



Similarity searching for wafer bin maps by measuring shape, location, and size similarities of defect patterns

Min-Su Kang ^a, Jin-Su Shin ^{b,c}, Dong-Hee Lee ^{a,*}

^a Department of Industrial Engineering, Sungkyunkwan University, 2066 Seobu-ro, Jangan-gu, Suwon 16419, Republic of Korea

^b Department of Semiconductor and Display Engineering, Sungkyunkwan University, 2066, Seobu-ro, Suwon-si, Gyeonggi-do 16419, Republic of Korea

^c Memory Division, Samsung Electronics Co, Ltd., 1-1 Samsungjeonja-ro Hwaseong-si, Gyeonggi-do 18448, Republic of Korea

ARTICLE INFO

Keywords:

Wafer bin map
Similarity searching
Semiconductor manufacturing
Unsupervised learning
Defect pattern

ABSTRACT

A wafer bin map (WBM) is a visual representation of the spatial distribution of defective chips on a wafer. WBMs showing specific defect patterns are usually a result of process-assignable causes; thus, it is important to identify them to eliminate assignable causes. With advances in semiconductor manufacturing technology, identifying new defect patterns, and diagnosing their causes have become critical. However, most existing methods for WBM analysis use a supervised learning approach, which only detects previously known defect patterns. The similarity-search approach is a suitable alternative for defining new defect patterns. The proposed method uses an unsupervised approach to search for similar WBMs by measuring the similarities of three spatial features of defect patterns—shape, location, and size—which are useful for defining new defect patterns and diagnosing their causes. These three similarities are achieved using tensor voting and the mountain function for shape similarity, Euclidean distance for location similarity, and a combination of defect count and average radius for size similarity. The overall similarity was assessed using the weighted average of the three similarities. The weights are determined by quantifying the uncertainty of each similarity based on information entropy theory to better distinguish between similar patterns. The experimental results demonstrate the effectiveness of the proposed method compared to existing methods and highlight its capability to identify and describe the spatial features of defect patterns.

1. Introduction

After the intricate process of semiconductor manufacturing, where integrated circuits (ICs) are created on wafers, each chip undergoes an electrical die sorting test using probes to assess its basic functionality. The outcomes of these tests are mapped to corresponding wafer bin numbers. By utilizing these wafer bin numbers, a visualization of interest for defective areas is created, forming what is known as a wafer bin map (WBM) (Kyeong & Kim, 2018). The shapes or patterns of the combined chips in this WBM provide insights into process variability, equipment malfunctions, material defects, design flaws, and environmental factors. Therefore, accurate detection and comparison of defect patterns in WBM present an opportunity for improvement in semiconductor manufacturing (Chen & Liu, 2000; Hsu & Chien, 2007).

WBM analysis can be broadly categorized into two branches: supervised and unsupervised. Recent studies have relied predominantly on supervised learning techniques (Gallo & Capozzi, 2020; Kahng & Kim, 2021; Lee & Kim, 2020; Park & You, 2023; Wang & Ni, 2023), which involve pre-existing knowledge of defect patterns, labeled data, and extensive training processes. However, as technology advances rapidly, semiconductor manufacturing processes evolve swiftly, leading to the emergence of new defect patterns in WBM (Kyeong & Kim, 2018; Liao et al., 2014). Supervised learning methods are limited to identifying only patterns that are known beforehand, posing a significant constraint in the dynamic environment of semiconductor manufacturing, where new defect patterns frequently emerge (Lee et al., 2023).

The adoption of similarity search methods is necessary to detect new and unknown defect patterns effectively (Lee et al., 2021). These

Abbreviations: ICs, integrated circuits; WBM, wafer bin map; SVM, support vector machine; CNN, convolutional neural network; PCA, principal component analysis; GANs, generative adversarial networks; DPGMM, Dirichlet Process Gaussian Mixture Model; DCNN, Deep Convolutional Neural Network; WSCN, WaferSegClassNet.

* Corresponding author.

E-mail address: dhee@skku.edu (D.-H. Lee).

<https://doi.org/10.1016/j.cie.2024.110486>

Received 10 March 2024; Received in revised form 26 July 2024; Accepted 13 August 2024

Available online 14 August 2024

0360-8352/© 2024 Elsevier Ltd. All rights are reserved, including those for text and data mining, AI training, and similar technologies.

approaches involve comparing defect patterns across different wafers, identifying recurring issues and trends that are crucial for rapid diagnostics, and making efficient process improvements. According to [Hwang & Kim \(2020\)](#), analyzing the process histories of similar failures within a database can provide information for defect diagnosis and root cause analysis, even in the absence of predefined classes. Additionally, [\(Park et al. \(2021\)\)](#) suggest that using a similarity search on a WBM can help detect new categories of previously unidentified defect patterns.

In recent years, several studies using similarity search techniques for WBM analysis have been conducted. [Liao et al. \(2014\)](#) were the first to propose an algorithm that utilizes label information to perform similarity searches and rank WBMs based on similarity. Subsequently, methods employing supervised learning, which require label information, were developed for similarity searching ([Kong & Ni, 2022](#); [Nakazawa & Kulkarni, 2018](#); [Yu et al., 2019](#)). The supervised learning approaches used in these studies generally provide excellent performance. However, they have the limitation of identifying only previously recognized patterns.

To overcome this limitation, the use of unsupervised learning for similarity searches was proposed ([Wu et al., 2015](#)). Subsequent research focused on different aspects of defect patterns. For instance, [Hsu et al. \(2020\)](#) focused on capturing the density of defect clusters, while [Lee et al. \(2021\)](#) identified the locations of the defect clusters. Recently, [Wang & Wang \(2022\)](#) emphasized capturing the shapes of the defect clusters.

However, existing research on similarity searches fails to sufficiently capture the spatial features. For instance, [Hsu et al. \(2020\)](#) focused on capturing the density of defect clusters, while [Lee et al. \(2021\)](#) identified the locations of the defect clusters of defect patterns, thereby not providing adequate information about defects related to semiconductor manufacturing processes. Fundamentally, the spatial features of the defect patterns can be described in three dimensions: shape, size, and location. For an effective process diagnosis, it is crucial to identify the precise causes of defects and provide detailed information, such as lot or machine numbers, to better understand the defects. In this context, we believe that a method presenting three spatial aspects of defect patterns offers a higher likelihood of providing more detailed information about defects, along with defect suspicions, compared to considering only one or two aspects. Therefore, it is essential to simultaneously search for WBMs with similar defect patterns in all three aspects. However, the existing methods lack the ability to identify detailed spatial information regarding the shape, size, and location of defect patterns, thereby limiting their effectiveness in diagnosing the causes and improving the occurrence of defect patterns in semiconductor manufacturing processes.

For example, during photolithography, patterns are imprinted onto the wafer surface using light ([Nguyen et al., 2022](#)), stage movement, or stepper motor errors. These errors are often observed as linear defect patterns in WBMs. Most linear defect patterns are unique because they have different locations and sizes, and these differences result from different causes. In such cases, searching for linear defect patterns with similar locations and sizes is useful for tracking and diagnosing the causes. However, the existing methods focus on searching for linear-shaped patterns without considering their size and location. This claim is true not only for linear shapes but also for other shapes such as arcs, donuts, clusters, and rings. Chemical-mechanical planarization, a major semiconductor manufacturing process, can cause uneven slurry distribution, pad wear, and inconsistent pressure ([Basim & Moudgil, 2002](#)). These failures result in cluster-shaped patterns such as hotspots. Similarly, existing methods focus on searching for cluster-shaped patterns without considering their location and size in WBMs.

This study proposes a similarity search method that independently measures shape, location, and size similarities. This method can be applied to evaluate and explain the detailed features of defect patterns (i.e., shape, location, and size), thereby addressing the limitations of existing methods. Moreover, it enables easy identification of new defect

patterns based on these three similarities. We used tensor voting and the mountain function to measure the shape similarity, the Euclidean distance of the defect center between the two WBMs for location similarity, and the variance of the number of defects, in addition to the average radius of the defect cluster for size similarity. The overall similarity was then measured using a weighted average. Each weight was measured based on the uncertainty of its weight similarity using information entropy. This method enables a detailed comparison of how each WBM resembles the reference WBM in specific aspects, thus allowing engineers to adjust their weights to emphasize the features of interest. Unlike previous studies that failed to adequately consider shape, location, and size, our proposed method integrates all three dimensions.

This study presents the following specific contributions and novelties: Firstly, it proposes a similarity search method capable of independently measuring the similarity in shape, location, and size. This enables more precise analysis of defect patterns within complex WBM (Wafer Bin Map) images. Secondly, it introduces a method that utilizes information entropy to set weights according to the uncertainty of similarity measurements. This contributes to enhancing the accuracy of similarity assessments. Thirdly, by proposing an unsupervised learning-based WBM image classification method, it addresses the limitations associated with labeling and presents the possibility of detecting unknown defect types. This approach is expected to overcome the limitations of existing methods and significantly improve the efficiency of defect pattern analysis and detection in semiconductor manufacturing processes.

The remainder of this paper is structured as follows: [Section 2](#) explores the related literature. [Section 3](#) details the proposed method for similarity search. Experiments utilizing different datasets are presented in [Section 4](#). [Section 5](#) addresses the potential limitations of our approach and concludes.

2. Literature review

[Table 1](#) presents a review of various literatures on defect analysis in Wafer Bin Maps (WBM). WBM defect analysis research can be broadly divided into two main categories: one focusing on the accurate class classification of defect patterns, and the other utilizing similarity with target wafer maps to identify more clear and specific causes of defects, thereby aiding in engineering solutions. The existing studies on WBM defect analysis primarily employ supervised learning algorithms ([Kim & Behdinan, 2023](#)), particularly those studies on defect pattern classification aim to learn predefined defect type classes to detect them accurately and efficiently. However, these studies require pre-labeled data and have the limitation of potential misclassification when new, unlearned defect types occur. With the miniaturization of semiconductor processes and the introduction of new processes and equipment, there is an increasing demand for methods capable of identifying defect patterns similar to those of interest (typically new patterns). This demand has stimulated research on similarity search methods. The research on similarity comparison can be divided into supervised and unsupervised learning-based studies, with supervised learning methods using labeled data to search for WBMs similar to reference maps. Various studies have been conducted on this approach, highlighting its importance and effectiveness in identifying comparable defect patterns in WBMs.

[Liao et al. \(2014\)](#) proposed the concept of similarity search and designed a morphology-based support vector machine (SVM) for similarity retrieval that integrates morphological operations and SVM classification to detect defect patterns in WBMs. Morphological operations generate synthetic samples that simulate various defect patterns and enrich training datasets. The SVM classifier trained on this synthetic dataset identified the separation hyperplane between the defect classes.

[Nakazawa & Kulkarni \(2018\)](#) developed and applied a convolutional neural network (CNN) for the classification and image retrieval of defect patterns in WBM. This method leverages synthetic WBMs to train a CNN, achieving high classification accuracy and demonstrating the ability to

Table 1
Literature review.

Reference	Object	Learning type	Algorithm	Method of Studies
(S. Chen et al., 2022)	Classification	Supervised	Dual DCNN	WBM Defect Pattern class Classification
(Xu et al., 2022)			ResNet 18	
(E. S. Kim et al., 2021)			ResNet50	
(J. Yu et al., 2021)			DenseNet-GCForest	
(Bae & Kang, 2023)			Modified VGG16	
(Nag et al., 2022)			WSCN	WBM Defect Pattern class Classification & Segmentation
(Shinde et al., 2022)			YOLO V4	WBM Defect Pattern class Classification & Localization Detect
(Liao et al., 2014)			M-SVM	SVM trained on set of WBMs constructed by domain experts (Engineer's domain)
(Nakazawa & Kulkarni, 2018)			CNN	Binary classification (whether similar or not)
(N. Yu et al., 2019)			CNN/PCA	
(Kong & Ni, 2022)	Similarity Search	Unsupervised	CNN/GAN	Similarity using Euclidean Distance between Feature vector
(Wu et al., 2015)			Radon transform/SVM	Similarity using Euclidean Distance between Feature vector
(C. Y. Hsu et al., 2020)			Mountain function	Similarity using Weighted Modified Hausdorff Distance
(J. H. Lee et al., 2021)			DPGMM/ Hierarchical clustering	Similarity using the Jensen-Shannon divergence
(Wang & Wang, 2023)			Tensor voting	Similarity using best-buddies similarity score
Proposed	Similarity Search	Unsupervised	Mountain function/ Tensor voting	Similarity search using shape, location, and size

efficiently retrieve similar defect patterns from a large database.

Yu et al. (2019) presented a method for recognizing and searching for defect patterns in WBMs using CNNs and principal component analysis. This study introduces two CNN models, an 8-layer model for the initial inspection of defect patterns on WBMs and a more detailed 13-layer model dedicated to classifying specific defect patterns. For post-classification, the method extracts features using a CNN and then employs principal component analysis (PCA) for dimensionality reduction. In addition, it retrieves similar defect patterns and ranks them according to their similarities.

Kong & Ni (2022) proposed a deep CNN-based retrieval method resembling generative adversarial networks (GANs) to identify WBMs with unknown defect patterns by exploring the similarities between the defect patterns in WBMs. This approach is designed to work effectively even with a very limited number of samples for unknown defect patterns, addressing a common challenge in WBM defect pattern analysis.

Generally, these supervised learning-based methods perform well in searching for WBMs with defect patterns similar to the reference pattern. However, they present practical difficulties because they require a long training time, which can lead to inefficiency in the manufacturing process. Engineers are also required to manually label the defect patterns. This is time-consuming and labor-intensive, and it presents a major shortcoming in achieving process automation. Moreover, as mentioned in Section 1, they are constrained to identifying only previously known patterns, which is a significant limitation in the dynamic environment of semiconductor manufacturing, where new defect patterns emerge frequently (Lee et al., 2023).

To overcome these limitations, several studies have been conducted that do not require label information. Wu et al. (2015) proposed a system that extracts radon/geometry-based features and uses a SVM for classification. In addition, a wafer map similarity ranking based on the Euclidean distance of the extracted features was designed to retrieve defect patterns on WBMs, similar to a given query wafer map, which can assist engineers in identifying the root causes of similar failure patterns.

Hsu et al. (2020) employed a modified mountain function for the feature transformation. This method enhances defect features based on clustering density. To measure the similarity between defect patterns on WBMs, weighted modified Hausdorff distance was utilized, and metrics for binary classification were used to assess the accuracy of the model in identifying similar and dissimilar defect patterns in WBMs.

Lee et al. (2021) applied the polar coordinate transform, Dirichlet Process Gaussian Mixture Model (DPGMM), and the hierarchical clustering method to cluster similar defect patterns. In addition, weight vectors with respect to the clusters were derived to serve as reduced representations of the defect patterns in WBMs. The similarity

between the defect patterns was measured using Jensen-Shannon divergence based on the derived weight vectors.

Wang & Wang (2023) utilized tensor voting to remove noise and highlight structural information. They then used the weighted best-buddies similarity to measure the similarity between the defect patterns. The effectiveness of their method was assessed through binary classification to determine whether a WBM contained a specific query pattern. Wang & Wang (2023) addressed the challenges in recognizing rare- and mixed-type defects by proposing a similarity-searching approach.

Numerous similarity search studies have been conducted; however, there is a lack of comprehensive consideration of shape, location, and size. While Hsu et al. (2020) focused on density, they did not adequately address the spatial characteristics of defect patterns, leading to analyses that missed essential details regarding defect locations. This oversight is particularly crucial for the accurate evaluation of ring patterns, where understanding the spatial context is crucial for accuracy. Wu et al. (2015) and Lee et al. (2021) considered the geometric features of defect patterns and the positioning of clusters but did not fully explore the shape attributes. Additionally, Wang & Wang (2023) did not thoroughly address the location or size of the defect patterns. This limitation is particularly significant for cluster-shaped patterns, defined as circular defect clusters, where the size plays a crucial role in differentiating them from ring-shaped patterns, with variations in size suggesting different root causes. Overall, there is a gap in existing research on the integration of spatial features in similarity search methodologies.

The proposed method attempts to independently assess shape, location, and size before synthesizing these dimensions to establish a comprehensive similarity ranking and ensure thorough consideration of spatial characteristics. As highlighted in Section 1, we emphasize the critical need to consider shape, size, and location to distinguish between linear and cluster shapes. Despite the significance of these factors, existing supervised and unsupervised methods have struggled to address them adequately. Our method addresses these gaps by providing a comprehensive approach that integrates shape, location, and size into the similarity search process, thereby facilitating more precise root cause analysis and similarity assessment.

3. Proposed method

The proposed method aims to compute the similarity between the defect patterns of two WBMs. The first corresponds to a reference WBM that typically has a new defect pattern of interest. The other corresponds to the candidate WBM, which is compared with the reference WBM. The reference and candidate WBM are denoted as P and Q, respectively. To

compute the similarity between P and Q, three individual similarities of shape, location, and size between P and Q are computed and then aggregated into the overall similarity. We assume that P and Q have the same number of dies and are denoted by N. The notations associated with P are listed in Table 2. The notation in Table 2 is also related to Q if we regard P as Q.

The overall flowchart of the proposed method operates as shown in Fig. 1. Initially, noise is filtered through two preprocessing stages of the WBM. Next, shape similarity is assessed using the product of density similarity and stick-ball similarity. The modified mountain function is utilized to determine density similarity, while tensor voting is used to derive stick-ball saliency. Subsequently, location similarity is computed based on the distance between the centroids of the patterns in the two WBMs. Size similarity is determined by the product of the similarity in the number of defects and the radius of the patterns. Finally, overall similarity is derived by summing all individual similarities using a weighted average approach.

3.1. Noise filtering

The proposed method includes a two-step noise-filtering process that involves the sequential use of c-mean filtering and tensor-voting algorithms. This eliminates unnecessary information from the WBMs images and extracts more accurate defect patterns to help the proposed similarity searching process focus only on defect patterns and not on noise. In WBMs, some defective chips deviate significantly from the defect patterns; therefore, these cases are considered noise and must be removed to enhance the performance of the proposed method. The effect of noise filtering is directly related to the improvement in algorithm performance, and this two-step noise filtering process differs from that

Table 2
List of notations.

Symbol	Description	Value
i, j	Index of the i th (j th) die on P	$i, j \in \{1, 2, \dots, N\}$
P_i	Coordinate of the i th die on P	$(x_i, y_i)^T \in \mathbb{Z}_{\geq 0}^2$
$Bin(P_i)$	Bin value of the i th die on P . 0 for background, 1 for normal die, 2 for defective die	$Bin(P_i) \in \{0, 1, 2\}$
P_{def}	Set of defective dies on P	$P_{def} = \{i Bin(P_i) = 2\}$
$ P_{def} $	Cardinality of P_{def}	$ P_{def} \in \mathbb{Z}_{\geq 0}$, $ P_{def} \leq N$
n_i	Set of neighbors around P_i	$n_i \subset \{P_1, \dots, P_N\}$
t_c	Threshold for c-mean filtering	$t_c \in \mathbb{R}_{>0}$
K_i	Geometric tensor at P_i	$K_i \in \mathbb{R}^{2 \times 2}$
S_{ij}	Tensor vote cast by P_j to P_i	$S_{ij} \in \mathbb{R}^{2 \times 2}$
c_{ij}	Weighting coefficient using in tensor voting	$c_{ij} \in \mathbb{R}_{\geq 0}$
R_{ij}	Rotation matrix aligns the tensor vote	$R_{ij} \in \mathbb{R}^{2 \times 2}$
R_{ij}'	Modified rotation matrix aligns the tensor vote	$R_{ij}' \in \mathbb{R}^{2 \times 2}$
$d(P_i, P_j)$	Euclidean distance between P_i and P_j	$d(P_i, P_j) \in \mathbb{R}_{\geq 0}$
σ_d	Scale parameter of c_{ij}	$\sigma_d \in \mathbb{R}_{\geq 0}$
r_{ij}	Unit vector pointing from P_j to P_i	$r_{ij} \in \mathbb{R}^2$
λ_1	Largest eigenvalue of K_i	$\lambda_1 \in \mathbb{R}$
λ_2	Second largest eigenvalue of K_i	$\lambda_2 \in \mathbb{R}$
\vec{e}_1	Eigenvector corresponding to the eigenvalue λ_1	$\vec{e}_1 \in \mathbb{R}^2_{\geq 0}$
\vec{e}_2	Eigenvector corresponding to the eigenvalue λ_2	$\vec{e}_2 \in \mathbb{R}^2_{\geq 0}$
λ_s	Stick saliency of K_i	$\lambda_s \in \mathbb{R}_{\geq 0}$
λ_b	Ball saliency of K_i	$\lambda_b \in \mathbb{R}_{\geq 0}$
T_s	Stick tensor of K_i	$T_s \in \mathbb{R}^{2 \times 2}$
T_b	Ball tensor of K_i	$T_b \in \mathbb{R}^{2 \times 2}$
t_t	Threshold for tensor voting filtering	$t_t \in \mathbb{R}_{>0}$
$M(P_i)$	Mountain value of P_i	$M(P_i) \in \mathbb{R}_{\geq 0}$
β	Normalizing factor for mountain function	$\beta \in \mathbb{R}_{\geq 0}$
$P_{centroid}$	Coordinate of centroid of P_{def}	$P_{centroid} \in \{P_1, \dots, P_N\}$
P_{center}	Coordinate of center of P	$P_{center} \in \{P_1, \dots, P_N\}$
λ_{P_i}	Saliencies of P_i	$(\lambda_s, \lambda_b)^T \in \mathbb{R}^2_{\geq 0}$
R_P	The radius of defect pattern in P	$R_P \in \mathbb{R}_{\geq 0}$
H	Uncertainty derived by informational entropy	$H \in \mathbb{R}_{\geq 0}$

of previous studies.

First, c-mean filtering proposed by (Yu et al., 2021) was conducted. This filtering method is effective in maintaining linear patterns by specifying the areas surrounding each faulty die, such as scratch patterns. It specifically targets the area surrounding each defective die, proving to be efficient in maintaining linear patterns, such as scratch patterns. According to (Xu et al., 2022), this method prevents edge blurring or the misclassification of normal dies as defective. As listed in Table 2, each defective die has a bin value of two. Based on the following equation, the bin value of each chip was updated: If the updated bin value is 1, the corresponding die is considered noise.

$$Bin(P_i) = \begin{cases} 2 \text{if } \frac{1}{|n_i|} \sum_{j \in n_i} Bin(P_j) \geq t_c & \text{for } i \in P_{def} \\ 1 \text{others} \end{cases}$$

Second, noise filtering was conducted through tensor voting, which is a pattern-recognition algorithm (Medioni et al., 2000). This method involves gathering votes S_{ij} from nearby pixels to capture spatial characteristics, which are represented as a second-order symmetric tensor K_i (T.-P. Wu et al., 2016) as follows:

$$K_i = \sum_j S_{ij} \text{ for } i, j \in P_{def}$$

Subsequently, the tensor-voting process at point P_i , influenced by the tensor K_j located at P_j is described by the following formula:

$$S_{ij} = c_{ij} R_{ij} K_j R_{ij}'$$

Matrix R_{ij} is a rotation matrix that modifies the orientation of a tensor based on the directional relationship between points P_i and P_j , with unit vector r_{ij} where $R_{ij} = I - 2r_{ij}r_{ij}^T$ and $R_{ij}' = (I - 2r_{ij}r_{ij}^T)R_{ij}^{-T}$. The $R_{ij}' = (I - 2r_{ij}r_{ij}^T)R_{ij}^{-T}$ combines a reflection operation $I - 2r_{ij}r_{ij}^T$, with the transpose of R_{ij} to make the final orientation and contribution of K_j to P_i through S_{ij} . Here, I represents the identity matrix and r_{ij} is a unit vector pointing from P_j to P_i . $c_{ij} = \exp\left(-\frac{d(P_i, P_j)}{\sigma_d}\right)$ is the normalization factor with σ_d , the scale parameter which affects how c_{ij} decreases with distance. In our approach, K_i initially begins as an identity matrix and is subsequently updated immediately after gathering all tensor votes.

After the process of aggregating tensor votes is completed, K_i is decomposed into

$$K_i = \lambda_1 \vec{e}_1 \vec{e}_1^T + \lambda_2 \vec{e}_2 \vec{e}_2^T$$

where \vec{e}_1 and \vec{e}_2 are the eigenvectors of K_i with corresponding eigenvalues λ_1 , λ_2 ($\lambda_1 \geq \lambda_2$). K_i can be further split into stick and ball components as follows:

$$K_i = \lambda_s T_s + \lambda_b T_b$$

where $T_s = \vec{e}_1 \vec{e}_1^T - \vec{e}_2 \vec{e}_2^T$ and $T_b = \vec{e}_2 \vec{e}_2^T$. Here, $\lambda_s = \lambda_1 - \lambda_2$ represents stick saliency, indicating the confidence of the direction encoded. $T_b = \lambda_2$ reflects ball saliency, measuring the extent to which the point's structure is non-oriented. A higher λ_s value suggests a concentration of surrounding points forming a line, while a higher λ_b indicates a clustering of points into a ball shape. By retaining the points satisfying $\lambda_s \geq t_t \times \max(\lambda_s)$ and $\lambda_b \geq t_t \times \max(\lambda_b)$, patterns resembling lines and spheres can be isolated (Wang & Wang, 2023). Wang & Wang (2023) demonstrated that conducting tensor voting exclusively on defective dies in WBMs can effectively identify defect patterns. λ_s and λ_b values derived from tensor voting on defective dies provide insights into the patterns formed, which are instrumental for the next phase of similarity measurement.

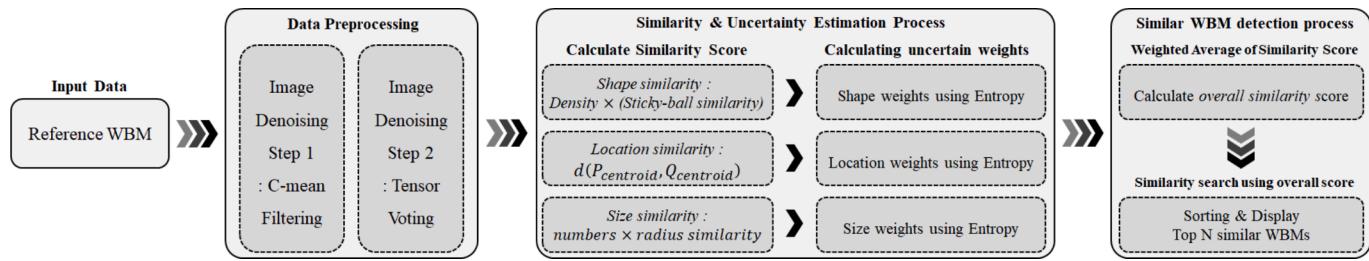


Fig. 1. Overall Flowchart of Proposed Method.

3.2. Shape similarity

Density and stick-ball features are essential indicators for describing the shape of defect patterns, with ball-like shapes indicating clustered defects due to contamination and stick-like shapes indicating systematic defects from photolithography or etching processes. Therefore, the shape similarity was calculated as the product of the density similarity and stick-ball similarity as follows: This approach is based on the premise that both the concentration of defects and the presence of distinct stick-ball patterns are critical for characterizing the shapes of defect clusters. Combining these two measures into a single metric allows for a precise evaluation of shape similarity and captures the complexity of defect patterns.

$$\text{shape similarity} = \text{density similarity} \times (\text{stick-ball similarity}).$$

The density similarity was assessed by comparing the defect densities at the centroid of P . We employed a modified mountain function to calculate the density at the centroid, which approximates the probability density function of the feature according to [Hsu et al. \(2020\)](#). The density similarity is defined as follows:

$$\text{density similarity} = 1 - \left| \frac{M(P_{\text{centroid}})}{\max_{i \in \{1, \dots, N\}} (M(P_i))} - \frac{M(Q_{\text{centroid}})}{\max_{i \in \{1, \dots, N\}} (M(Q_i))} \right|,$$

$$\text{where } M(P_i) = \sum_{j \in P_{\text{def}}} \exp(-\beta \cdot d(P_i, P_j)) \text{ and } P_{\text{centroid}} = \frac{1}{|P_{\text{def}}|} \sum_{i \in P_{\text{def}}} P_i.$$

Here, β serves as a normalizing factor for distances, calculated as $\beta = \left(\frac{\sum_j d(P_j, P_{\text{center}})}{N} \right)^{-1}$. By normalizing the mountain function value to its maximum within the WBM and calculating the difference, we normalize and compute the density similarity within the range [0, 1]. The stick-ball Similarity is determined based on the saliency of the stick and ball shapes derived during the tensor-voting process explained in [Section 3.1](#). This was calculated using the following formula:

$$\text{stick-ball similarity} = \exp \left(\frac{-1}{|P_{\text{def}}| |Q_{\text{def}}|} \sum_{i \in P_{\text{def}}} \sum_{j \in Q_{\text{def}}} d(\lambda_{P_i}, \lambda_{Q_j}) \right).$$

Here, $\frac{1}{|P_{\text{def}}| |Q_{\text{def}}|} \sum_{i \in P_{\text{def}}} \sum_{j \in Q_{\text{def}}} d(\lambda_{P_i}, \lambda_{Q_j})$ represents the average Euclidean distance between the saliencies of each defective die in two WBMs, with λ_{P_i} denoting the saliency (λ_s, λ_b) of the i th die on P . This approach causes the stick-ball similarity to fall within the range [0, 1] and decrease rapidly as the differences grow.

3.3. Location similarity

Location similarity is calculated by determining the Euclidean distance between the centroids of the defective dies on two WBMs. This distance is then used in an exponential decay function, where the decay constant, denoted by α , influences the rate at which the function declines. The formula used to calculate location similarity is as follows:

$$\text{location similarity} = \exp(-\alpha \cdot d(P_{\text{centroid}}, Q_{\text{centroid}})).$$

where $d(P_{\text{centroid}}, Q_{\text{centroid}})$ represent the Euclidean distances between the centroids of the defective dies in the two compared WBMs. The constant α is crucial as it dictates how rapidly the similarity value diminishes with increasing distance. Based on empirical observations, we have chosen an α value of 0.1 to gauge location similarity across WBMs with both similar and dissimilar patterns, adjusting α to refine our understanding of how it affects these comparisons.

3.4. Size similarity

The size of a defect pattern within a WBM can be defined by two critical factors: the number of defective dies it encompasses and its average radius. This allows us to differentiate between the sizes of the two defect patterns, even if they contain the same number of defective dies, because the radii of these patterns may vary. Therefore, the size of the pattern is determined by both the number of defective dies and their average radii. Consequently, we define the size similarity as the product of the defect quantity and radius similarities as follows:

$$\text{size similarity} = \text{defect numbers similarity} \times \text{radius similarity}.$$

Defect numbers similarity is computed as follows:

$$\text{defect numbers similarity} = \exp \left(\frac{\min(|P_{\text{def}}|, |Q_{\text{def}}|)}{\max(|P_{\text{def}}|, |Q_{\text{def}}|)} \right)$$

This calculation involves the ratio of the total number of defective dies in the defect pattern in Q to the defect pattern in P and applying the result within an exponential decay function to gauge the similarity in defect quantity. **Radius similarity** was assessed based on the ratio of the radii of the two patterns:

$$\text{radius similarity} = \frac{\min(R_p, R_q)}{\max(R_p, R_q)}.$$

where R_p and R_q represent the average distances from each defective die to P_{centroid} and Q_{centroid} , respectively. To represent the radius of a pattern with sporadic width changes, R_p and R_q are computed as follows:

$$R_p = \frac{1}{|P_{\text{def}}|} \sum_{i \in P_{\text{def}}} d(P_i, P_{\text{centroid}}),$$

$$R_q = \frac{1}{|Q_{\text{def}}|} \sum_{i \in Q_{\text{def}}} d(Q_i, Q_{\text{centroid}}).$$

3.5. Overall similarity

To calculate the overall similarity between P and Q , we used a weighted average of the shape, location, and size similarities as follows:

$$\begin{aligned} \text{Overall similarity} = & w_{\text{shape}} \times \text{shape similarity} + w_{\text{loc}} \\ & \times \text{location similarity} + w_{\text{size}} \times \text{size similarity}. \end{aligned}$$

Here, the weights w_{shape} , w_{location} and w_{size} account for the differing sig-

nificance of each similarity aspect, with the sum of weights equal to 1 ($\sum w = 1$), and each weight being a positive real number ($w \in \mathbb{R}^+$). This weighted average allows each type of similarity to contribute according to its importance, with the overall similarity score ranging from 0 to 1, making it easy to compare overall similarities when the reference WBM changes. Engineers can manually set the values of the weights according to their subjective judgment of the importance of the similarities. This is recommended when engineers have prior knowledge of the defect patterns. If this is not the case, we recommend setting the weight based on information entropy theory (Gray, 2011; Pan & Deng, 2020) to improve the discrimination between similar patterns. This ultimately leads to a search for Q s that are more similar Q s to a given P . Information entropy is a measure of uncertainty or unpredictability in a dataset, and the weights can be determined based on this uncertainty. In general, information entropy can be represented as

$$H = - \sum_x p(x) \log p(x)$$

In this formula, $p(x)$ represents the probability of occurrence of x and H denotes the overall uncertainty. A lower value of H indicates higher certainty. In the proposed method, each similarity measure (shape, location, and size) can be interpreted as the likelihood of a given Q s matching P . Suppose there exist U Q s, and each Q is compared to P . Then, each Q has three similarities according to the formulations in Sections 3.2, 3.3, and 3.4. Let us denote the three similarities of the u th Q as $shapemimilarity_u$, $locationsimilarity_u$, and $sizesimilarity_u$ for $u = 1, 2, \dots, U$. Then, the likelihood of each similarity for the u th Q is defined as

$$p(shapemimilarity_u) = \frac{shapemimilarity_u}{\sum_u^U shapemimilarity_u},$$

$$p(locationsimilarity_u) = \frac{locationsimilarity_u}{\sum_u^U locationsimilarity_u},$$

$$p(sizesimilarity_u) = \frac{sizesimilarity_u}{\sum_u^U sizesimilarity_u}.$$

According to Gray (2011), the certainty of each similarity can be quantified by calculating its respective entropy, H .

$$H_{shape} = - \sum_u^U p(shapemimilarity_u) \times \log p(shapemimilarity_u),$$

$$H_{location} = - \sum_u^U p(locationsimilarity_u) \times \log p(locationsimilarity_u),$$

$$H_{size} = - \sum_u^U p(sizesimilarity_u) \times \log p(sizesimilarity_u).$$

Entropy (H) quantifies the uncertainty associated with each similarity. High entropy means that the uncertainty is high; thus, the weight has a small value. For example, suppose that the shape entropy (H_{shape}) is high. According to the entropy equation, this indicates that the shape likelihood values (i.e., $p(shapemimilarity_u)$ for $u = 1, 2, \dots, U$) are likely uniform. These uniform values imply that the uncertainty is high, and they are not helpful for distinguishing Q s from P . For this reason, we assign a small value to the weight when the entropy is high. For the same reason, a large value is assigned to the weight when the entropy is low. Considering this inverse relationship between H and the weight, the inverse of H is used in the weight calculation. The softmax function was applied to these inverse H values to highlight the differences in certainty between the similarity measures, thereby making the weights more sensitive to changes. The weights were calculated as follows:

$$w_{shape} = \frac{\exp(-H_{shape})}{\exp(-H_{shape}) + \exp(-H_{location}) + \exp(-H_{size})},$$

$$w_{location} = \frac{\exp(-H_{location})}{\exp(-H_{shape}) + \exp(-H_{location}) + \exp(-H_{size})},$$

$$w_{size} = \frac{\exp(-H_{size})}{\exp(-H_{shape}) + \exp(-H_{location}) + \exp(-H_{size})}.$$

4. Experiment

This section presents several experiments conducted using the proposed method on the WM-811 K dataset. The WM-811 K dataset contains 811,457 WBMs from actual semiconductor manufacturing processes (M. J. Wu et al., 2015). It categorizes WBMs into eight defect pattern types, as listed in Table 3. They are Center, Donut, Edge-local, Edge-Ring, Loc, Random, Scratch, and Near full of 632 different sizes. Among the entire WBM dataset, about 21.3 % are labeled, and the remaining 78.7 % are unlabeled. While previous supervised learning-based studies utilized only the labeled data (21.3 %) for accuracy comparison in classification and similarity measurement, this study conducted similarity measurement experiments using the entire WM-811 K dataset, as it employs rule-based unsupervised learning that does not require labeling.

Section 4.1 illustrates the process of computing the three similarities for P and Q selected from the WM-811 K dataset. Section 4.2 reports the results of the similarity search for the eight defect pattern types. A similarity search was conducted without size normalization to avoid distortion of the spatial information of the WBM. Section 4.3 confirms the results of ablation studies to confirm the effectiveness of each similarity comparison algorithm and uncertainty estimation using entropy. Section 4.4 compares the performance of the proposed method with that of an existing similarity-search method.

4.1. Illustration of computing shape, location, and size similarities

We selected two WBMs from the WM-811 K dataset—one for P and the other for Q , and computed the three similarities. The first step was a two-step noise-filtering process. The first column of Table 4 displays P and Q , whose defect pattern types are center and scratch, respectively. The second column shows P and Q after c-mean filtering. The third column shows the saliency map of λ_s and λ_b after applying the tensor voting process. In the saliency map, a brighter color indicates a higher value. As mentioned in Section 3.1, c-mean filtering followed by tensor voting effectively removed most noise while preserving the scratch pattern. The last column lists the P and Q values after tensor voting. The edge threshold for c-mean, the sigma value, and saliency threshold for tensor voting were set to $t_c = 1.25$, $\sigma_d = 3$, $t_t = 0.3$ respectively, following the same values described in Wang & Wang (2023) and Yu et al. (2021).

After filtering, the shape, location, and size similarities were measured. Firstly, the average distance is calculated using the saliency values which were derived by the tensor voting as $\sum_{i \in P_{def}} \sum_{j \in Q_{def}} d(\lambda_{P_i}, \lambda_{Q_j}) = 11185$, $|P_{def}| = 145$, $|Q_{def}| = 29$. Using these values, the stick-ball similarity is derived as follows:

$$\text{stick-ball similarity} = \exp \left(\frac{-1}{|P_{def}| |Q_{def}|} \sum_{i \in P_{def}} \sum_{j \in Q_{def}} d(\lambda_{P_i}, \lambda_{Q_j}) \right) = 0.12$$

The density similarity was measured using the normalized mountain value at the WBM centroid. Fig. 2 represents the peak values of P and Q and their centroids.

The brighter color indicates a larger mountain value, and the green dot represents the centroids of P and Q . Below are the mountain values at the centroids and the maximum values of P and Q .

$$M(P_{centroid}) = 3.93, M(Q_{centroid}) = 2.39$$

$$\max_{i \in \{1, \dots, N\}} (M(P_i)) = 3.93, \max_{i \in \{1, \dots, N\}} (M(Q_i)) = 2.84$$

Table 3

Defect pattern types of WM-811 K.

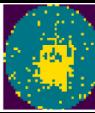
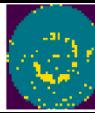
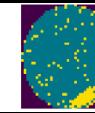
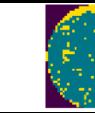
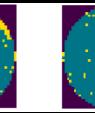
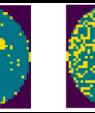
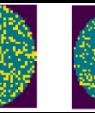
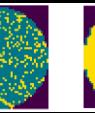
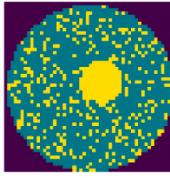
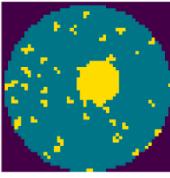
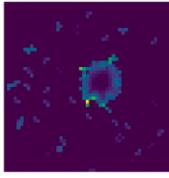
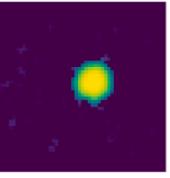
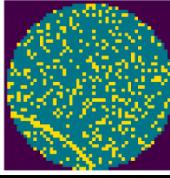
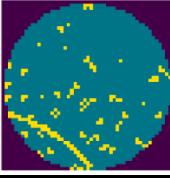
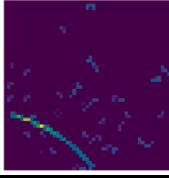
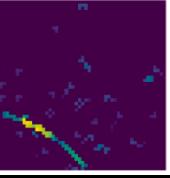
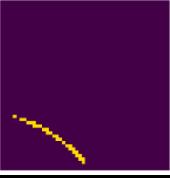
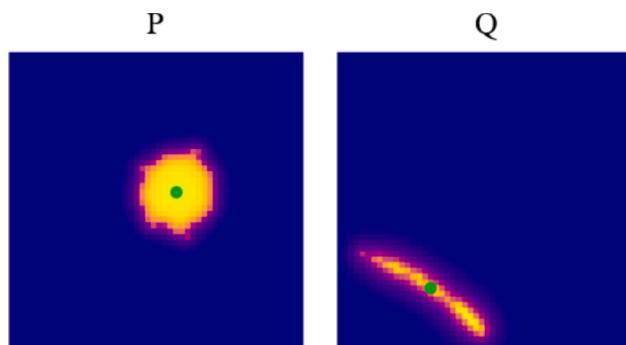
Type	Center	Donut	Edge-local	Edge-Ring	Loc	Random	Scratch	Near-full
WBM								

Table 4A two-step noise filtering process of P and Q .

WBM	Original	c-mean	Tensor voting		Result
			λ_s	λ_b	
P					
					

**Fig. 2.** Mountain values of P and Q

Note that $M(P_{centroid})$ and $M(Q_{centroid})$ are the values indicated by the green dots. Using the above values, the density similarity can be calculated as follows:

$$\text{densitySimilarity} = 1 - \left| \frac{M(P_{centroid})}{\max_{i \in \{1, \dots, N\}} (M(P_i))} - \frac{M(Q_{centroid})}{\max_{i \in \{1, \dots, N\}} (M(Q_i))} \right| = 0.84$$

The shape similarity can be calculated by multiplying the density similarity and stick-ball similarity as follows:

$$\begin{aligned} \text{ShapeSimilarity} &= \text{DensitySimilarity} \times \text{StickballSimilarity} = 0.55 \times 0.12 \\ &= 0.07 \end{aligned}$$

The location similarity is calculated as the Euclidean distance between two centroids, which is the same as the distance between the two green dots above:

$$d(P_{centroid}, Q_{centroid}) = 21.84.$$

As we set the α as 0.1, the location similarity is calculated as:

$$\text{locationSimilarity} = \exp(-\alpha \times d(P_{centroid}, Q_{centroid})) = 0.11.$$

Size similarity was calculated as the product of the similarity of the defect numbers and radii. Because we have already calculated the $|P_{def}|$, $|Q_{def}|$, the defect similarity can be derived as

$$\text{defectnumberssimilarity} = \exp\left(\frac{\min(|P_{def}|, |Q_{def}|)}{\max(|P_{def}|, |Q_{def}|)}\right) = 0.20.$$

Radius similarity can be derived from the average radius of the centroid. The average radii of P and Q are as follows:

$$R_P = \frac{1}{|P_{def}|} \sum_{i \in P_{def}} d(P_i, P_{centroid}) = 4.57,$$

$$R_Q = \frac{1}{|Q_{def}|} \sum_{i \in Q_{def}} d(Q_i, Q_{centroid}) = 6.53,$$

where $\sum_{i \in P_{def}} d(P_i, P_{centroid}) = 662.77$ and $\sum_{i \in Q_{def}} d(Q_i, Q_{centroid}) = 189.30$. Using above, radius similarity is calculated as:

$$\text{radiussimilarity} = \frac{\min(R_P, R_Q)}{\max(R_P, R_Q)} = 0.70$$

Size similarity was measured by aggregating the defect number similarity and radius similarity by product.

$$\begin{aligned} \text{Sizesimilarity} &= \text{Defectnumberssimilarity} \times \text{RadiusSimilarity} \\ &= 0.20 \times 0.70 = 0.14 \end{aligned}$$

Briefly, the three similarities were computed as (ShapeSimilarity,

Locationsimilarity, Sizesimilarity) = (0.07, 0.11, and 0.14). This score was rather low and was expected before computing the similarities, because P and Q were quite different, as shown in Table 4. This result implies that the proposed method performs well.

4.2. Similarity search using WM-811 K

By substituting Q with every WBM in the WM-811 K dataset and

repeating the process described in Section 4.1, the three similarities of each WBM in the WM-811 K were computed. These similarity values were then used to calculate the weights w_{shape} , $w_{location}$, w_{size} , which in turn were used to compute the overall similarity and perform the ranking.

Fig. 3 shows the results of similarity matching. For each defect pattern type listed in Table 3, we selected a single WBM and regarded it as P . The selected P for each defect pattern type is shown in the first row,

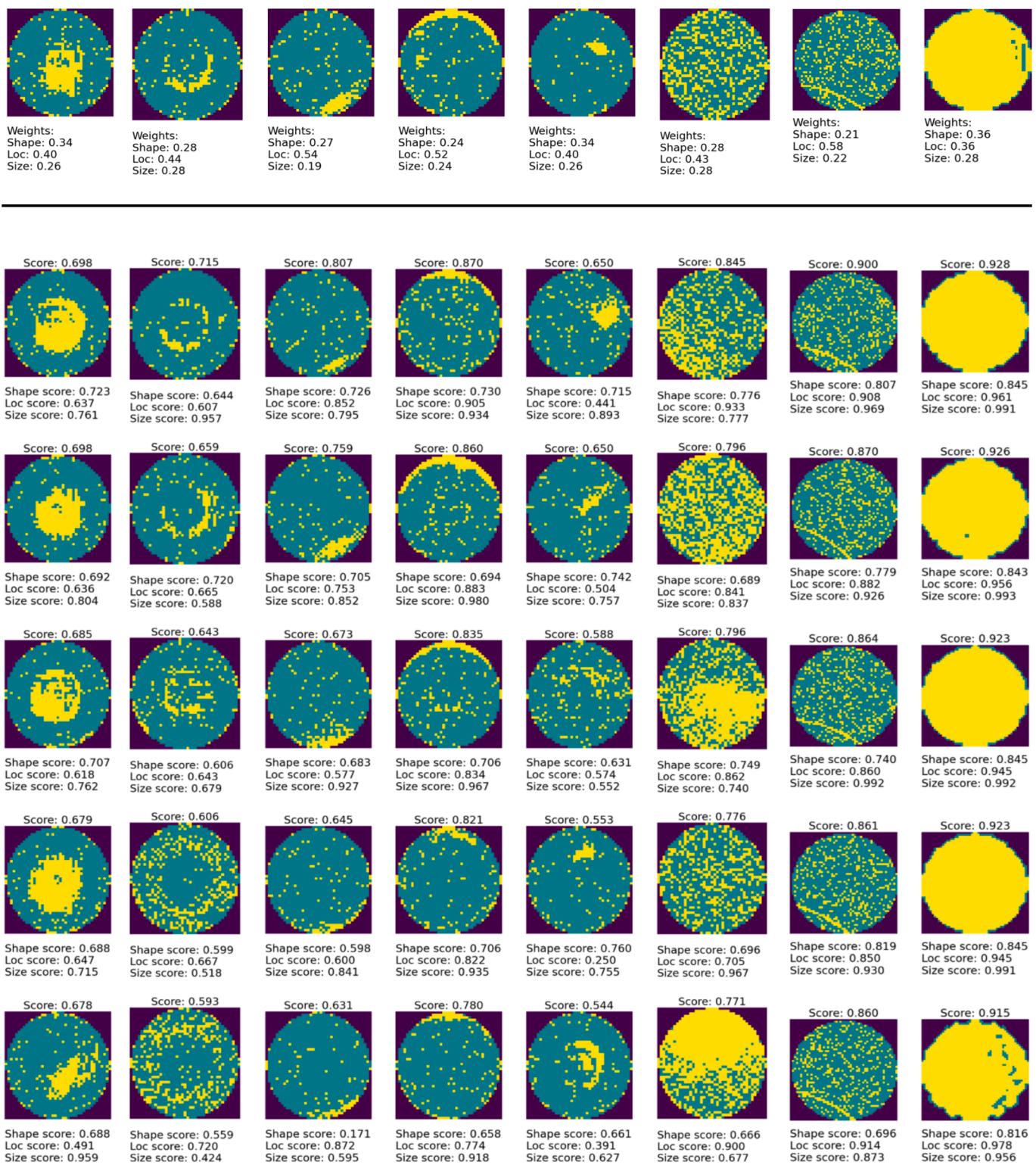


Fig. 3. Similarity ranking result of single pattern.

whereas the Q s for each defect pattern type are shown in the second to sixth rows. We show only the top five Q s in order of overall similarity. Overall, the five Q s are similar to P for all eight defect pattern types, implying that the proposed method works well.

In addition to the three similarities, the weights of the three similarities and the overall similarity for each Q were calculated and presented. The three weights for each defect pattern type are shown in the first row of Fig. 3. These were obtained based on the information entropy theory described in Section 3.4. It is noteworthy that the three weights are determined relatively evenly, which implies that the similarity search considers these three aspects in a balanced manner. Nevertheless, it is noteworthy that $w_{location}$ was the largest among the three weights. This means that the location aspect plays an important role in distinguishing similar patterns to search for Q s that are more similar to a given P . Let us compare two groups: one with small $w_{location}$ values and the other with large $w_{location}$ values. The center, Random, and Near-full types (first, sixth, and last columns in Fig. 3) had smaller $w_{location}$ values, whereas the Edge-local, Edge-Ring, and Scratch types (third, fourth, and seventh columns in Fig. 3) had larger $w_{location}$ values. The three defect pattern types of the first group tended to be located at similar positions on the WBM, implying that the location aspect did not play an important role in discriminating between similar patterns. By contrast, the other three defect pattern types can be located at various positions on the WBM; thus, the location aspect plays a critical role in searching for Q s that are more similar Q s to a given P .

Table 5 presents the results of the computational time taken by the proposed model to compare a single reference WBM with a candidate WBM in this study. The experimental setup was conducted on a server equipped with an AMD EPYC 3rd generation (16Core/32Thread) Model 7313 at 3.00 GHz, 128 MB, 155 W, under a Linux OS, with Sci-kit Learn and NumPy libraries. The results showed that it took approximately 0.3 s to compare a reference WBM with a candidate WBM. Of this time, the image preprocessing phase occupied about 0.26 s, making up the majority of the time consumed. It was observed that the process of calculating and comparing the similarities in shape, size, and position had a relatively lower computational complexity. Moreover, it was noted that the comparison of size and location similarities required a significantly lower computational time (about 0.0008 s in total), indicating that there is no substantial difference in computation time and complexity when comparing similarities based solely on shape.

4.3. Ablation studies

Fig. 4 presents the results of visualizing the top 5 WBMs with the highest similarity scores to the reference WBM after measuring similarity scores using only a single individual similarity to investigate effect of each individual similarity. Fig. 4 show the results of similarity comparison based solely on the shape of defects (a), the location of defects (b), the size of defects (c), and the results of a similarity search method that considers shape, size, and location applying equal weight to the three individual scores (d). In Fig. 4(a), when the reference pattern is Edge Loc, scratch defect patterns with similar curved shapes exhibit high shape similarity scores. This is an expected result since the two patterns share similar shapes. However, it can be observed that the defect locations are not fixed. In Fig. 4(b), scratch defects located at the bottom-left of the reference WBM are often confused with other defect patterns in the same location. This occurs because, when considering only the

defect location, different shaped defects occurring at the same location can also have high similarity scores. Similarly, in Fig. 4(c), when only size similarity is considered, defect patterns such as Edge-Loc and scratch defects with similar sizes are detected. Finally, in Fig. 4(d), the top five images of the center defect class, which is the same defect class as the reference WBM, are presented. However, it can be observed that the third and fifth candidate WBM images differ from the reference WBM image. On the other hand, when the proposed weighting scheme based on information entropy theory is applied, as shown in Fig. 3, defect types that are often confused, such as Edge-Loc and scratch defects, are correctly identified. Additionally, even in the case of center defects that were not appropriately retrieved in Fig. 4(d), it can be confirmed that the proposed method correctly identified similar types of defects up to the fourth image. These results demonstrate the necessity of each module proposed in this study and the importance of a comprehensive approach to similarity measurement and retrieval.

Consequently, through Fig. 4, it is confirmed that the shape, location, and size similarity algorithms in the proposed methodology faithfully find candidate WBMs according to the purpose of each similarity measurement. This indicates that our methodology operates appropriately. However, considering each similarity comparison method independently or assigning equal weights to each similarity score to measure a composite score can retrieve similar WBMs that differ from the reference WBM. To prevent this, it is important to apply the uncertainty estimation and weight assignment method based on the information entropy theory proposed in the study, using the weighted average of each similarity score.

4.4. Comparison to existing methods

This section compares our method with the conventional similarity search method proposed in Wang & Wang (2022), which conducts an experiment using a public dataset, MixedWM38K. Fig. 5(a) and (b) respectively show the comparison of similarity search results for Loc and Scratch defect patterns between the proposed method and the existing method. In these figures, the top (bottom) shows the results obtained by the existing (proposed) method, and the two most left WBM images are the reference WBM (P), while the rest are the WBMs (Q s) found to have high similarity by the proposed and existing methods. Looking at the results in detail, in the case of the traditional similarity search method, when recommending the top 10 WBM images based on similarity scores, although images from the same defect pattern class as P were presented, they mainly depended on shape. In particular, as can be seen in Fig. 5(a), despite the Loc defect pattern located at the top left of the wafer, the existing study selected a WBM at the top right with different location and size as the most similar image. Moreover, as shown in Fig. 5(a), the fourth and eighth Q s from the left have smaller defect patterns than P , showing significant differences in shape, size, and location with the reference WBM in about 6–8 out of the 10 recommended images. Similarly, the results for the Scratch defect type in Fig. 5(b) also show that all 10 images belong to the same defect class, but there are slight differences in the location and size where the defects occurred.

On the other hand, all Q s found using the proposed method show high similarity with the reference WBM in all aspects, such as shape, form, and size. This is because the proposed method independently evaluates similarity regarding shape, location, and size, providing a more precise similarity ranking. Especially, thanks to its ability to

Table 5
Computational complexity of proposed method.

–	Data Preprocessing Process		Similarity & Uncertainty Measurement Process				Total time
	C-Mean Filtering	Tensor Voting	Shape Similarity	Location Similarity	Size Similarity	Uncertainty Estimation	
Calculate Time (ms / WBM)	65.5096	191.1688	40.5684	0.725	0.0615	17.8129	315.8462
Total Time	256.6784		59.1678				

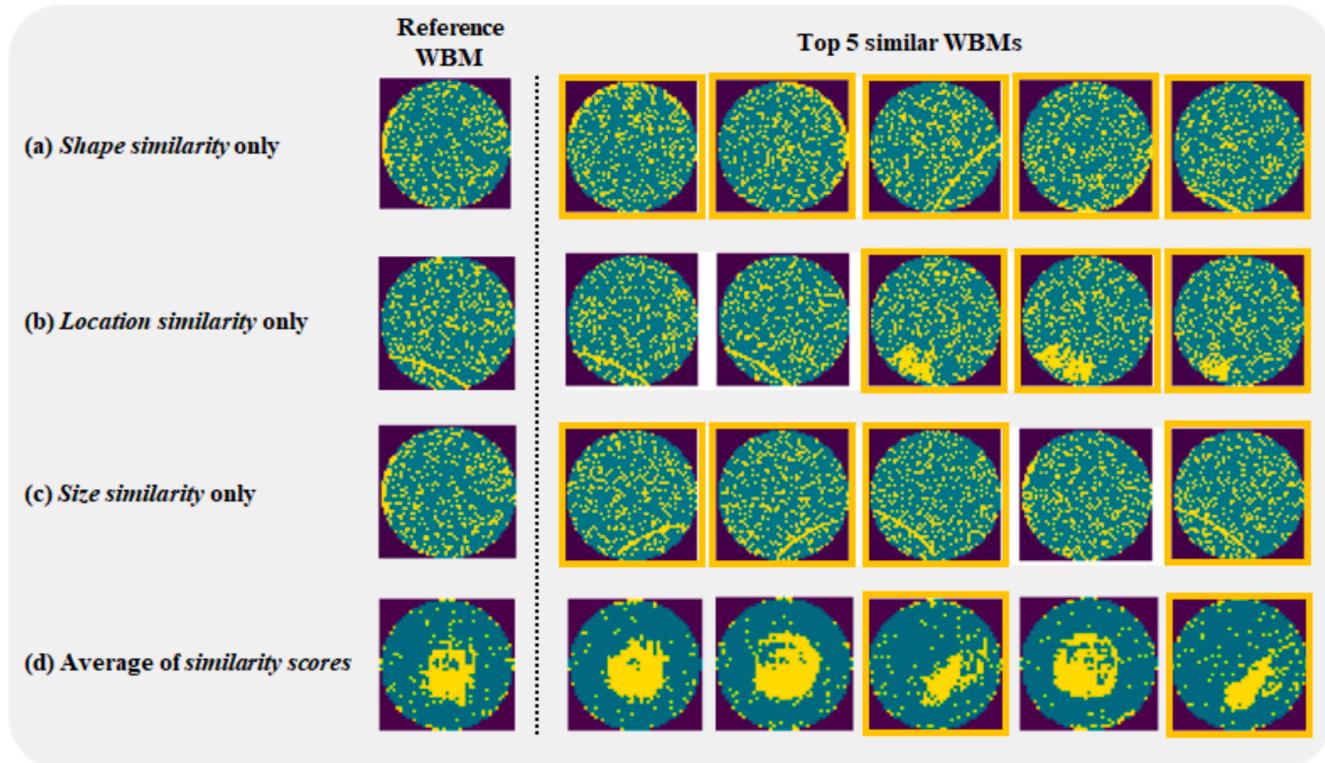


Fig. 4. Similarity search results using each method.

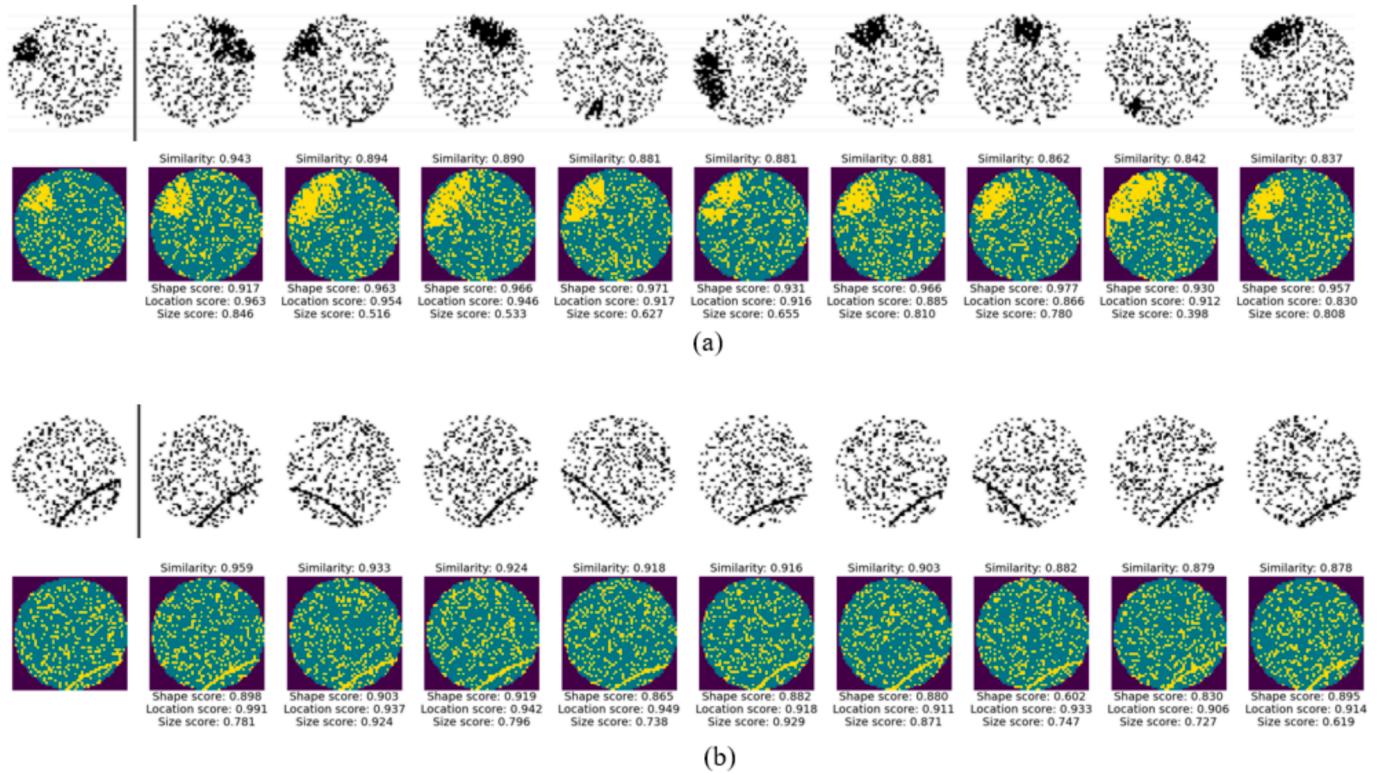


Fig. 5. Comparison of the proposed and existing search methods.

precisely identify spatial features, it excels at determining the exact location and size of each pattern. Through this analysis, this study demonstrates that focusing not only on a single element of defect patterns but on more detailed and specific defect similarity search can

contribute to early detection and cause identification of defects.

5. Concluding remarks

In this study, we propose a similarity ranking method for WBMs that considers the combined aspects of defect pattern shape, location, and size, which have been overlooked in existing similarity search research. To measure shape similarity, we employ tensor voting and a mountain function, while Euclidean distance of the center of mass between two WBM is used for location similarity. Additionally, we incorporate the average radius of defect clusters and the variance in defect counts for size similarity. The overall similarity is calculated using a weighted average, with each weight determined based on the uncertainty of each similarity score using information entropy. By employing these methods, our proposed approach can independently evaluate shape, location, and size characteristics, automatically identify dominant features, and derive integrated overall similarity scores.

The proposed method does not require predefined labels for defects or a training process, enabling swift and accurate comparison of defects even for previously unknown patterns. Moreover, it effectively detects spatial features, allowing for a more nuanced comparison within the same defect pattern class. These achievements are particularly effective for defects like Edge-Loc and Scratch, where location and size information are crucial. However, the proposed method faces challenges when dealing with mixed-pattern defects, where various shapes and sizes of defects coexist on a single wafer, making comparison difficult due to their combined shapes, locations, and sizes. Additionally, the adoption of two unique preprocessing methods to compare defect shapes, locations, and sizes may increase computational complexity. Therefore, future research should focus on developing efficient preprocessing methods, especially for handling diverse types of mixed defects. Furthermore, exploring multimodal learning approaches that integrate defect image and equipment information could enhance convenience for engineers.

CRediT authorship contribution statement

Min-Su Kang: Writing – original draft, Software, Methodology, Formal analysis. **Jin-Su Shin:** Writing – review & editing, Supervision, Investigation. **Dong-Hee Lee:** Writing – review & editing, Funding acquisition, Conceptualization.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request.

Acknowledgement

This work was supported by the National Research Foundation of Korea (NRF) grant funded by the Korea government (MSIT) (No. NRF-2022R1C1C1011743).

References

- Bae, Y., & Kang, S. (2023). Supervised contrastive learning for wafer map pattern classification. *Engineering Applications of Artificial Intelligence*, 126. <https://doi.org/10.1016/j.engappai.2023.107154>
- Basim, G. B., & Moudgil, B. M. (2002). Effect of Soft Agglomerates on CMP Slurry Performance. *Journal of Colloid and Interface Science*, 256(1), 137–142. <https://doi.org/10.1006/JCIS.2002.8352>
- Chen, F.-L., & Liu, S.-F. (2000). A Neural-Network Approach To Recognize Defect Spatial Pattern In Semiconductor Fabrication. In *IEEE Transactions On Semiconductor Manufacturing* (Vol. 13, Issue 3).
- Chen, S., Zhang, Y., Hou, X., Shang, Y., & Yang, P. (2022). Wafer map failure pattern recognition based on deep convolutional neural network. *Expert Systems with Applications*, 209. <https://doi.org/10.1016/j.eswa.2022.118254>
- Gray, R. M. (2011). *Entropy and Information Theory*. Springer Science & Business Media.
- Hsu, C. Y., Chen, W. J., & Chien, J. C. (2020). Similarity matching of wafer bin maps for manufacturing intelligence to empower Industry 3.5 for semiconductor manufacturing. *Computers and Industrial Engineering*, 142. <https://doi.org/10.1016/j.cie.2020.106358>
- Hsu, S. C., & Chien, C. F. (2007). Hybrid data mining approach for pattern extraction from wafer bin map to improve yield in semiconductor manufacturing. *International Journal of Production Economics*, 107(1), 88–103. <https://doi.org/10.1016/j.ijpe.2006.05.015>
- Hwang, J., & Kim, H. (2020). Variational Deep Clustering of Wafer Map Patterns. *IEEE Transactions on Semiconductor Manufacturing*, 33(3), 466–475. <https://doi.org/10.1109/TSIM.2020.3004483>
- Kim, E. S., Choi, S. H., Lee, D. H., Kim, K. J., Bae, Y. M., & Oh, Y. C. (2021). An oversampling method for wafer map defect pattern classification considering small and imbalanced data. *Computers and Industrial Engineering*, 162. <https://doi.org/10.1016/j.cie.2021.107767>
- Kim, T., & Behdinan, K. (2023). Advances in machine learning and deep learning applications towards wafer map defect recognition and classification: a review. *Journal of Intelligent Manufacturing*, 34(8), 3215–3247. <https://doi.org/10.1007/s10845-022-01994-1>. Springer.
- Kong, Y., & Ni, D. (2022). A one-shot learning approach for similarity retrieval of wafer bin maps with unknown failure pattern. *IEEE Transactions on Semiconductor Manufacturing*, 35(1), 40–49. <https://doi.org/10.1109/TSIM.2021.3123290>
- Lee, H., Lee, J., & Kim, H. (2023). Semi-supervised learning for simultaneous location detection and classification of mixed-type defect patterns in wafer bin maps. *IEEE Transactions on Semiconductor Manufacturing*, 36(2), 220–230. <https://doi.org/10.1109/TSIM.2023.3264279>
- Lee, J. H., Moon, I. C., & Oh, R. (2021). Similarity search on wafer bin map through nonparametric and hierarchical clustering. *IEEE Transactions on Semiconductor Manufacturing*. <https://doi.org/10.1109/TSIM.2021.3102679>
- Liao, C. S., Hsieh, T. J., Huang, Y. S., & Chien, C. F. (2014). Similarity searching for defective wafer bin maps in semiconductor manufacturing. *IEEE Transactions on Automation Science and Engineering*, 11(3), 953–960. <https://doi.org/10.1109/TASE.2013.2277603>
- Nag, S., Makwana, D., Mittal, S., & Mohan, C. K. (2022). WaferSegClassNet - A light-weight network for classification and segmentation of semiconductor wafer defects. *Computers in Industry*, 142. <https://doi.org/10.1016/j.compind.2022.103720>
- Nakazawa, T., & Kulkarni, D. V. (2018). Wafer map defect pattern classification and image retrieval using convolutional neural network. *IEEE Transactions on Semiconductor Manufacturing*, 31(2), 309–314. <https://doi.org/10.1109/TSIM.2018.2795466>
- Pan, L., & Deng, Y. (2020). Probability transform based on the ordered weighted averaging and entropy difference. *International Journal of Computers, Communications and Control*, 15(4). <https://doi.org/10.15837/IJCCC.2020.4.3743>
- Park, S., Jang, J., & Kim, C. O. (2021). Discriminative feature learning and cluster-based defect label reconstruction for reducing uncertainty in wafer bin map labels. *Journal of Intelligent Manufacturing*, 32(1), 251–263. <https://doi.org/10.1007/s10845-020-01571-4>
- Shinde, P. P., Pai, P. P., & Adiga, S. P. (2022). Wafer defect localization and classification using deep learning techniques. *IEEE Access*, 10, 39969–39974. <https://doi.org/10.1109/ACCESS.2022.3166512>
- Medioni, G., Tang, C.-K., & Lee, M.-S. (2000). Tensor Voting: Theory and Applications. In *Proceedings of RFIA*, 2000.
- Wang, R., & Wang, S. (2022). Tensor Voting Based Similarity Matching of Wafer Bin Maps in Semiconductor Manufacturing. *2022 5th International Conference on Data Science and Information Technology, DSIT 2022 - Proceedings*. <https://doi.org/10.1109/DSIT55514.2022.9943882>
- Wang, R., & Wang, S. (2023). Similarity searching for fault diagnosis of defect patterns in wafer bin maps. *Computers & Industrial Engineering*, 185, Article 109679. <https://doi.org/10.1016/j.CIE.2023.109679>
- Wu, M. J., Jang, J. S. R., & Chen, J. L. (2015). Wafer map failure pattern recognition and similarity ranking for large-scale data sets. *IEEE Transactions on Semiconductor Manufacturing*, 28(1), 1–12. <https://doi.org/10.1109/TSIM.2014.2364237>
- Xu, Q., Yu, N., & Essaf, F. (2022). Improved wafer map inspection using attention mechanism and cosine normalization. *Machines*, 10(2). <https://doi.org/10.3390/machines10020146>
- Yu, J., Shen, Z., & Wang, S. (2021). Wafer map defect recognition based on deep transfer learning-based densely connected convolutional network and deep forest. *Engineering Applications of Artificial Intelligence*, 105. <https://doi.org/10.1016/j.engappai.2021.104387>
- Yu, N., Gong, X. Q., Wang, H. L., & Lin, J. (2021). Wafer bin map inspection based on DenseNet. *Journal of Central South University*, 28(8), 2436–2450. <https://doi.org/10.1007/s11771-021-4778-7>
- Yu, N., Xu, Q., & Wang, H. (2019). Wafer defect pattern recognition and analysis based on convolutional neural network. *IEEE Transactions on Semiconductor Manufacturing*, 32(4), 566–573. <https://doi.org/10.1109/TSIM.2019.2937793>