objective lenses were used for capturing images, and the images were obtained under identical base quality wafers. Those taken with the 100× lens exhibited distorted backgrounds and were named "low-resolution wafer images" in this study. Conversely, images captured with the alternative lens displayed relatively smoother surfaces and were referred to as "high-resolution wafer images." All acquired images underwent normalization and resizing to a resolution of 1536 × 1024 pixels. Figure 1(*b*) presents a selection of these pre-processed images, wherein the region of interest (ROI) is demarcated by the inner perimeter of a red circle.

4 Methodology

This proposed wafer inspection framework consists of four stages: classification of high- and low-resolution wafer images, identification of defective wafer images, extraction of data set with autonomous data annotation, and localization of defects. Figure 2 provides an overview of the entire workflow. This investigation was done with a system operated on Ubuntu 20.04.5 LTS with Intel i7-11700K CPU, 64 GB RAM, and Nvidia RTX A5000 GPU. PyTorch 2.0.1 with GPU acceleration [38] was used for training and executing the deep learning model. To initiate the investigation from the ground up, two unclassified sets of images—train and test sets—were generated by combining the collected high- and low-resolution wafer images, both with and without defects. The two sets were used independently for training and evaluation processes. In other words, explored results and models during the classification of the train set—normal and defective wafers, then normal and defective

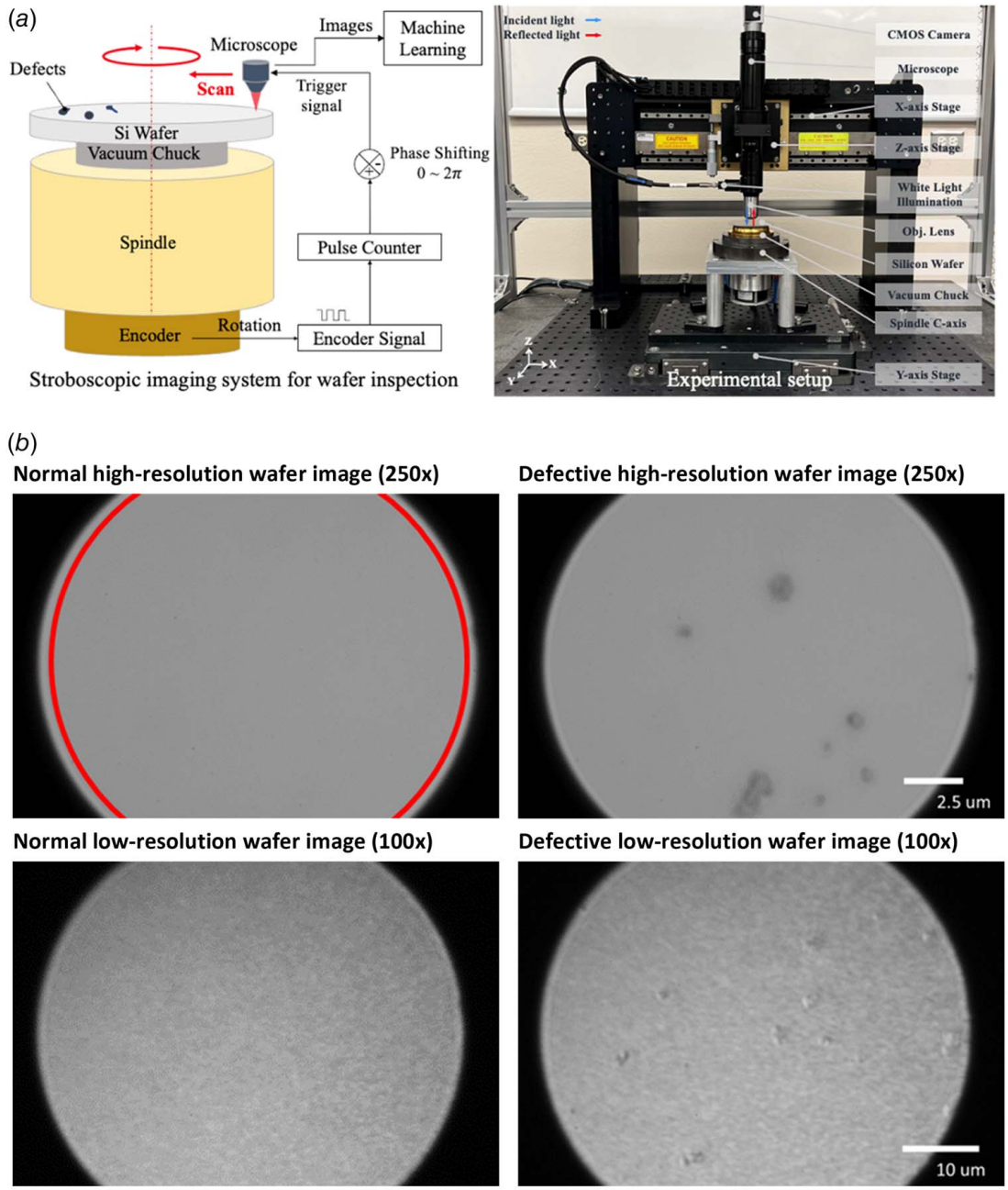

Fig. 1 (*a*) Experimental setup of stroboscopic wafer inspection and (*b*) example of the wafer images

surfaces—were evaluated on the test set. The detailed configuration of the image sets is described in Table 1.

4.1 Wafer Classification. A histogram is effective for abstracting and analyzing features of an image, like the intensity and distribution of brightness [39,40]. Figure 3 depicts the histogram distribution of whole wafer images, displaying the number of their pixels according to intensity. The distribution revealed two distinct feature clusters. Given that low-resolution wafer images were expected to exhibit greater variance compared to high-resolution wafer images, the dotted and solid lines in the histogram were interpreted as representing high- and low-resolution wafer images, respectively. This result indicated that textures represented different features on each wafer image type. Although the histogram analysis provided a brief description of each group, it was insufficient to generalize the features of each group and separate them thoroughly, and further textures of wafers were analyzed.

Gray-level co-occurrence matrix (GLCM) [41] is a statistical approach representing features and texture distributions through a two-dimensional histogram of paired pixel values within a specific spatial range. GLCM has found extensive applications in various fields, including detection, segmentation, and classification tasks such as landscape and geomorphological pattern recognition [42–44], medical image analysis [45,46], and fabric defect detection [47]. Moreover, Yu et al. [48,49] presented its applicability in detecting defects on a wafer map that represents defect patterns of a semiconductor wafer graphically [50]. Utilizing GLCM, a range of statistical metrics can be derived to adequately characterize texture features adapted to each specific scenario.

Based on a GLCM g and its (i, j)th component g_{ij}, several statistics were proposed. This study employed three specific GLCM statistics—homogeneity, energy, and correlation [51] with a GLCM parameter distance of 16 and angles at 0, 45, 90, and 135 deg—tailored to different use cases. Homogeneity quantifies image uniformity, assigning higher values when there are smaller differences

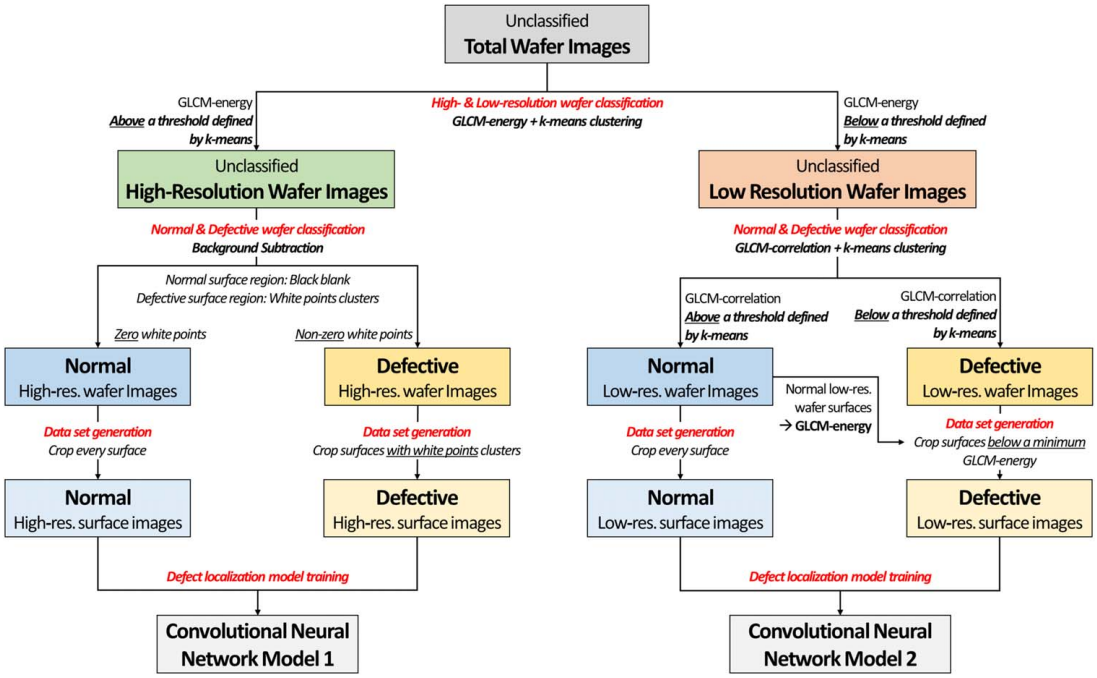

Fig. 2 Pipelines of the developed framework

Table 1 Configuration of the data set

Unclassified train set			Unclassified test set		
Low-resolution wafer images	Defective	400 images	Low-resolution wafer images	Defective	200 images
	Normal	400 images		Normal	200 images
High-resolution wafer images	Defective	400 images	High-resolution wafer images	Defective	200 images
	Normal	400 images		Normal	200 images

in gray tones between paired elements, and can be expressed as Eq. (1).

$$\text{Homogeneity} = \sum_i \sum_j \frac{1}{1+(i-j)^2} g_{ij} \tag{1}$$

The correlation feature assesses the degree of linear dependencies between gray tones in the image and is defined as follows:

$$\text{Correlation} = \frac{\sum_i \sum_j (ij) g_{ij} - \mu_x \mu_y}{\sigma_x \sigma_y} \tag{2}$$

where μ and σ are the mean and standard deviation of g, and x and y indicate the row and column directions of a GLCM. Energy gauges textural uniformity through the repetition of pixel pairs, serving as an indicator for irregularities in textures and can be calculated as Eq. (3).

$$\text{Energy} = \sum_i \sum_j g_{ij}^2 \tag{3}$$

Figure 4 describes a GLCM homogeneity analysis applied to the unclassified wafer images. Based on the partitioning boundary by k-means clustering, the wafer images were effectively and

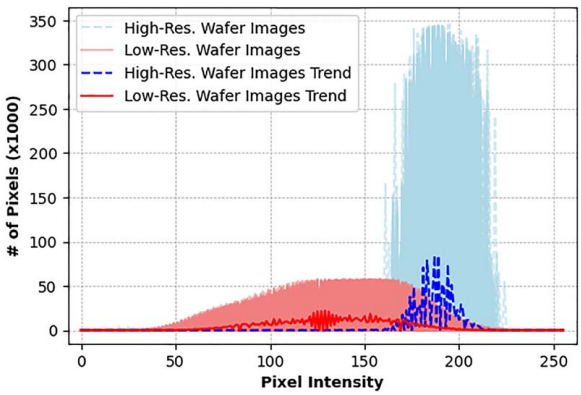

Fig. 3 Histograms of the unclassified wafer images

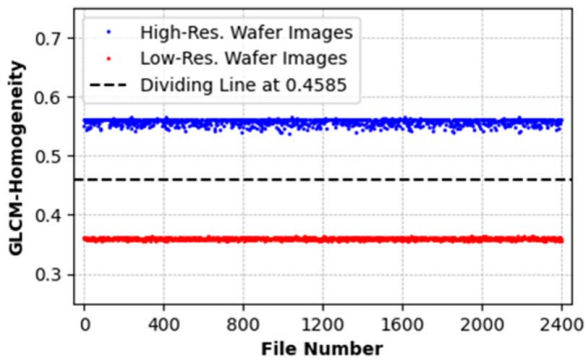

Fig. 4 GLCM homogeneity analysis on the unclassified wafer images

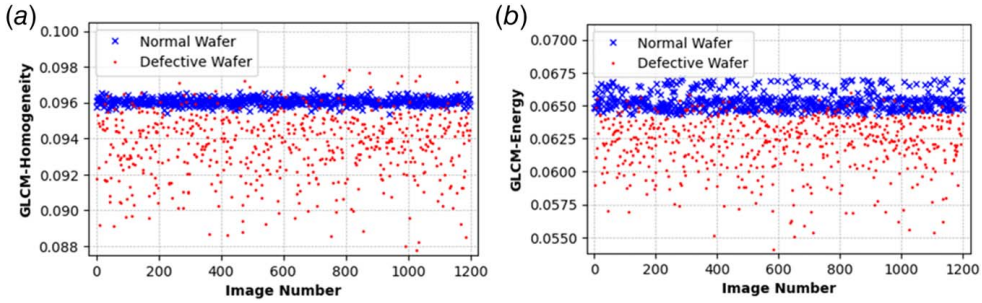

Fig. 5 GLCM analysis on the classified high-resolution wafer images: (*a*) GLCM homogeneity and (*b*) GLCM energy

automatically segregated into two distinct groups. Images presumed to be high-resolution wafer images were positioned above the dividing line, while the remaining images fell below it. Since GLCM homogeneity indicates image uniformity, this analysis validated the earlier hypothesis formed during the histogram examination and successfully differentiated the unclassified wafer images. In short, unclassified wafers were divided into high- and low-resolution wafer images as the automatically defined threshold by k-means clustering under the GLCM homogeneity analysis.

4.2 Defective and Normal Wafer Classification. The classifications of defective and normal wafers were conducted under two separate scenarios: high- and low-resolution wafer images. This chapter outlines the classification processes of defective and normal wafers for both scenarios. First, normal high-resolution wafer images generally displayed higher GLCM homogeneity and energy values than defective high-resolution wafer images, as illustrated in Fig. 5. While GLCM analysis provided some insights into the characteristics of high-resolution wafer images, it fell short of fully distinguishing between normal and defective high-resolution wafer images as some defective wafers exhibited atypical features represented as the red points overlapped on the blue points. Thus, a computer vision algorithm—background subtraction—was suggested to address this challenge.

The wafer image sets had similar repetitive patterns [3] and could be likened to video scenes with consistent backgrounds. In this context, defects would be treated as foreground objects, and a background subtractor was employed to identify the presence of defects on a high-resolution wafer image. The background subtraction methods are popular in applications such as video surveillance, where the scenes remain largely consistent except for a few moving objects [52]. Typically, these methods initially construct a background model based on a selected number of frames. Then, they identify foreground objects by comparing this established background model with a target frame, subsequently updating the background model during the process of foreground detection [53].

This research opted BackgroundSubtractorMOG2 class in OpenCV [54] owing to its satisfactory performance and computational efficiency [52,55]. This class is based on a Gaussian mixture model [56], which is improved by applying constantly updated parameters and simultaneously selecting components for each pixel [57]. Among the parameters of the class, "*setHistory*," which defines a target frame until building a background model, equaled 800, matching the total number of training image sets, while default values were retained for all other parameters. Salt-and-pepper noise was observed in the background-subtracted images. To denoise this, the median blur function from OpenCV was applied with a kernel size of 3. Figure 6 illustrates an example of original, background-subtracted, and denoised images. If a denoised image contained no white points, it was categorized as a normal high-resolution wafer image. Conversely, any image with white points was identified as a defective high-resolution wafer image. Consequently, all high-resolution wafer images in the training and test set were classified automatically and thoroughly via background subtraction and denoising techniques.

In contrast, the use of background subtraction and denoising techniques proved insufficient for classifying normal and defective low-resolution wafer images due to the presence of highly distorted backgrounds. As shown in Fig. 7, these distorted backgrounds generated noise, making it challenging to distinguish between normal and defective low-resolution wafer images. GLCM analysis was conducted in conjunction with k-means clustering to investigate the characteristics of low-resolution wafer surface images. Figure 8 demonstrates the outcome: low-resolution wafer images were successfully separated into two groups by the dividing line automatically defined by the k-means clustering analysis. According to GLCM correlation attributes, it is concluded that the group above the line consists of normal wafers, while the group below contains defective wafers. This assumes that normal surfaces would exhibit higher linear dependencies between gray tones in the image than defective surfaces.

4.3 Defects Localization 1—High-Resolution Wafer Image Case. Background subtraction and denoising techniques effectively represented defects on high-resolution wafer images as clusters of white points. Therefore, the ratio of a white point cluster area, WP ratio, is considered a critical parameter in this study. A WP ratio was evaluated by the number of white pixels in a certain size of an

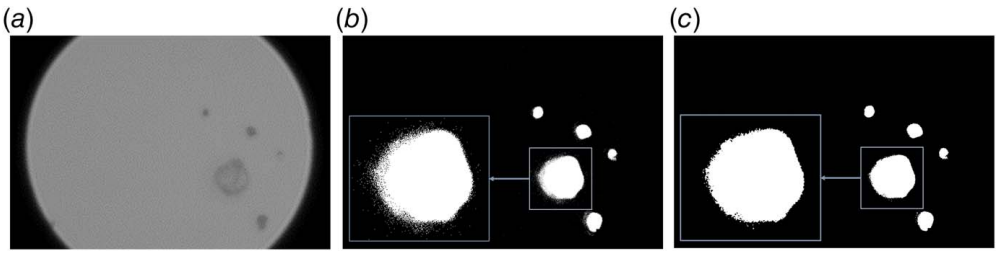

Fig. 6 Example of processed defective high-resolution wafer images: (*a*) defective wafer image—original, (*b*) background subtraction, and (*c*) denoise

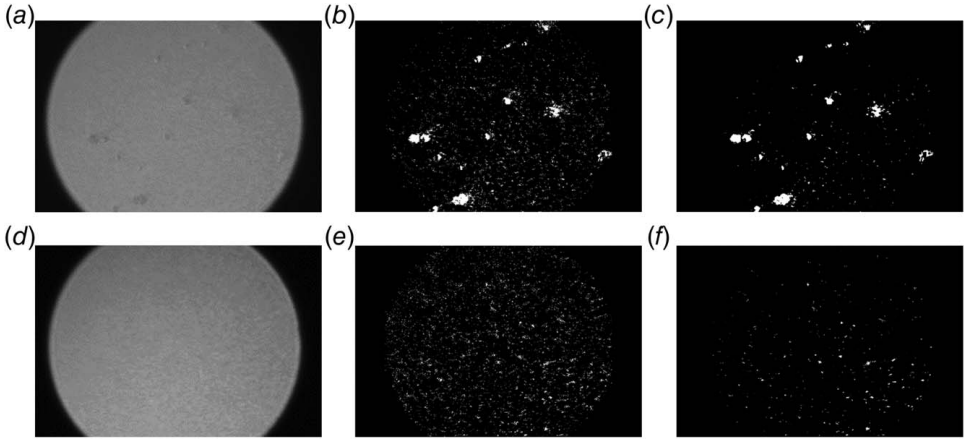

Fig. 7 Examples of processed low-resolution wafer images: (*a*) defective wafer image—original, (*b*) background subtraction, (*c*) denoise, (*d*) normal wafer image—original, (*e*) background subtraction, and (*f*) denoise

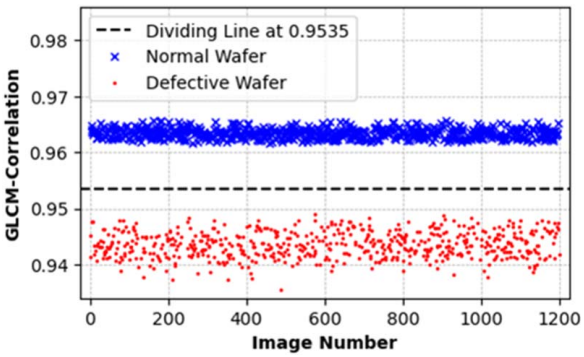

Fig. 8 GLCM analysis on the classified low-resolution wafer images

image (Fig. 9). Considering factors like particle size, computational cost, and detection accuracy, the image size was set as a 32 × 32-pixel square, and the permissible error range was set to 32 pixels accordingly. A prediction was performed with a 32-pixel striding step to avoid overlaps. As illustrated in Fig. 10, the degree of detection was varied with WP ratios. Thus, the performance of the defect localization method based on a white points cluster was evaluated by varying WP ratios.

For evaluation, a confusion matrix was defined as Fig. 11. In this context, the terms "Positive" and "Negative" refer to the prediction of the defect and normal classes, respectively. For instance, the area was estimated as "Positive" when the WP ratio of an area exceeded a determined criterion. The evaluation metrics were precision, recall, accuracy, and F1 score, which is calculated in Table 2. As the ratio between normal and defect regions was imbalanced, the precision and accuracy metrics could not evaluate the performance of a model adequately. Therefore, the recall and F1 score metrics were implemented to properly evaluate the performance of predicting defect regions as well as normal regions.

Fig. 12 shows that the method achieved a maximum F1 score of 99.4% with a WP ratio of 3%. At this ratio, precision, recall, and accuracy were 99.7%, 99.2%, and 100.0%, respectively, without a manual data labeling process. It is important to note, however, that the ideal ratio for achieving the best evaluation metrics may differ depending on various factors and environmental conditions. Therefore, CNN was implemented to devise a robust method independent of WP ratios for defect localization.

A range of WP ratios—5%, 15%, 25%, 35%, 50%, 65%, 75%, 85%, and 100%—were applied to acquire defect images. These images were extracted as 32 × 32-pixel square images from defective high-resolution wafer images with a 16-pixel stride. If a WP ratio in a 32 × 32-pixel window exceeded a specified number while striding an image, the scene in the window was saved as a defect image for training. Normal surface images were gathered as the same size as images from normal high-resolution wafer images by striding 32 pixels. The number of images in each class was configured to be taken over by the balance. Table 3 shows the configurations of generated data sets with autonomous data annotation. CNN models were trained for each data set by following the neural network configuration delineated in Fig. 13. Throughout the training process, the cross-entropy loss function and Adam optimizer with a learning rate ranging from 10^{-4} to 10^{-5} were applied.

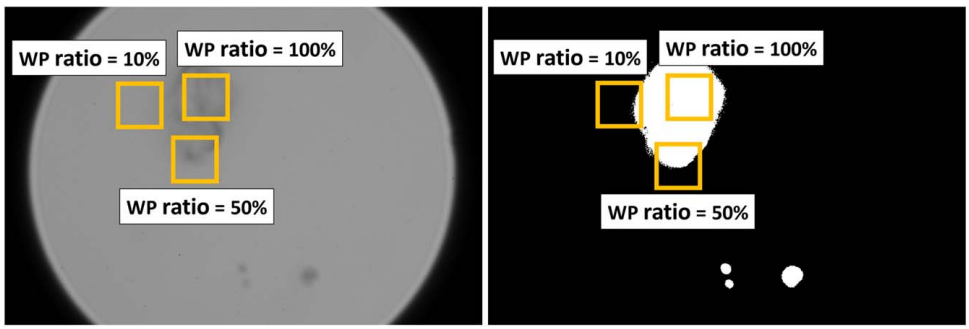

Fig. 9 Example of WP ratios by regions of a defective high-resolution wafer image

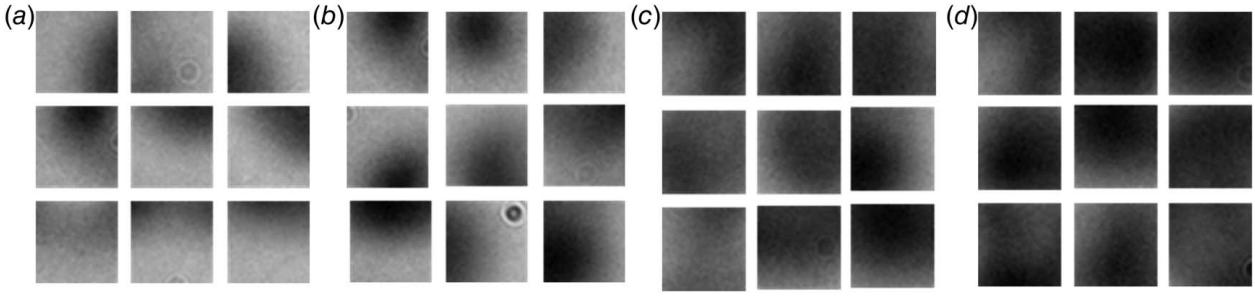

Fig. 10 Defect samples of high-resolution wafer images by WP ratios: (*a*) ratio: 5%, (*b*) ratio: 50%, (*c*) ratio: 75%, and (*d*) ratio: 100%

The batch size was 16, and 20% of the data sets were used for validation. Furthermore, the data sets were augmented by the "AutoAugment" PyTorch library with the "IMAGENET" policy providing an image augmentation strategy automatically [58].

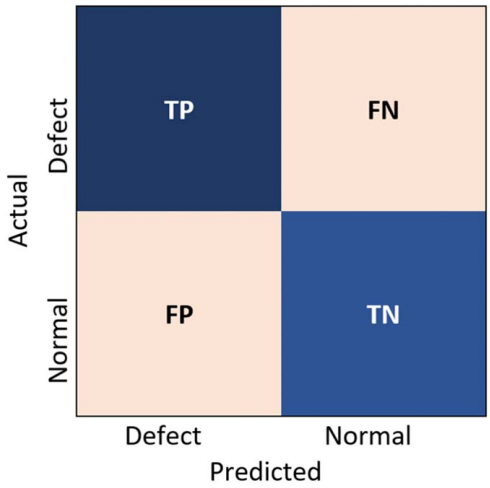

Fig. 11 Confusion matrix for defects localization

Table 2 Metrics and formulae in evaluation

Metric	Formula
Precision	$TP/(TP+FP)$
Recall	$TP/(TP+FN)$
Accuracy	$(TP+TN)/(TP+TN+FP+FN)$
F1 score	$2/(\text{Recall}^{-1}+\text{Precision}^{-1})$

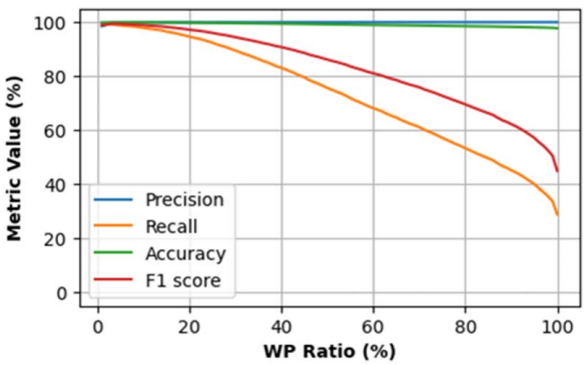

Fig. 12 Defects localization results by WP ratios on defective high-resolution wafer images

Its performance was evaluated by the same method for the white point cluster-based defects localization method after the training. The models classified a surface by striding 32 pixels with a 32 × 32-pixel window, and the error range was also 32 pixels. As in Fig. 14, the CNN models yielded consistent and solid performance, keeping all metric values over 97.2% throughout every data set. The best F1 score of 99.7% with a recall of 99.5%, an accuracy of 100.0%, and a precision of 99.8% was observed on the data set generated by a WP ratio of 85%. Notably, the CNN approach exhibited a significant improvement at a WP ratio of 100%: a recall of 98.6%, an accuracy of 99.9%, a precision of 99.9%, and an F1 score of 99.3%. In conclusion, this result emphasizes that it is possible to build an accurate CNN model regardless of a WP ratio as long as the white point cluster method guarantees defect detection with high precision and accuracy to get a proper defect image set for CNN training.

4.4 Defects Localization 2—Low-Resolution Wafer Image Case. In contrast to the high-resolution wafer image case, a background subtraction approach with denoising was less effective at highlighting defects as groups of white points in low-resolution wafer images. As Fig. 15 demonstrates, the white point cluster-based defects localization approach in the low-resolution wafer image case achieved a maximum F1 score of 77.0% with a WP ratio of 8%. At the same ratio, precision, recall, and accuracy were 79.1%, 75.1%, and 99.3% respectively. The distorted surfaces of low-resolution wafer images contributed to the worsened precision and recall values compared to the high-resolution wafer image scenario. Hence, CNN models were developed to improve defect localization performance in low-resolution wafer images. Before training, however, it was necessary to build a data set comprising two groups. Since the white point cluster method yielded low precision inappropriate for extracting defect images, detailed texture analysis on low-resolution wafer images was investigated by GLCM energy to figure out the way with higher precision to extract and annotate data sets automatically.

GLCM energy was calculated for every 32 × 32-pixel size surface with a stride of 32 pixels for both normal and defective low-resolution wafer images. As depicted in Fig. 16, the calculation result revealed a general trend that normal low-resolution wafer image surfaces tended to have higher GLCM energy values than those of defective low-resolution wafer image surfaces. Even though it did not allow for a complete separation of two surfaces, there were instances where the GLCM energy for some parts of defective wafer surfaces was always lower than the lowest value of normal wafer surfaces. To validate an assumption that defective wafer surfaces below the dividing line were defects, defects localization performance using the lowest GLCM energy value (0.0634) observed for normal wafer surfaces was evaluated. Specifically, the surface was estimated to have a defect when its GLCM energy was below the lowest GLCM energy value, and vice versa was assumed as a normal surface.

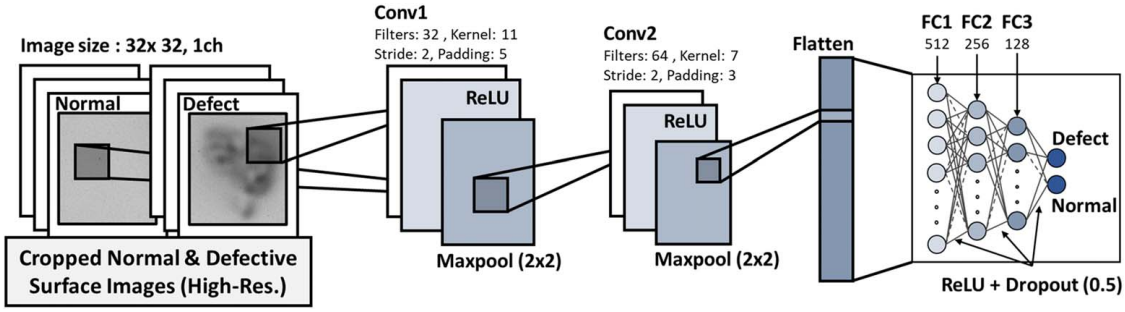

Fig. 13 CNN configuration for defect localization on high-resolution wafer images

Table 3 Configurations of data sets

WP ratio	5%	15%	25%	35%	50%	65%	75%	85%	100%
No. of defect images	66,246	56,242	49,870	44,499	38,268	31,229	27,377	23,081	13,212
No. of normal images	66,969	57,402	51,024	45,709	37,571	31,890	27,638	23,386	13,819

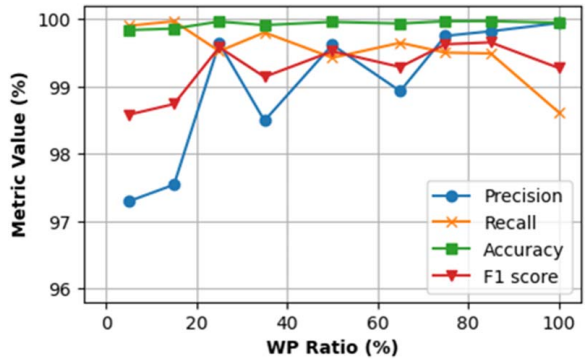

Fig. 14 CNN model-based defects localization results on defective high-resolution wafer images

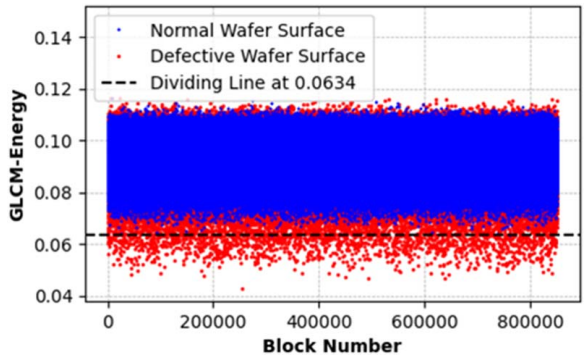

Fig. 16 GLCM analysis on the normal and defective low-resolution wafer image surfaces

This method yielded a recall of 61.0%, an accuracy of 99.4%, a precision of 91.6%, and an F1 score of 73.2%, showing a low false positive rate. It meant that defect images could be obtained with a probability exceeding 91% by the criterion. Moreover, normal surface images could be gathered from classified normal low-resolution wafer image surfaces without errors. Consequently, a defect image set for CNN training was generated by moving a 32 × 32-pixel window with a 16-pixel stride on the train set of defective low-resolution wafer image surfaces. Normal low-resolution wafer image surface set was extracted by the same size of a window with a 32-pixel stride on the train set of normal low-resolution wafer images. A total of 4664 defect images and 5315 normal surface images were acquired and annotated autonomously in this process.

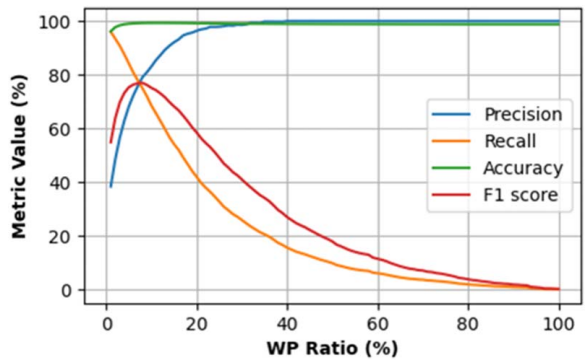

Fig. 15 Defects localization results by WP ratios on defective low-resolution wafer images

Using the train data set, a CNN model was built based on the configuration illustrated in Fig. 17. This model was trained with the Adam optimizer with a learning rate ranging between 10^{-4} and 10^{-5}, the cross-entropy loss function, and a batch size of 16. The developed CNN model showed progress in performance compared to other methods. It yielded a recall of 87.4%, an accuracy of 99.6%, a precision of 83.6%, and an F1 score of 85.4%, respectively. Compared to the former two approaches, the CNN model showed improved performance even on the distorted surface with an automatically labeled data set.

5 Discussion

In the previous section, it was described that GLCM methods classified unclassified images into high- and low-resolution wafer images thoroughly. Moreover, defective low-resolution wafer images were completely identified, and defect images in defective low-resolution wafer images for CNN training were extracted by GLCM analysis. The trained CNN model localized defects on defective low-resolution wafer images with an F1 score of 85.4%. For the high-resolution wafer image scenario, a background subtraction technique was remarkably effective in distinguishing between normal and defective high-resolution wafer images. This

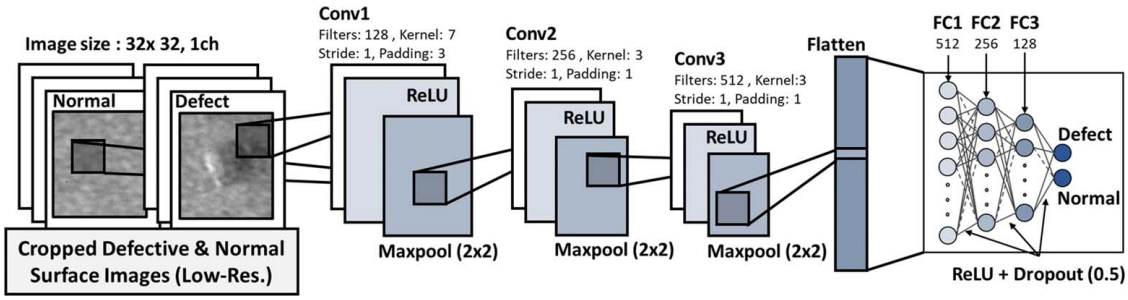

Fig. 17 CNN configuration for defect localization on low-resolution wafer images

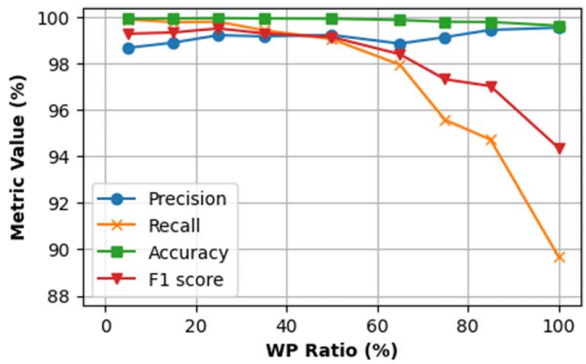

Fig. 18 CNN model performance without image augmentation on high-resolution wafer images

technique represented a defect as a cluster of white points as well and achieved a maximum F1 score of 99.4%. To compensate for the limitation of this method that should explore the optimal WP ratio, CNN models were developed using a train data set generated by referring to WP ratios. The CNN models showed a consistent and robust result with an F1 score over 98.6%, regardless of which WP ratio was used to extract train data set in defect localization on defective high-resolution wafer images. In summary, the developed framework suggests the following steps.

For unclassified wafers,

- *Step 1:* Classification between high- and low-resolution wafer images following a threshold derived by GLCM homogeneity analysis with k-mean clustering.

For high-resolution wafer images,

- *Step 2:* Identification of defective high-resolution wafer images based on the number of white points after subtracting the background.
- *Step 3:* Extraction of train data set on a region of a defective high-resolution wafer image with a WP ratio of 100%.
- *Step 4:* CNN model training and defects localization on defective high-resolution wafer images.

For low-resolution wafer images,

- *Step 5:* Identification of defective low-resolution wafer images based on GLCM correlation analysis with k-means clustering.
- *Step 6:* Determination of a threshold by the minimum GLCM energy value of normal low-resolution wafer image surfaces.
- *Step 7:* Extraction of train data set on a region of a defective low-resolution wafer image showing a lower GLCM energy value than the defined threshold.
- *Step 8:* CNN model training and defects localization on defective low-resolution wafer images

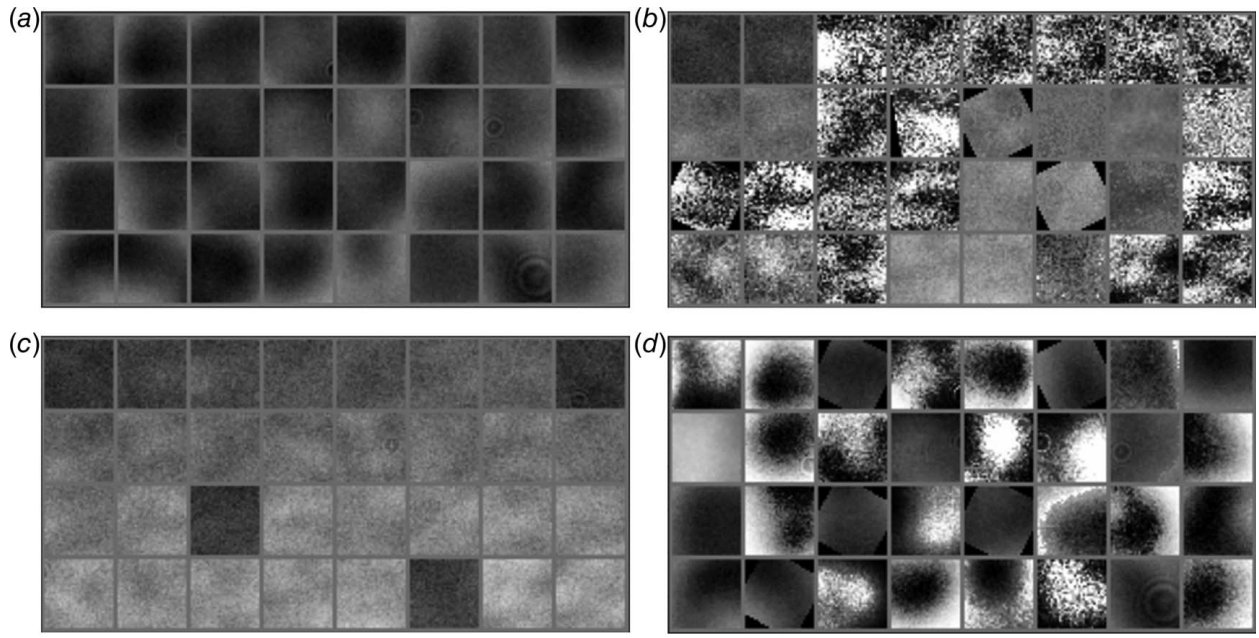

Fig. 19 Original and augmented defect images: (*a*) original defect images of high-resolution wafer images, (*b*) augmented defect images of high-resolution wafer images, (c) original defect images of low-resolution wafer images, and (*d*) augmented defect images of low-resolution wafer images

According to this procedure, the classification of wafers, identification of defectiveness, and localization of defects could be achieved without a manual data labeling process.

The next discussion is about the effectiveness of implementing image augmentation for high- and low-resolution wafer images. The image augmentation process plays a key role in improving the performance of the defect localization CNN model for defective high-resolution wafer images. As described in Fig. 18, the performance of the CNN model trained without the augmentation library for defective high-resolution wafer images showed a steep decline in WP ratios from 65% to 100% for defect localization. When increasing a WP ratio, only defects occupying the large area could be confirmed and extracted (Fig. 11), resulting in many defect images being excluded from the train data set. Therefore, a CNN model trained by a set from a high WP ratio detected only defects occupying a large area in a 32×32 square. In contrast, the image augmentation technique provided diversity in the data set (Figs. 19(*a*) and 19(*b*)) and enabled to generate a superior model with a limited data set when applying a high WP ratio and provided robustness and consistency in the development of models regardless of the WP ratio. However, augmenting images was not effective on low-resolution wafer images, which had more intricate patterns than high-resolution wafer images. Unlike defects on high-resolution wafer images, complicated surface features of low-resolution wafer images resulted in vague boundaries between normal surfaces and defects (Fig. 19(*c*)). The increased degree of complexity in augmented images caused divergence during training (Fig. 19(*d*)).

This study adopted Gradient-Weighted Class Activation Mapping (Grad-CAM) [59], which is based on the concept of explainable artificial intelligence (XAI), to validate the effectiveness of image augmentation for each case. Most deep learning models operate as black boxes, hiding their decision-making processes. This not only impedes understanding of these processes but also undermines the reliability of models [60]. XAI seeks to address these issues by providing greater transparency in decision-making, thereby enhancing the robustness and applicability of deep learning models [61–66]. As a result, this study delves into the interpretability of deep learning models, employing Grad-CAM to visualize activated pixels after the CNN layers, thereby enhancing the robustness and applicability of deep learning models.

Figure 20 presents the heat map of Grad-CAM analysis on surfaces of high-resolution wafer images. The map shows the number of activated pixels, which indicates that the CNN model captured the overall surface features of the original images effectively. As a result, the augmented images with differentiated angles, contrast, and brightness enhanced the training by offering more diverse features. For high-resolution wafer images, image augmentation significantly improved the recall of the model from 89.7% to 98.6% at a WP ratio of 100%. Nonetheless, the augmentation technique was not effective for low-resolution wafer image

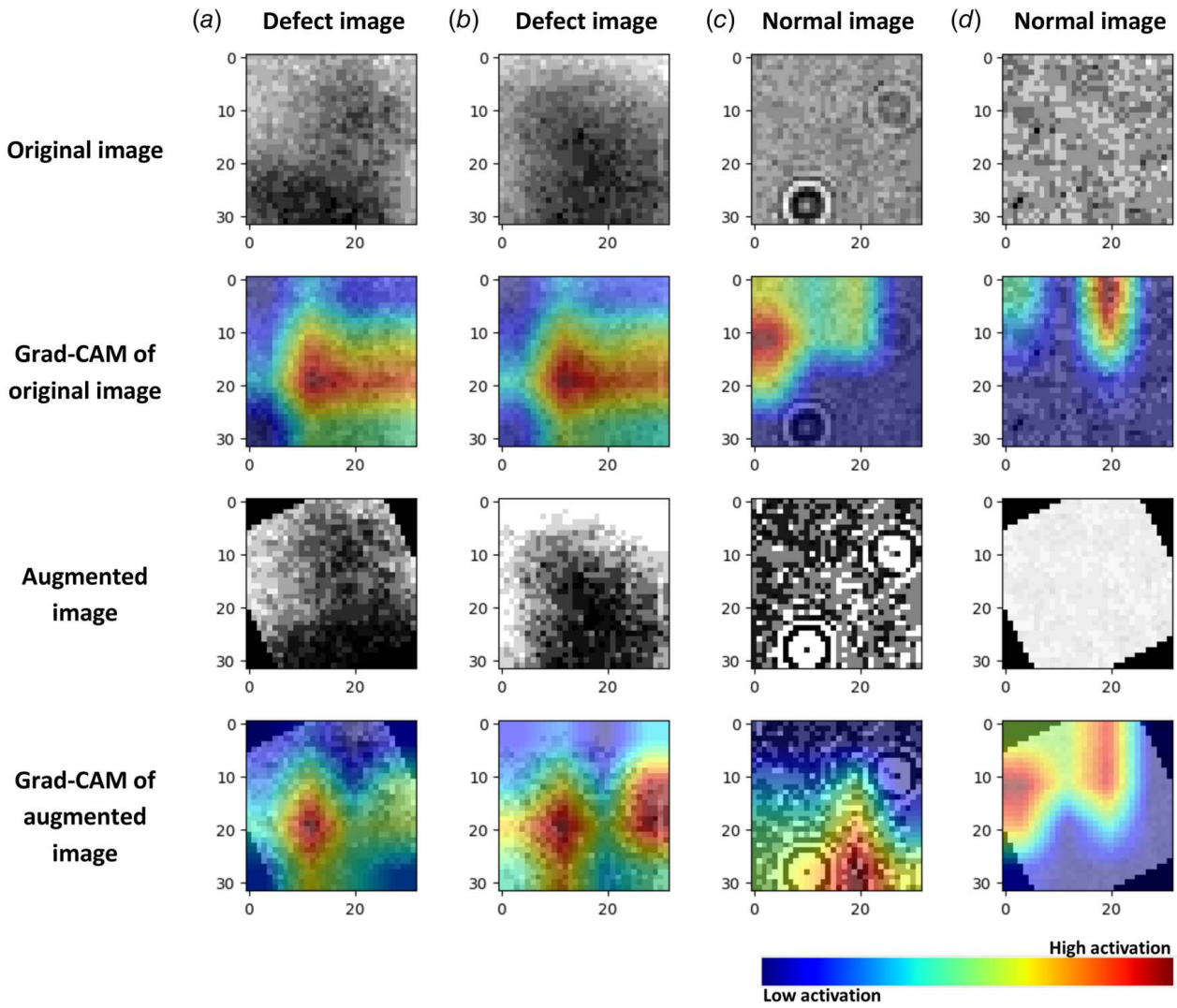

Fig. 20 Grad-CAM heat map of surfaces of high-resolution wafer images: (*a*) defect image, (*b*) defect image, (c) normal image, and (*d*) normal image

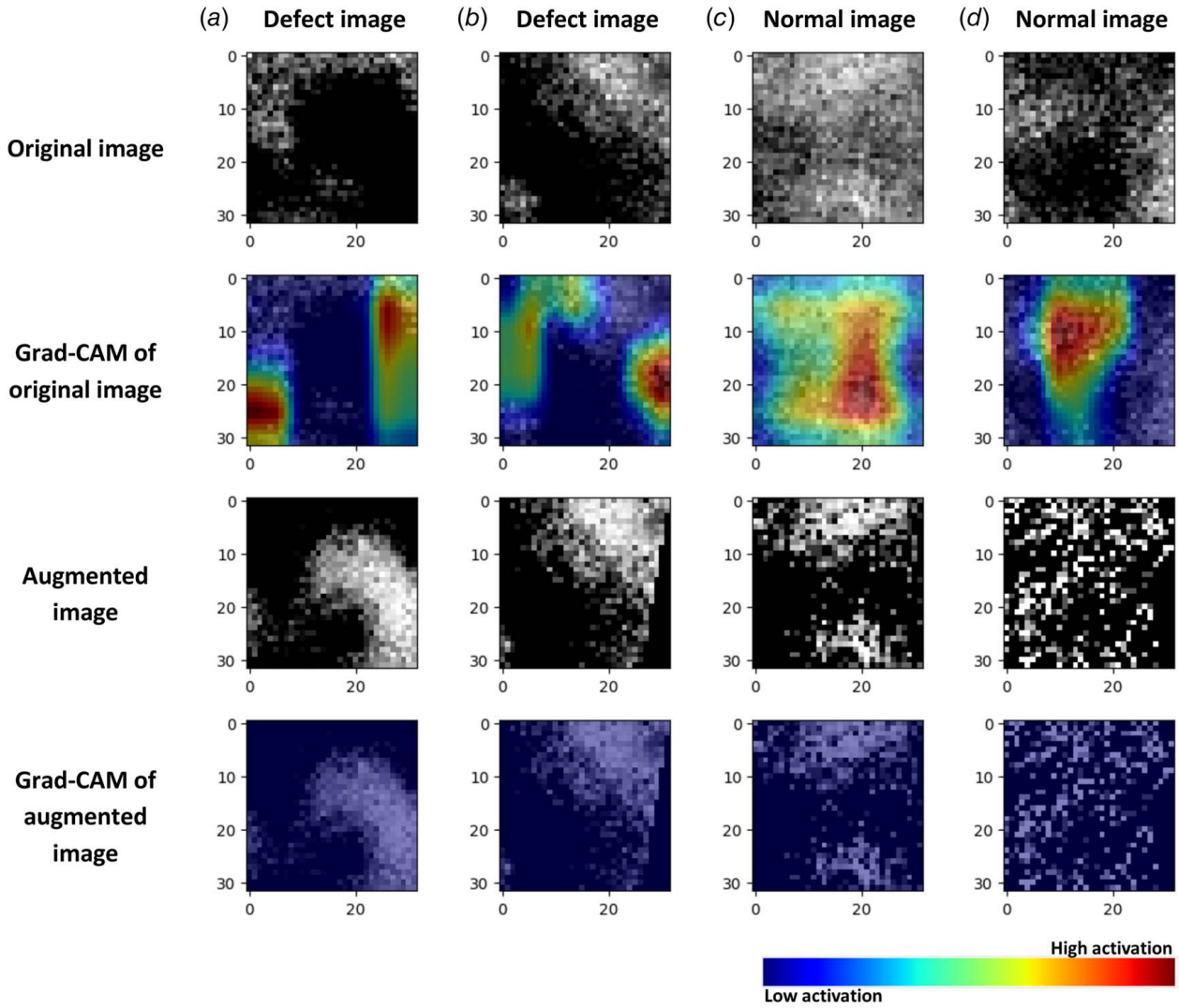

Fig. 21 Grad-CAM heat map of surfaces of low-resolution wafer images: (*a*) defect image, (*b*) defect image, (*c*) normal image, and (*d*) normal image

surfaces, and the trained model with the augmentation showed an F1 score of 78.6%. As shown in Fig. 21, the intricate surfaces of low-resolution wafer images resulted in deactivated pixels after augmenting; the model could not abstract features from highly transformed surfaces. Therefore, this study did not apply the "Auto-Augment" PyTorch library in the low-resolution wafer image case and sought another approach to augment images effectively.

Perez and Wang [67] combined CNN training with an augmentation neural network and improved image classifier performance. Unlike the previous augmentation technique consisting of rotating, inverting, and solarizing methods, an augmentation neural network with CNN layers provided image augmentation and feature abstraction simultaneously. Figure 22 illustrates the configuration of the augmentation neural network. Two grayscale images of the same

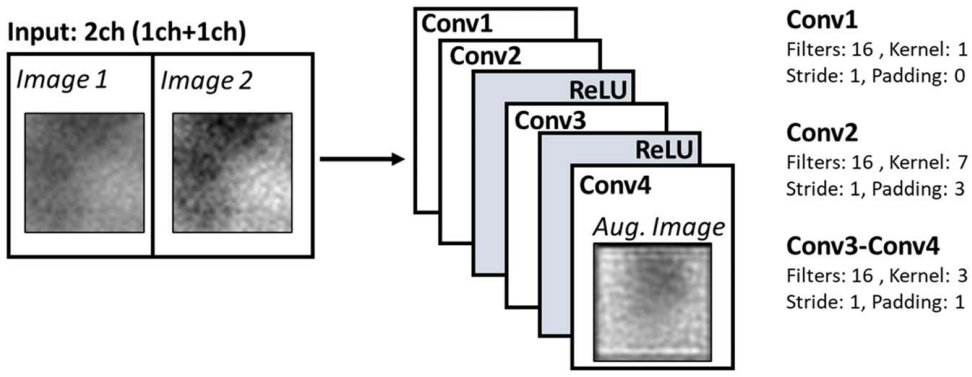

Fig. 22 Configuration of the augmentation neural network

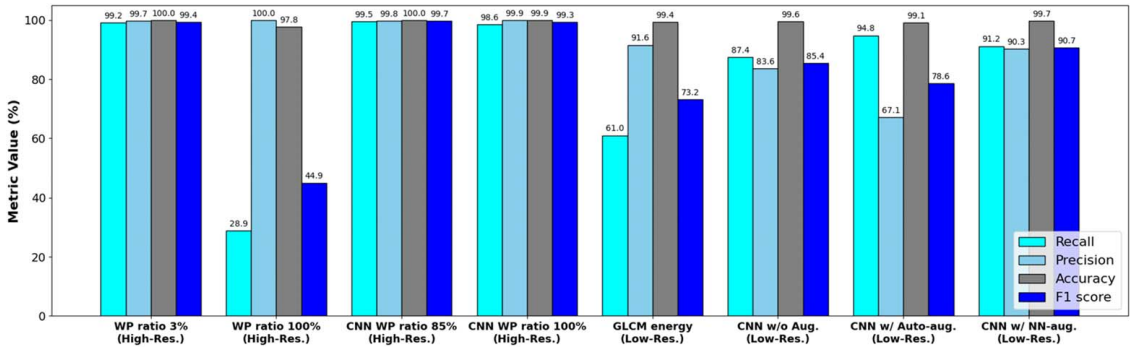

Fig. 23 Performance of the developed models

class, including one original image and another image with or without auto-augmentation, were concatenated and inputted to the network. The network was trained by evaluating mean square error with the Adam optimizer compared to an input original image. With calculated weights, the augmentation neural network generated one-channel augmented images. The augmented images via the neural network improved the F-score from 85.4% to 90.1% for defect localization. In other words, implementing an augmentation neural network was effective in improving the performance of defect localization in the low-resolution wafer image case. Finally, Fig. 23 summarizes the performance of each model.

6 Conclusion

Based on GLCM, background, subtraction, and deep learning methods, this study presents an automated visual semiconductor wafer inspection framework for classifying high- and low-resolution wafer images, identifying the defectiveness of wafers, and localizing defects with self-labeling techniques. The GLCM methods revealed the characteristics of wafers and surfaces, and the thresholds were defined by k-means clustering autonomously. Furthermore, the background subtraction technique revealed the presence of a defect as a white point cluster. Depending on WP ratios, the evaluation metrics showed varied performances. After implementing a CNN approach, the dependency of WP ratios was resolved by maintaining performance. While this framework successfully processed every step, yielding F1 scores for localizing defects over 90.1% for low-resolution wafer images and 99.3% for high-resolution wafer images, it has still several limitations as follows.

First, this investigation considered only two types of wafers. When handling multiple types of wafers, it may cause unclear separation on GLCM analysis with k-means clustering. Therefore, an additional algorithm would be necessary to address the issue. One tentative solution is the application of deep learning. By implementing an AE, it would be possible to extract representative features of each group and data sets for CNN training. Then, a trained CNN model is expected to separate each group. Next, this study solved the binary classification problems. Nevertheless, there should be classification problems with multiple classes. The features of defects in this research also could be divided into multiple classes by shapes and reflections. As a result, it will be the future study of this investigation.

In addition, the detection speed is also important for real-world applications. The speed of localizing defects was from 1.00 s per image to 1.75 s per image. For real-time inspection, the speed should reach around 0.03 s per image [68]. YOLO or Region-Based Convolutional Neural Network (R-CNN) with optimization will be implemented to tackle this problem. Moreover, the performance of defect localization for low-resolution wafer images should be improved to implement this framework in real-world cases, and the framework and models will be optimized in future research.

Acknowledgment

This material is based upon work supported by the National Science Foundation under Grant AM-2125826 (Purdue University) and Grant CMMI-2124999 (Texas A&M University).

Conflict of Interest

There are no conflicts of interest.

Data Availability Statement

The datasets generated and supporting the findings of this article are obtainable from the corresponding author upon reasonable request.

References

[1] Kang, H. S., Lee, J. Y., Choi, S., Kim, H., Park, J. H., Son, J. Y., Kim, B. H., and Noh, S. D., 2016, "Smart Manufacturing: Past Research, Present Findings, and Future Directions," Int. J. Precis. Eng. Manuf. Green Technol., **3**(1), pp. 111–128.
[2] Shankar, N. G., and Zhong, Z. W., 2005, "Defect Detection on Semiconductor Wafer Surfaces," Microelectron. Eng., **77**(3–4), pp. 337–346.
[3] Huang, S. H., and Pan, Y. C., 2015, "Automated Visual Inspection in the Semiconductor Industry: A Survey," Comput. Ind., **66**, pp. 1–10.
[4] Harris, D. H., 1969, "The Nature of Industrial Inspection," Hum. Factors, **11**(2), pp. 139–148.
[5] Chin, R. T., and Harlow, C. A., 1982, "Automated Visual Inspection: A Survey," IEEE Trans. Pattern Anal. Mach. Intell., **PAMI-4**(6), pp. 557–573.
[6] Megaw, E. D., 1979, "Factors Affecting Visual Inspection Accuracy," Appl. Ergon., **10**(1), pp. 27–32.
[7] Babic, M., Farahani, M. A., and Wuest, T., 2021, "Image Based Quality Inspection in Smart Manufacturing Systems: A Literature Review," Procedia CIRP, **103**, pp. 262–267.
[8] Psarommatis, F., May, G., Dreyfus, P.-A., and Kiritsis, D., 2019, "Zero Defect Manufacturing: State-of-the-Art Review, Shortcomings and Future Directions in Research," Artic," Int. J. Prod. Res., **58**(1), pp. 1–17.
[9] LeCun, Y., Bengio, Y., and Hinton, G., 2015, "Deep Learning," Nature, **521**(7553), pp. 436–444.
[10] Rusk, N., 2016, "Deep Learning," Nat. Methods, **13**(1), pp. 35–35.
[11] Wang, D., and Shang, Y., 2014, "A New Active Labeling Method for Deep Learning," Proceedings of the International Joint Conference on Neural Networks, Beijing, China, July 6–11, pp. 112–119.
[12] Chow, J. K., Su, Z., Wu, J., Tan, P. S., Mao, X., and Wang, Y. H., 2020, "Anomaly Detection of Defects on Concrete Structures With the Convolutional Autoencoder," Adv. Eng. Inform., **45**, p. 101105.
[13] Kozamernik, N., and Bračun, D., 2020, "Visual Inspection System for Anomaly Detection on KTL Coatings Using Variational Autoencoders," Procedia CIRP, **93**, pp. 1558–1563.
[14] Kouteva-Arguirova, S., Arguirov, T., Wolfframm, D., and Reif, J., 2003, "Influence of Local Heating on Micro-Raman Spectroscopy of Silicon," J. Appl. Phys., **94**(8), pp. 4946–4949.
[15] Kumar, A., and Arnold, W., 2022, "High Resolution in Non-Destructive Testing: A Review," J. Appl. Phys., **132**(10), p. 100901.
[16] Hasegawa, K., and Saito, H., 2016, "Synthesis of a Stroboscopic Image From a Hand-Held Camera Sequence for a Sports Analysis," Comput. Vis. Media, **2**(3), pp. 277–289.
[17] Guo, X., and Lee, C. B., 2021, "Preliminary Study of Phase-Shifting Strobo-Stereoscopy for Cutting Tool Monitoring," J. Manuf. Process., **64**, pp. 1214–1222.

[18] Zhang, L., Tham, Z. W., Chen, Y. F., Tan, C. Y., Cui, F., Mutiargo, B., and Ke, L., 2022, "Defect Imaging in Carbon Fiber Composites by Acoustic Shearography," Compos. Sci. Technol., **223**, p. 109417.
[19] Grasland-Mongrain, P., Zorgani, A., Nakagawa, S., Bernard, S., Paim, L. G., Fitzharris, G., Catheline, S., and Cloutier, G., 2018, "Ultrafast Imaging of Cell Elasticity With Optical Microelastography," Proc. Natl. Acad. Sci. U. S. A., **115**(5), pp. 861–866.
[20] Yu, Y., Yang, Q., and Wang, X., 2013, "3D Imaging Application in the Studies of Micro Air Vehicles," Comput. Ind., **64**(9), pp. 1178–1185.
[21] Versluis, M., 2013, "High-Speed Imaging in Fluids," Exp. Fluids, **54**(2), pp. 1–35.
[22] Brigham, E. O., and Morrow, R. E., 1967, "The Fast Fourier Transform," IEEE Spectr., **4**(12), pp. 63–70.
[23] Graps, A., 1995, "An Introduction to Wavelets," IEEE Comput. Sci. Eng., **2**(2), pp. 50–61.
[24] Singh, S. A., and Desai, K. A., 2023, "Automated Surface Defect Detection Framework Using Machine Vision and Convolutional Neural Networks," J. Intell. Manuf., **34**(4), pp. 1995–2011.
[25] Yun, J. P., Shin, W. C., Koo, G., Kim, M. S., Lee, C., and Lee, S. J., 2020, "Automated Defect Inspection System for Metal Surfaces Based on Deep Learning and Data Augmentation," J. Manuf. Syst., **55**, pp. 317–324.
[26] Park, J. K., Kwon, B. K., Park, J. H., and Kang, D. J., 2016, "Machine Learning-Based Imaging System for Surface Defect Inspection," Int. J. Precis. Eng. Manuf. Green Technol., **3**(3), pp. 303–310.
[27] Kulkarni, A., and Xu, C., 2021, "A Deep Learning Approach in Optical Inspection to Detect Hidden Hardware Trojans and Secure Cybersecurity in Electronics Manufacturing Supply Chains," Front. Mech. Eng., **7**.
[28] Nakazawa, T., and Kulkarni, D. V., 2018, "Wafer Map Defect Pattern Classification and Image Retrieval Using Convolutional Neural Network," IEEE Trans. Semicond. Manuf., **31**(2), pp. 309–314.
[29] Phua, C., and Theng, L. B., 2020, "Semiconductor Wafer Surface: Automatic Defect Classification with Deep CNN," IEEE Region 10 Annual International Conference Proceedings/TENCON, Osaka, Japan, November, pp. 714–719.
[30] Wen, G., Gao, Z., Cai, Q., Wang, Y., and Mei, S., 2020, "A Novel Method Based on Deep Convolutional Neural Networks for Wafer Semiconductor Surface Defect Inspection," IEEE Trans. Instrum. Meas., **69**(12), pp. 9668–9680.
[31] Yun, H., Kim, H., Jeong, Y. H., and Jun, M. B. G., 2023, "Autoencoder-Based Anomaly Detection of Industrial Robot Arm Using Stethoscope Based Internal Sound Sensor," J. Intell. Manuf., **34**(3), pp. 1427–1444.
[32] Erhan, D., Bengio, Y., Courville, A., Manzagol, P.-A., and Vincent, P., 2010, "Why Does Unsupervised Pre-Training Help Deep Learning?," J. Mach. Learn. Res., **11**, pp. 625–660.
[33] Feng, S., Yu, H., and Duarte, M. F., 2020, "Autoencoder Based Sample Selection for Self-Taught Learning," Knowl. Based Syst., **192**, p. 105343.
[34] Fan, S. K. S., Hsu, C. Y., Jen, C. H., Chen, K. L., and Juan, L. T., 2020, "Defective Wafer Detection Using a Denoising Autoencoder for Semiconductor Manufacturing Processes," Adv. Eng. Inform., **46**, p. 101166.
[35] Yu, J., Zheng, X., and Liu, J., 2019, "Stacked Convolutional Sparse Denoising Auto-Encoder for Identification of Defect Patterns in Semiconductor Wafer Map," Comput. Ind., **109**, pp. 121–133.
[36] Wang, C. H., 2009, "Separation of Composite Defect Patterns on Wafer Bin Map Using Support Vector Clustering," Expert Syst. Appl., **36**(2), pp. 2554–2561.
[37] Schlosser, T., Friedrich, M., Beuth, F., and Kowerko, D., 2022, "Improving Automated Visual Fault Inspection for Semiconductor Manufacturing Using a Hybrid Multistage System of Deep Neural Networks," J. Intell. Manuf., **33**(4), pp. 1099–1123.
[38] Paszke, A., Gross, S., Massa, F., Lerer, A., Bradbury, J., Chanan, G., Killeen, T., et al., 2019, "PyTorch: An Imperative Style, High-Performance Deep Learning Library," Advances in Neural Information Processing Systems 32, Vancouver, Canada, Dec. 8–14, H. Wallach, H. Larochelle, A. Beygelzimer, F. d'Alché-Buc, E. Fox, and R. Garnett, eds., Curran Associates, Inc, pp. 8024–8035.
[39] Holub, O., and Ferreira, S. T., 2006, "Quantitative Histogram Analysis of Images," Comput. Phys. Commun., **175**(9), pp. 620–623.
[40] Salem, N., Malik, H., and Shams, A., 2019, "Medical Image Enhancement Based on Histogram Algorithms," Procedia Comput. Sci., **163**, pp. 300–311.
[41] Haralick, R. M., 1979, "Statistical and Structural Approaches to Texture," Proc. IEEE, **67**(5), pp. 786–804.
[42] Hall-Beyer, M., 2017, "Practical Guidelines for Choosing GLCM Textures to Use in Landscape Classification Tasks Over a Range of Moderate Spatial Scales," Int. J. Remote Sens., **38**(5), pp. 1312–1338.
[43] Franklin, S. E., Hall, R. J., Moskal, L. M., Maudie, A. J., and Lavigne, M. B., 2000, "Incorporating Texture Into Classification of Forest Species Composition From Airborne Multispectral Images," Int. J. Remote Sens., **21**(1), pp. 61–79.
[44] Ozdemir, I., Mert, A., and Senturk, O., 2012, "Predicting Landscape Structural Metrics Using Aster Satellite Data," J. Environ. Eng. Landsc. Manag., **20**(2), pp. 168–176.
[45] Mathew, A. R., Anto, P. B., and Thara, N. K., 2017, "Brain Tumor Segmentation and Classification Using DWT, Gabour Wavelet and GLCM," 2017 International Conference on Intelligent Computing, Instrumentation and Control Technologies (ICICICT), Kerala, India, July 6–7, pp. 1744–1750.
[46] Hussain, A., and Khunteta, A., 2020, "Semantic Segmentation of Brain Tumor From MRI Images and SVM Classification Using GLCM Features," 2020 Second International Conference on Inventive Research in Computing Applications (ICIRCA), Coimbatore, India, July 15–17, pp. 38–43.
[47] Raheja, J. L., Kumar, S., and Chaudhary, A., 2013, "Fabric Defect Detection Based on GLCM and Gabor Filter: A Comparison," Optik, **124**(23), pp. 6469–6474.
[48] Yu, J., 2019, "Enhanced Stacked Denoising Autoencoder-Based Feature Learning for Recognition of Wafer Map Defects," IEEE Trans. Semicond. Manuf., **32**(4), pp. 613–624.
[49] Yu, J., and Lu, X., 2016, "Wafer Map Defect Detection and Recognition Using Joint Local and Nonlocal Linear Discriminant Analysis," IEEE Trans. Semicond. Manuf., **29**(1), pp. 33–43.
[50] Saqlain, M., Jargalsaikhan, B., and Lee, J. Y., 2019, "A Voting Ensemble Classifier for Wafer Map Defect Patterns Identification in Semiconductor Manufacturing," IEEE Trans. Semicond. Manuf., **32**(2), pp. 171–182.
[51] Gadkari, D., 2004, "Image Quality Analysis Using GLCM," Electron. Theses Diss. Univ. Cent. Fla.
[52] Sobral, A., and Vacavant, A., 2014, "A Comprehensive Review of Background Subtraction Algorithms Evaluated With Synthetic and Real Videos," Comput. Vis. Image Underst., **122**, pp. 4–21.
[53] Bouwmans, T., 2012, "Background Subtraction For Visual Surveillance: A Fuzzy Approach," Handb. Soft Comput. Video Surveill., pp. 103–134.
[54] Bradski, G., 2000, "The OpenCV Library," Dr Dobbs J. Softw. Tools.
[55] Trnovszký, T., Sýkora, P., and Hudec, R., 2017, "Comparison of Background Subtraction Methods on Near Infra-Red Spectrum Video Sequences," Procedia Eng., **192**, pp. 887–892.
[56] Friedman, N., and Russell, S., 2013, Image Segmentation in Video Sequences: A Probabilistic Approach," CoRR, abs/1302.1539.
[57] Zivkovic, Z., and Heijden, F. V. D., 2006, "Efficient Adaptive Density Estimation per Image Pixel for the Task of Background Subtraction," Pattern Recognit. Lett., **27**(7), pp. 773–780.
[58] Cubuk, E. D., Zoph, B., Mane, D., Vasudevan, V., and Le, Q. V., 2019, "Autoaugment: Learning Augmentation Strategies From Data," Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit., Long Beach, CA, June, pp. 113–123.
[59] Selvaraju, R. R., Cogswell, M., Das, A., Vedantam, R., Parikh, D., and Batra, D., 2020, "Grad-CAM: Visual Explanations From Deep Networks via Gradient-Based Localization," Int. J. Comput. Vis., **128**(2), pp. 336–359.
[60] Gunning, D., Stefik, M., Choi, J., Miller, T., Stumpf, S., and Yang, G. Z., 2019, "XAI-Explainable Artificial Intelligence," Sci. Robot., **4**(37), p. eaay7120.
[61] Wang, J., Fu, P., and Gao, R. X., 2019, "Machine Vision Intelligence for Product Defect Inspection Based on Deep Learning and Hough Transform," J. Manuf. Syst., **51**, pp. 52–60.
[62] Wang, J., Li, Y., Gao, R. X., and Zhang, F., 2022, "Hybrid Physics-Based and Data-Driven Models for Smart Manufacturing: Modelling, Simulation, and Explainability," J. Manuf. Syst., **63**, pp. 381–391.
[63] Cooper, C., Zhang, J., Huang, J., Bennett, J., Cao, J., and Gao, R. X., 2023, "Tensile Strength Prediction in Directed Energy Deposition Through Physics-Informed Machine Learning and Shapley Additive Explanations," J. Mater. Process. Technol., **315**, p. 117908.
[64] Wang, F., Zhao, Z., Zhai, Z., Shang, Z., Yan, R., and Chen, X., 2023, "Explainability-Driven Model Improvement for SOH Estimation of Lithium-Ion Battery," Reliab. Eng. Syst. Saf., **232**, p. 109046.
[65] Liu, T., Lough, C. S., Sehhat, H., Ren, Y. M., Christofides, P. D., Kinzel, E. C., and Leu, M. C., 2022, "In-Situ Infrared Thermographic Inspection for Local Powder Layer Thickness Measurement in Laser Powder Bed Fusion," Addit. Manuf., **55**, p. 102873.
[66] Lee, J., Noh, I., Lee, J., and Lee, S. W., 2022, "Development of an Explainable Fault Diagnosis Framework Based on Sensor Data Imagification: A Case Study of the Robotic Spot-Welding Process," IEEE Trans. Ind. Inform., **18**(10), pp. 6895–6904.
[67] Perez, L., and Wang, J., 2017, "The Effectiveness of Data Augmentation in Image Classification Using Deep Learning." arXiv preprint.
[68] Lee, J., Wang, P., Xu, R., Dasari, V., Weston, N., Li, Y., Bagchi, S., and Chaterji, S., 2021, "Virtuoso: Video-Based Intelligence for Real-Time Tuning on SOCs," arXiv preprint.