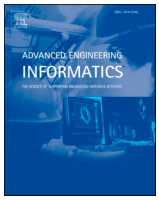
Advanced Engineering Informatics 60 (2024) 102540

Contents lists available at ScienceDirect

Advanced Engineering Informatics

journal homepage: www.elsevier.com/locate/aei



Full length article

# Development of a spatial dimension-based taxonomy for classifying the defect patterns in a wafer bin map

Seung-Hyun Choi [a], Dong-Hee Lee [b], Eun-Su Kim [c], Young-Mok Bae [a,c], Young-Chan Oh [c], Kwang-Jae Kim [a,*]

[a] Department of Industrial and Management Engineering, Pohang University of Science and Technology, 77, Cheongam-ro, Nam-gu, Pohang, Gyung-buk 37673, Republic of Korea
[b] Department of Systems Management Engineering, Sung Kyun Kwan University, 2066, Seobu-ro, Jangan-gu, Suwon, Gyeong-gi 16419, Republic of Korea
[c] SK Hynix Inc., 337, Jikji-daero, Heungdeok-gu, Cheongju, Chung-buk 28436, Republic of Korea

ARTICLE INFO

*Keywords:*
Semiconductor manufacture
Wafer bin map
Defect pattern
Classification
Taxonomy

ABSTRACT

A wafer bin map (WBM) represents the locational information of defective chips on the wafer. The spatial correlation of defects on the wafer provides crucial information for the root cause diagnosis of defects in wafer fabrication. The spatial correlation is classified as a defect pattern for efficient diagnostics. A defect pattern taxonomy should be defined in advance for coherent classification of defect patterns. Various taxonomies are used in previous studies, but they share common limitations in that the differentiation among defect patterns is unclear, the set of predefined defect patterns is insufficient, and they cannot accommodate newly-emerged defect patterns. A concept of spatial dimension-based defect pattern taxonomy and its development procedure are proposed. Defect patterns are defined by three spatial dimensions, namely, *Shape*, *Size*, and *Location*. The development procedure is applied to a major NAND flash memory semiconductor manufacturer for two years. Results show that spatial dimension-based taxonomy can improve the performance of the defect pattern classification system by alleviating common existing limitations. Moreover, meaningful defect patterns for diagnostics are retained through the engineers' involvement in the development procedure.

## 1. Introduction

The semiconductor manufacturing process comprises several hundred steps to fabricate hundreds to thousands of chips on a wafer. After the wafer fabrication, electrical probe tests are conducted to determine whether a chip has the proper quality for a packaging process. Defective chips (referred to as defects) fail at least one of the probe tests. The test results are recorded by the binary bin code, which expresses whether a chip is functional or defective [1], or by the categorical bin code, which expresses the type of failed tests [2]. A wafer bin map (WBM) arranges the position of chips and their bin codes to represent the probe test results of a wafer.

A WBM may comprise random defects, non-random defects, or both. Determining whether a given defect (*i.e.*, a specific defect) is a random or non-random defect is generally impossible due to the large number of chips on a wafer and the various factors influencing chip quality throughout the wafer fabrication process [3]. However, in the literature, if a group of defects exhibits a spatial correlation on the WBM, the defects are attributed to an assignable cause and categorized as non-random defects [4]. For instance, line-shaped and circle-shaped defects are attributed to the hardening of the pad and the non-uniformity problem during chemical–mechanical planarization, respectively [5]. On the other hand, defects that are randomly distributed on a WBM without a spatial correlation are considered random defects.

The types of spatial correlation among defects are referred to as defect patterns. They manifest themselves in various forms, such as a *Line*, *Circle*, and *Ring* [6]. The relationship between the wafer fabrication failure and defect patterns is identified through manual monitoring [7], statistical analysis [8], pattern mining [9], and rule-based inference [10].

The fundamental assumption of diagnostics is that defect patterns are coherently classified, but this assumption is invalid in practice. In manual classification, a single engineer's long-term repeatability is less than 95 %, and the agreement between engineers is less than 45 % [11].

* Corresponding authors.
*E-mail addresses:* seunghyun.choi@postech.ac.kr (S.-H. Choi), dhee@skku.edu (D.-H. Lee), eunsu381.kim@sk.com (E.-S. Kim), ymbae@postech.ac.kr (Y.-M. Bae), youngchan.oh@sk.com (Y.-C. Oh), kjk@postech.ac.kr (K.-J. Kim).

https://doi.org/10.1016/j.aei.2024.102540
Received 28 February 2023; Received in revised form 13 March 2024; Accepted 9 April 2024
Available online 12 April 2024
1474-0346/© 2024 Elsevier Ltd. All rights reserved.

Automatic classification models have recently replaced manual classification [1,5,9,12–16], but incoherently classified datasets degrade the performance of the classification models [17]. The performance of the models heavily depends on the quality of input data [18,19]; thus, the acquisition of a coherently classified WBM dataset is important.

A defect pattern taxonomy refers to a set of predefined defect patterns, and a WBM dataset is commonly classified using the taxonomy in practice. For instance, non-random defects in the WM811K dataset [20] are classified as one of the nine defect patterns of the WM811K taxonomy, namely, *Center, Donut, Location, Edge-Loc, Edge-Ring, Scratch, Random, None,* and *Near-Full*. Fig. 1 depicts the WM811K taxonomy, where functional and defective chips are indicated in gray and white colors, respectively.

As mentioned earlier, a WBM can be represented using either the binary bin code or the categorical bin code. However, in the literature, defect pattern taxonomies have been developed predominantly for the case of binary bin code. This is because a defect pattern taxonomy aims to classify defect patterns based solely on the spatial features of defects [21–23], where the binary bin code proves sufficient for representing these spatial features. Moreover, a WBM with the categorical bin code can be easily transformed into one with the binary bin code [24]. Hence, without the loss of generality, this study assumes that the WBM comprises the binary bin code.

A poorly defined defect pattern taxonomy is one of the reasons for incoherent classification [1,25]. For coherent classification, the defect pattern taxonomy should possess three desirable properties of a general taxonomy, namely, robustness, comprehensiveness, and extendibility [26]. The robustness refers to a clear differentiation between the defect patterns. Accurately classifying spatial correlations requires differentiation criteria that facilitate clear differentiation between defect patterns. The comprehensiveness means that defect patterns are sufficiently defined to classify spatial correlations among defects in the WBM dataset. When defects show spatial correlations not defined in a defect pattern taxonomy, there is a risk of neglecting these spatial correlations during diagnostics or misclassifying them as a different defect pattern within the taxonomy. This may lead to inaccurate identification of the root cause [27]. The extendibility means the capability to add newly-emerged defect patterns in the future. A new defect pattern may indicate a previously unidentified root cause of wafer fabrication failure that requires investigation [9]. Consequently, the defect pattern taxonomy should be adaptable to include new patterns, facilitating ongoing root cause analysis in the wafer fabrication process.

Data quality has grown to prominence in smart manufacturing systems [28], and previous studies assume that the WBM dataset is coherently classified. However, previous taxonomies used in automatic defect pattern classification studies lack at least one of the three desirable properties, which are critical for coherent defect pattern classification. The research gap between the acquisition of a coherently classified WBM dataset and defect pattern classification lies in the absence of a defect pattern taxonomy that possesses all three desirable properties.

This paper presents the concept and a development procedure of a spatial-dimension-based WBM taxonomy that possesses the three desirable properties of a general taxonomy and integrates domain expertise. The proposed taxonomy comprises three spatial dimensions, namely, *Shape*, *Size*, and *Location*. The dimension levels represent the spatial features of the defects in each dimension. Defect patterns are defined by the combination of dimension levels of the three spatial dimensions. The robustness, comprehensiveness, and extendibility are possessed through the methodical definition of defect patterns on the three-dimensional structure. We applied the proposed procedure to the development of a defect pattern taxonomy of NAND flash memory for two years and demonstrated the advantages of the developed taxonomy.

This paper is organized as follows. A review of the existing taxonomies for defect pattern classification is provided in Section 2. The spatial dimension-based taxonomy and its development procedure are proposed in Section 3. The proposed taxonomy is validated through manual and automatic defect pattern classification in Section 4. Finally, concluding remarks and future research issues are discussed in Section 5.

## 2. Literature review

The taxonomies developed in the early stages of defect pattern classification studies comprise only a few defect patterns. For instance, Cunningham and MacKinnon [4] defined only one defect pattern, *Scratch*. Chih-Hsuan Wang *et al.* [29] considered three defect patterns: *Zone*, *Line*, and *Ring*. *Line* corresponds to *Scratch* in Cunningham and MacKinnon [4] taxonomy. Similar taxonomies were proposed in other studies [9,30–36].

The taxonomies in the aforementioned category effectively classify a few defect patterns. However, the spatial correlation within defects becomes complicated because the wafer size is enlarged, and the chip size is decreased [25]. Thus, the taxonomies in modern wafer fabrication processes are extended to accommodate more complicated defect patterns through two approaches. These two approaches will henceforth be referred to as ‘pattern subdivision’ and ‘pattern addition’.

One defect pattern may be related to several root causes [8,37], and the pattern subdivision approach mitigated such relationships by subdividing the existing defect pattern into more detailed ones. For instance, Chen and Liu [38] subdivided ring-shaped and scratch-shaped defect patterns; ring-shaped defect patterns were subdivided in accordance with the distance from the wafer center and scratch-shaped defect patterns were subdivided by their location. Bae *et al.* [39] subdivided scattered-shaped defect patterns by their location, namely, *random scattering*, *radial scattering*, and *scattering in one corner*. Li and Huang [3] and Liukkonen and Hiltunen [40] used a similar approach.

The pattern subdivision approach enhances the robustness by

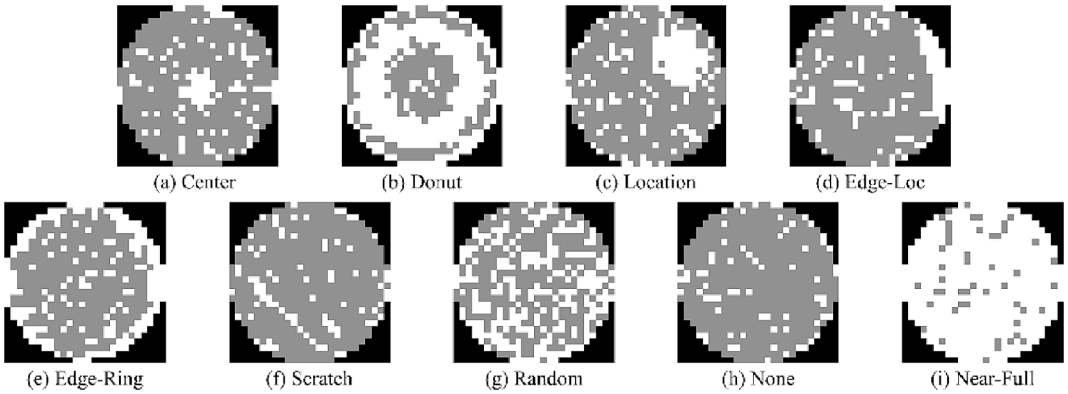



**Fig. 1.** WM811K taxonomy.

improving the differentiation between defect patterns. However, it has the disadvantage of limiting the taxonomy's comprehensiveness, as pattern subdivision can only extend the taxonomy based on existing defect patterns. For instance, the taxonomy introduced in Chen and Liu [38] cannot accommodate scattered-shaped defect patterns through subdivision because only ring- and scratch-shaped defect patterns existed in the taxonomy from the beginning. The extendibility of the taxonomy is limited in that this approach only allows the subdivision of existing defect patterns.

In order to accommodate a newly-emerged defect pattern, the pattern addition approach added a newly-emerged defect pattern to an existing defect pattern taxonomy [41–49]. For instance, *Hat* pattern (Fig. 2-(a)) was added in Cheng *et al.* [50], Ooi *et al.* [49], and Ooi *et al.* [45], and *Checker-board* (Fig. 2-(b)) was added in Chien *et al.* [48] and S.-C. Hsu and Chien [43] to distinguish new defect patterns from existing defect patterns (*i.e.*, *Ring*, *Line*, and *Zone*). The addition of a new defect pattern can accommodate newly-emerged defect patterns, wherein the comprehensiveness and extendibility are possessed.

However, the pattern addition approach has the disadvantage of compromising the taxonomy's robustness. Adding a new defect pattern in an ad hoc manner may lead to inconsistent differentiation criteria between defect patterns, resulting in unclear differentiation between defect patterns. For example, in the WM811K taxonomy, *Donut* indicates a shape of defect pattern, while *Center* indicates a location of defect pattern. Thus, the donut-shaped defects in the center location can be classified as one of the two defect patterns depending on the engineer's subjective judgment.

The research gap is evident in the existing pattern subdivision and pattern addition approaches, which do not possess all three desirable properties of taxonomy as summarized in Table 1. To address this research gap, this study proposes a spatial dimension-based taxonomy and its development procedure, aimed at embodying the three desirable properties as shown in the rightmost column of Table 1. Clear differentiation between dimension levels in each spatial dimension enhances the taxonomy's robustness, while the addition and modification of dimension levels in spatial dimensions contribute to the comprehensiveness and extendibility. Moreover, domain expertise is integrated into the taxonomy through the involvement of engineers in the defect pattern development procedure.

## 3. Development of the spatial dimension-based taxonomy

This section describes the concept of the spatial dimension-based taxonomy and its development procedure with an application to a NAND flash memory semiconductor manufacturer in Korea. The WBM dataset comprises 1793 WBMs, each consisting of approximately 500 chips, manufactured in 2020. Each chip's binary bin codes and x- and y-positions on a WBM (integer, $[0, 25]^2$) are recorded in the dataset, along with the wafer ID. The WBM size is identical at $26 \times 26$, in which a $26 \times 26$ matrix is made for each wafer; the row and column of matrix correspond to the y-axis and x-axis of WBM, respectively. The matrices were converted into black-and-white images for manual inspection.

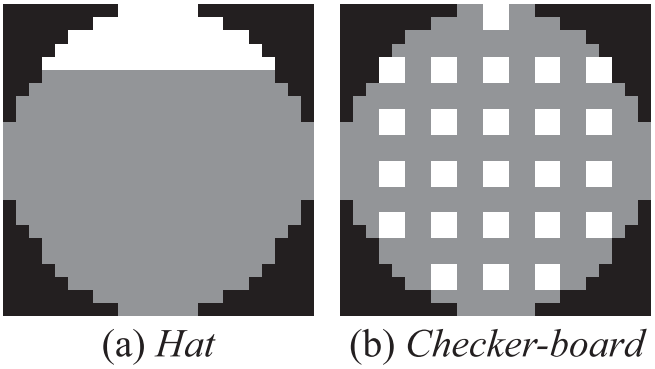



**Fig. 2.** Reproduced example of *Hat* and *Checker-board* patterns.

**Table 1**
Summary of existing WBM defect pattern taxonomies.

| | Pattern subdivision | Pattern addition | Proposed taxonomy |
|---|---|---|---|
| Robustness | ✓ | | ✓ |
| Comprehensiveness | | ✓ | ✓ |
| Extendibility | △ | ✓ | ✓ |
| Related studies | [3,38–40] | [41–50] | – |

✓: possess, △: restrictively possess

Inspired by the general taxonomy development procedure proposed by Nickerson *et al.* [26], the two-stage procedure is proposed. In Stage 1, spatial dimensions and their dimension levels are defined. Spatial features of non-random defects are extracted from a thorough review of existing defect pattern taxonomies as well as a real WBM dataset. The features extracted are then organized into dimensions and levels within those dimensions. In Stage 2, defect patterns for the taxonomy are defined by combining the dimension levels. Some defect patterns are merged or removed so that only meaningful defect patterns for diagnosing wafer fabrication failures are retained.

Three practicing engineers, each with over five years of experience in wafer quality monitoring and control, were involved in the taxonomy development procedure. They played a crucial role in Stages 1 and 2. In Stage 1, the engineers assisted in deciding whether to merge or subdivide dimension levels. They also contributed to establishing differentiation criteria between these levels. In Stage 2, their involvement extended to screening dimension level combinations, identifying those sharing the same root cause and removing irrelevant combinations. The engineers' involvement has been essential, integrating domain expertise to ensure the proposed taxonomy is both practically meaningful and effective.

### 3.1. Stage 1: Definition of dimensions and dimension levels

The spatial dimension-based taxonomy comprises multiple dimensions, each containing dimension levels that are mutually exclusive and collectively exhaustive. The main spatial features of non-random defects should be considered in defining the dimensions. In manual visual inspection of WBM, various spatial features such as shape, size, location, color, and orientation can be used to analyze non-random defects [4,31,32,38,42,51,52]. However, in the context of defect pattern taxonomy, color and orientation have been rarely considered. Color is not used because defect pattern taxonomies typically employ the binary bin code, as mentioned in Section 1. Orientation is disregarded because the same defect patterns with different orientations (*i.e.*, rotation-variant defect patterns) are known to be associated with the identical wafer fabrication failure [33]. Therefore, the shape, size, and location are considered to be sufficient in defining the meaningful spatial dimensions of defect patterns. Thus, the proposed taxonomy is developed based on the dimensions of shape, size, and location. The *Shape* dimension manifests the form of a defect pattern, such as a scratch or arc. The *Size* dimension is a scale indicator of the shape, so *Size* manifests itself in conjunction with *Shape*. For instance, length is a size measurement of a scratch shape. The *Location* dimension indicates the site of a defect pattern on a WBM.

Dimension levels represent the spatial feature of the defect patterns in each dimension. Dimension levels were defined on the basis of spatial features from the collected WBM dataset for each dimension. The spatial features of each dimension are identified and subsequently organized into dimension levels. Identification and organization are iteratively performed until every dimension level is unique within its dimension, and all spatial features of non-random defects in the WBM dataset are

represented by the dimension levels. Each dimension level is differentiated using an objective criterion, which formalizes engineers' domain knowledge and subjective judgment into a definitive form [49]. In addition, the threshold level between dimension levels should be established to effectively aid in root cause diagnosis.

The spatial features of the *Size* dimension were identified and organized in conjunction with those of the *Shape* dimension. Meanwhile, the spatial features of the *Location* dimension were identified and organized independently. Additionally, the criteria for differentiating dimension levels were established based on the number of defects, considering that the WBM dataset comprises wafers with the same number of chips. These criteria can be defined either by a relative scale, such as the percentage of total chips on a wafer [43], or by a physical scale, such as the length measured in centimeters [38].

The levels of the *Shape* dimension are *Scratch*, *Cluster*, *Ring*, and *Random*. They represent the defect patterns' forms, as indicated in those names. *Scratch* and *Cluster* were differentiated on the basis of the defect pattern's maximum width. The pattern is classified as *Scratch* when the maximum width is four chips. Otherwise, the pattern is classified as *Cluster*. The engineers determined the threshold level (*i.e.*, four chips) because the scratch-shaped pattern was attributed to the physical damage during the wafer fabrication. Thus, its width is generally lower than the four chips. Similar to taxonomies found in [20,36,42,44], the proposed taxonomy incorporates the "*Random*" level in the *Shape* dimension to characterize the defect pattern of a WBM without spatial correlation among defects.

The levels of the *Size* dimension were defined in conjunction with the shape of the defect pattern. The levels of the *Size* dimension for *Cluster* are *Big* and *Small*. The pattern is defined as *Cluster-Big* if the defect pattern comprises more than 40 chips and defined as *Cluster-Small* otherwise. The threshold level was determined at 40 chips, which is approximately 8 % of the total number of chips in a wafer, because the engineers believe that a cluster of more than 40 defective chips may degrade the wafer's overall quality. Similarly, the threshold for the *Size* dimension in the *Random* pattern was set at 150 chips, approximately 30 % of the total number of chips in a wafer. The threshold between *Scratch-Long* and *Scratch-Short* was determined at 10 chips, approximately 75 % of the wafer's radius. This decision took into account the rotation of the wafer during fabrication, which contributes to the formation of line-shaped defect pattern. For the *Ring* pattern, the *Size* dimension levels were established to encompass representative angles of 90°, 180°, 270°, and 360°. Each of the representative angle is allowed a leeway of ± 45° except *Ring-Whole*, where the arc angle is defined to be between 315° and 360°. Table 2 presents the dimension level description and representative images for the *Shape* and *Size* dimensions.

*Edge*, *Center*, *Others*, and *Scattered* levels were defined by the relative location of the defect patterns on the WBM in the *Location* dimension. The *Edge* level indicates that the defect pattern contacts with the wafer edge. Without that contact with the wafer edge, the *Center* level manifests that the defect pattern's center point is on the wafer center. The *Others* level manifests if the center point is not at the center. Further detailed location levels (*e.g.*, *Upper-right*, *Lower-left*) exist at the early stage of dimension level organization, but they were merged into *Others* because the engineers claimed that the rotational-variant features were not critical for diagnostics. The *Scattered* level was defined to represent the *Random* level location. This dimension level indicates that the defects are scattered across the wafer. Descriptions and representative images of the *Location* dimension are provided in Table 3.

### 3.2. Stage 2: Definition of defect patterns

Defect patterns were defined by the combination of the dimension levels in the three dimensions as *Shape–Size–Location* in the proposed taxonomy. For instance, *Scratch–Long–Center* refers to a scratch-shaped defect pattern with a length exceeding 10 chips on the wafer center. Out of 40 possible combinations, 26 physically feasible combinations

**Table 2**
Description of the *Shape* and *Size* dimensions.

| Dimension | *Shape* | | *Size* | | Image |
|---|---|---|---|---|---|
| | Level | Description | Level | Description | |
| Dimension Levels | *Scratch* | The defect pattern's shape is a thin line, and its maximum width is four chips. | *Long* | The length of the scratch exceeds 10 chips. |  |
| | | | *Short* | The length of the scratch is 10 chips at most. |  |
| | *Cluster* | The defect pattern's shape is a lump, and its maximum width exceeds four chips. | *Big* | The number of consisting chips exceeds 40. |  |
| | | | *Small* | The number of consisting chips is 40 at most. |  |
| | *Ring* | The defect pattern takes the form of an arc or a ring. | *One-quarter* | The arc angle is between 45° and 135°. |  |
| | | | *Half* | The arc angle is between 135° and 225°. |  |
| | | | *Three-quarters* | The arc angle is between 225° and 315°. |  |
| | | | *Whole* | The arc angle is between 315° and 360°. |  |
| | *Random* | The defective chips do not show a spatial correlation. | *Frequent* | The number of consisting chips exceeds 150. |  |
| | | | *Infrequent* | The number of consisting chips is 150 at most. |  |

**Table 3**
Description of the *Location* dimension.

| Level | Description | Image |
|---|---|---|
| *Edge* | The defect pattern contacts the wafer edge. |  |
| *Center* | The defect pattern's center point is on the center and does not contact the wafer edge. |  |
| *Others* | The defect pattern's center point is between the center and edge of the wafer and does not contact the edge. |  |
| *Scattered* | The defects are scattered across the wafer. |  |

remained. For instance, the *Ring* level from the *Shape* dimension cannot be combined with the *Scattered* level in the *Location* dimension.

The engineers reviewed the remaining combinations to reflect the domain knowledge on wafer fabrication failure, and Table 4 summarizes the results. Consequently, 10 out of the 26 defect patterns passed the practicality check. The engineers claimed that the root cause of the ring-shaped pattern is irrelevant to its size, wherein the related dimension levels in the *Size* dimension were merged into *All*. The *Others* and *Center* levels in the *Location* dimension were removed because the engineers knew from experience that the ring-shaped pattern occurred only at the edge of wafers. The engineers also argued that the root causes of cluster-shaped defects at the wafer center are similar. The *Cluster–Big–Center* and *Cluster–Small–Center* were merged into *Cluster–All–Center*. The location of the scratch-shaped defect pattern was also considered insignificant. Thus, the related dimension levels in the *Location* dimension were merged into *All*.

## 4. Validation of the taxonomy

In this section, the proposed taxonomy is validated through manual and automatic defect pattern classification. The WM811K taxonomy was selected for the benchmark taxonomy. The WM811K taxonomy is the most widely employed taxonomy for the validation of automatic defect pattern classification models [1,14–16,20,37,53–58]. Also, the WM811K taxonomy is the only taxonomy that is provided with a substantial labeled WBM dataset [20], which can be served as a reference for manual classification. The WBM dataset used in this section consists of 3,209 WBMs of NAND flash memory product family. The data format and WBM size were identical to those used in the Section 3.

**Table 4**
Defect patterns of the proposed taxonomy with representative images.

| Feasible combinations of dimension levels | | | Review results | Image |
|---|---|---|---|---|
| *Shape* | *Size* | *Location* | | |
| *Scratch* | *Long* | *Edge* | Merged to *Scratch–Long–All* | |
| | | *Center* | | |
| | | *Others* | | |
| | *Short* | *Edge* | Merged to *Scratch–Short–All* | |
| | | *Center* | | |
| | | *Others* | | |
| *Cluster* | *Big* | *Edge* | *Cluster–Big–Edge* | |
| | | *Others* | *Cluster–Big–Others* | |
| | *Small* | *Center* | Merged to *Cluster–All–Center* | |
| | | *Center* | | |
| | | *Edge* | *Cluster–Small–Edge* | |
| | | *Others* | *Cluster–Small–Others* | |
| *Ring* | *One-quarter* | *Edge* | Merged to *Ring–All–Edge* | |
| | *Half* | *Center* | Removed | |
| | *Three-quarters* | *Others* | Removed | |
| | *Whole* | | | |
| *Random* | *Frequent* | *Scattered* | *Random –Frequent–Scattered* | |
| | *Infrequent* | *Scattered* | *Random –Infrequent–Scattered* | |

### 4.1. Assessment of the robustness and the comprehensiveness

The robustness and comprehensiveness of the proposed taxonomy were assessed through manual classification in this section. Two raters independently classified the same WBM dataset using the proposed and WM811K taxonomies. The classification coherence between the raters was compared, and the incoherent classification results were reviewed to assess the robustness. Then, the number of WBMs classified as defect patterns without spatial correlation among defects (*i.e.*, *Random –Infrequent–Scattered* and *None*) in each taxonomy was compared to assess the comprehensiveness. A WBM is classified as them when the taxonomy cannot represent the spatial correlation among defects on a WBM. Thus, the number of WBMs classified as random defects is presumed to be an indicator of the comprehensiveness.

First, the inter-rater agreement was better in the proposed taxonomy. The inter–rater agreement was assessed by proportion agreement and Cohen's Kappa [59]. A proportion agreement is a proportion of WBMs that two raters classified as the same defect pattern. Kappa is a chance-corrected index that ranges from −1 to +1, where +1 indicates perfect agreement and 0 indicates that the agreement is the same as chance agreement. The inter-rater agreement is considered better as the proportion agreement and the Kappa value approaches 100 % and +1, respectively, in our context. Table 5 shows the inter-rater agreement results. Both indices were higher in the proposed taxonomy, which indicates that the two raters had more consistent differentiation results in the proposed taxonomy.

Second, a review of the disagreement WBMs revealed that the differentiation criteria were clearer in the proposed taxonomy. Proportions of disagreements were 20.2 % and 34.7 % in the proposed and WBM taxonomies, respectively (Table 5), and the disagreement WBMs were reconciled to confirm the defect pattern. Fig. 3 shows five examples of ambiguous WBM images that were successfully differentiated in the proposed taxonomy, but not in the WM811K taxonomy. Fig. 3-(a) was clearly classified as a *Cluster–Big–Edge* pattern in the proposed taxonomy, but could be classified as *Center*, *Edge-Loc*, or *Location* depending on the rater's subjective judgment using the WM811K taxonomy. The ambiguity arose due to the unclear differentiation criteria of the locational information in the WM811K taxonomy. Fig. 3-(b) was clearly classified as *Cluster–Small–Edge* in the proposed taxonomy because the differentiating criterion between *Cluster* and *Scratch* is objectively defined. However, the corresponding classification would be either *Scratch* or *Edge-Loc* in the WM811K taxonomy depending on the raters' subjectiveness. Fig. 3-(c) and (d) were clearly classified as *Cluster-All-Center* in the proposed taxonomy, whereas they could be classified as *Location* or *Donut* in the WM811K taxonomy.

Third, more spatial correlations among defects can be classified in the proposed taxonomy. The column of "Original" in Table 6 shows the number of WBMs in the reconciled result. The number of WBMs with *None* pattern in the WM811K taxonomy was more than twice that of *Random–Infrequent–Scattered* in the proposed taxonomy, indicating that more spatial correlations among defects can be represented in the proposed taxonomy.

In the taxonomy development procedure, spatial features from the WBM dataset were identified and then organized into the three spatial dimensions. Thus, the proposed taxonomy did not omit the spatial features in the dataset. For instance, Fig. 4-(a) was clearly classified as

**Table 5**
Inter–rater agreement results.

| Taxonomy | Proposed | WM811K |
|---|---|---|
| Proportion agreement | 79.8 % | 65.3 % |
| Kappa | 0.713 | 0.530 |

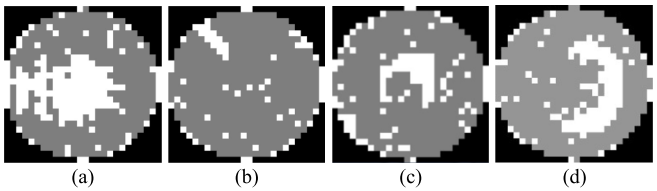



**Fig. 3.** Examples of ambiguous WBM images.

*Ring–All–Edge* in the proposed taxonomy. However, the *Size* dimension was not considered in the WM811K taxonomy; thus, the ring-shaped pattern with an angle less than 180° is undetermined in the WM811K taxonomy. Fig. 4-(b) was clearly classified as *Scratch–Short–All* in the proposed taxonomy, whereas it was not clearly defined in the WM811K taxonomy because the *Size* dimension of the *Line* pattern is not considered in the WM811K taxonomy.

### 4.2. Demonstration of the extendibility

This section demonstrates the extendibility of the proposed taxonomy. The taxonomy developed in Section 3 had been used in an automatic defect pattern classification system since 2020. A new technology for wafer fabrication was adopted in 2021, and new defect patterns emerged since then. The defect pattern classification system cannot classify the newly-emerged patterns. Thus, the proposed taxonomy was extended to accommodate them.

Two types of new defect patterns were derived from the unclassified WBMs. Fig. 5 illustrates the Type 1 defect pattern. The spatial features of Type 1 were identified in accordance with the three dimensions and compared with the existing dimension levels. The spatial features of Type 1 were mapped to existing dimension levels: *Ring* for *Shape*, *Three-quarters* for *Size*, and *Others* for the *Location* dimension. The combination of the three levels had been removed in the second stage of the development procedure. However, the *Ring–Three-quarters–Others* pattern was added, given that a new defect pattern occurred.

Fig. 6 illustrates the Type 2 defect pattern. The spatial features of Type 2 were also identified and organized into three spatial dimensions. The dimension level of *Location* was defined as the existing level *Center*, and a new dimension level, namely, *Double-cluster*, for the *Shape* dimension was defined. Given that the *Size* variation of *Double-cluster* was undetected in the dataset, a single level of *All* was defined in *Size*. Combining the dimension levels, the *Double-cluster–All–Center* pattern was added.

*Double-cluster* was consequently added in the *Shape* dimension, and two defect patterns were added to the existing taxonomy. Newly-observed spatial features were identified and added to the dimension levels; thus, the extendibility of the proposed taxonomy was realized.

### 4.3. Assessment of the classification coherence

The coherence of manual classification was assessed using the classification models in this section. The coherence in this section is focused on long-term repeatability. The performance of the classification model is better in the reliable training dataset [60], so the classification performance was presumed to indicate the coherence of the manually classified WBM dataset.

The dataset obtained after reconciliation in Section 4.1 was used in this section. The WBM dataset was preprocessed using simple preprocessing. Several studies denoised a WBM dataset to remove defects

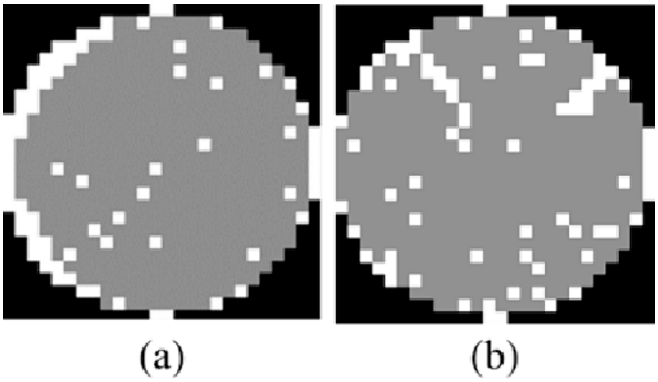



**Fig. 4.** Examples of WBM images undefined in the WM811K taxonomy.

**Fig. 5.** Newly-added defect pattern (Type 1).

**Fig. 6.** Newly-added defect pattern (Type 2).

**Table 6**
Manual classification results for the WBM dataset.

| Proposed taxonomy | | | WM811K taxonomy | | |
|---|---|---|---|---|---|
| Defect pattern | Original | Remain | Defect pattern | Original | Remain |
| *Cluster–Small–Edge* | 1417 | 1417 | *None* | 1272 | 1264 |
| *Ring–All–Edge* | 579 | 578 | *Edge-Ring* | 701 | 690 |
| *Random–Infrequent–Scattered* | 569 | 559 | *Edge-Loc* | 615 | 578 |
| *Scratch–Short–All* | 266 | 266 | *Scratch* | 365 | 363 |
| *Cluster–Small–Others* | 118 | 118 | *Location* | 175 | 163 |
| *Scratch–Long–All* | 103 | 102 | *Center* | 58 | 57 |
| *Cluster–All–Center* | 77 | 75 | *Random* | 23 | – |
| *Cluster–Big–Edge* | 45 | – | *Donut* | 0 | – |
| *Random –Frequent–Scattered* | 22 | – | *Near-Full* | 0 | – |
| *Cluster–Big–Others* | 13 | – | | | |
| Total | 3209 | 3115 | | 3209 | 3115 |

lacking spatial correlations [12,33,56,61]. However, denoising was not used in this study, as the two taxonomies define defect patterns for WBMs lacking spatial correlation among defects. For each wafer matrix, each chip was assigned 0 for the non-wafer area, 0.5 for functional chips, and 1 for defective chips, and zero padding was not used.

A convolutional neural network (CNN) [62] was utilized for the classification model in the experiments. The CNN model has an advantage over other alternative feature extraction-based methods in that it directly extracts the spatial features of defect patterns from WBM images [13]. Thus, the classification performance of different taxonomies can be clearly compared using the CNN model. However, the classification performance is affected by feature extraction methods as well as taxonomy using the feature extraction-based models [50,54]. For instance, rotational invariant moment-based features are effective when defect patterns are defined regardless of the rotational differences [45,50]. The CNN model is widely used in the defect pattern classification using the WM811K taxonomy, and many comparative studies revealed that the classification performance of the CNN model is superior to that of other models [2,14,55].

We referred to the traditional CNN architecture used in the WBM classification study because the input size of WBM was smaller than the commonly used image dataset [2,55]. However, a simpler architecture was used because the WBM size was smaller in our dataset. The CNN architecture used in this study is depicted in Fig. 7. Two convolutional layers and two fully-connected layers ("Conv" and "FC" in Fig. 7, respectively) constituted the CNN architecture. Each convolutional layer used a $3 \times 3$ convolution operation with a stride size of $1 \times 1$, and four convolution filters were present. In the second convolutional layer, $2 \times 2$ max-pooling with a stride size of 2 was conducted. The number of nodes in the first and second fully connected layers was 512. The number of output nodes was set to be the number of defect patterns in a taxonomy. Finally, each WBM was classified into the defect pattern with the highest node value. The rectified linear unit was used as an activation function from the first convolutional layer to the second fully connected layer, and the SoftMax function was used for output nodes. For training the model, categorical cross entropy was used for loss function, batch normalization was used with the batch size of 100, and an Adam [63] optimizer was used with a learning rate of 0.001.

Defect patterns with less than 50 WBMs were removed from the dataset because the classification model hardly extracts spatial features from defect patterns with such small sample sizes. Three defect patterns from the proposed taxonomy and one defect pattern from the WM811K taxonomy were consequently removed from the dataset. The WBMs that were removed from one taxonomy were also removed from the other taxonomy. For instance, the number of WBMs in *Location* was also decreased by the removal of *Cluster–Big–Others*. As a result, 3115 WBMs from 3209 remained and were used for the experiment. The column of "Remain" in Table 6 presents the numbers of WBMs after the removal of the defect patterns with small sample sizes.

The remaining dataset was randomly divided into training and validation datasets in a ratio of 8:2. The CNN model weight with the best accuracy on the validation set was then stored until 50 epochs. The maximum epoch was set according to the overfitting problem in preliminary experiments. The validation loss increased rapidly after around 10 epochs in both taxonomies (Fig. 8).

For the stored CNN model weights, the accuracy, precision, recall, and F-score of all defect patterns were examined. Macro-averaging of each metric was conducted in evaluating the performance of a classification model in consideration of the data imbalance. Macro-averaging calculates the average of a metric of defect patterns in a taxonomy, and all defect patterns have identical weights, regardless of sample size [64]. Table 7 provides the formulas of the metrics, where $l$ stands for the number of defect patterns in a taxonomy; $i$ denotes the index of the defect pattern; $tp_i$, $tn_i$, $fp_i$, and $fn_i$ are the true positive, true negative, false positive, and false negative counts for the defect pattern $i$, respectively. The $M$ index represents the macro-averaging.

The experiment was iterated 100 times, and the experimental results showed that the manual classification was more coherent in the proposed taxonomy than in the WM811K counterpart. Fig. 9 presents the box plots of the performance metrics. The box extended from Q1 to Q3 quartile values, and the whisker position was 1.5 interquartile range from the box edge. The median values of the average accuracy, macro-recall, and macro-F-score were higher in the proposed taxonomy. The median value of macro-precision was higher in the WM811K taxonomy, but more outliers exist than in the proposed taxonomy, which indicates that the classification performance was inconsistent during the iterated experiments. Table 8 presents the result of Welch's $t$-test for checking if the mean values of the metrics in 100 iterations from the two taxonomies were similar. The average accuracy, macro-recall, and macro-F-score were significantly greater in the proposed taxonomy.

The low recall of the WM811K taxonomy is attributed to low recalls of *Location* and *Center* patterns. Table 9 shows the average value of recall by each pattern for 100 iterated experiments, and *Location* and *Center* shows the extremely smaller recall than those of others. Table 10 presents the classified result for 100 iterated experiments of WBMs with *Location* and *Center* patterns. WBMs with *Location* and *Center* patterns were frequently classified as *Edge-Loc*, which results in low recall of *Location* and *Center* patterns. This result indicates that the differentiation criteria between *Center*, *Edge-Loc*, and *Location* patterns were untrained in the classification model based on the WM811K taxonomy. The result is consistent with the qualitative assessment (Section 4.1) result, in which the inconsistent manual classification arose due to the unclear differentiation criteria of the locational information in the WM811K taxonomy.

### 4.4. Discussion on expected benefits

This section discusses the expected benefits of the proposed

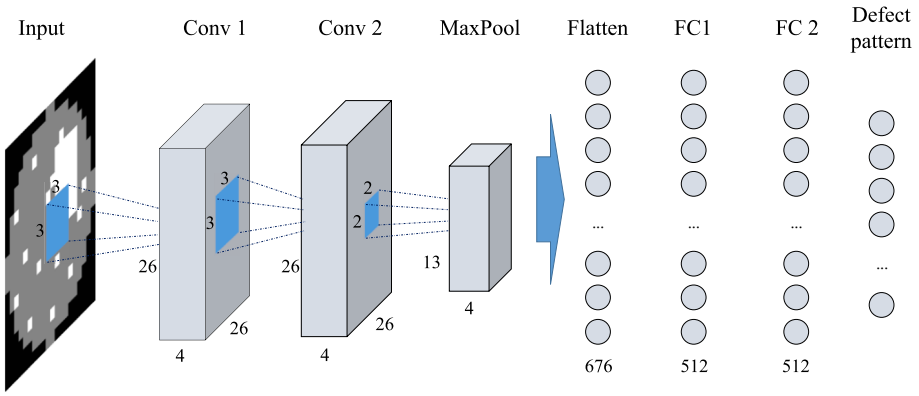
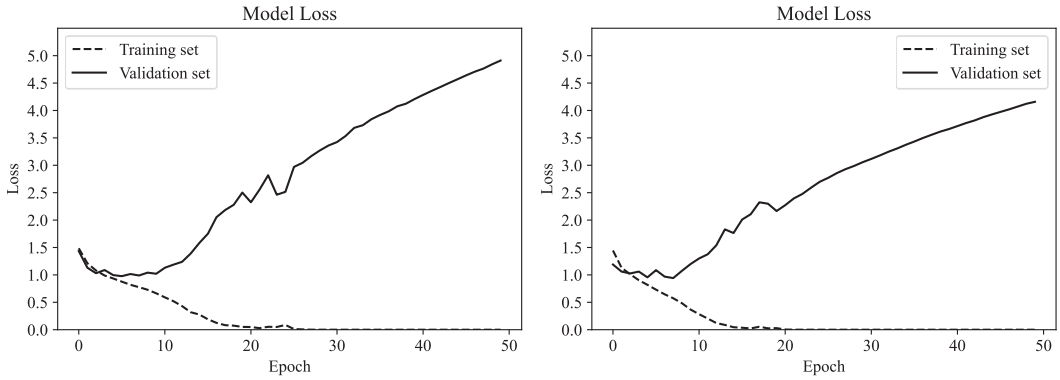



**Fig. 7.** CNN architecture.




(a) Proposed taxonomy (b) WM811K taxonomy

**Fig. 8.** Model loss function in the preliminary experiment.

**Table 7**
Performance metrics for defect pattern classification models.

| Measure | Formula |
|---|---|
| Average accuracy | $\frac{\sum_{i=1}^{l}\frac{tp_i+tn_i}{tp_i+tn_i+fp_i+fn_i}}{l}$ |
| $Precision_M$ | $\frac{\sum_{i=1}^{l}\frac{tp_i}{tp_i+fp_i}}{l}$ |
| $Recall_M$ | $\frac{\sum_{i=1}^{l}\frac{tp_i}{tp_i+fn_i}}{l}$ |
| $Fscore_M$ | $\frac{Precision_M Recall_M}{Precision_M + Recall_M}$ |

taxonomy within practical applications based on the experimental findings presented in Sections 4.1 through Sections 4.3. Firstly, the accuracy of manual defect pattern classification can be improved by using the proposed taxonomy. As discussed in Section 4.1, employing the proposed taxonomy led to an improvement in inter-rater agreement during manual classification. Given that higher inter-rater agreement is indicative of a more trustful ground truth [65], manual classification using the proposed taxonomy is considered more accurate than that using WM811K. Moreover, the proposed taxonomy facilitated the identification of more WBMs exhibiting spatial correlation among defects compared to WM811K. Accurate classification holds significant importance for root cause diagnosis in wafer fabrication [7–10] since the diagnostic performance relies on accurately classified WBM datasets.

Secondly, the proposed taxonomy can enhance the accuracy of automatic defect pattern classification. Without the need for model enhancements or additional WBM data collection, the automatic classification model demonstrated better accuracy when using the proposed taxonomy compared to WM811K (Section 4.3). This improvement can be attributed to the more cohesive classification of the WBM dataset used for model training with the proposed taxonomy. The increased accuracy in automatic defect pattern classification is advantageous as it aids practitioners in diagnosing the root causes of non-random defects.

Lastly, the proposed taxonomy is extendable, allowing for the

**Table 8**
Summary of the experimental results.

| Metric | Mean | | Welch's test $p$-value |
|---|---|---|---|
| | Proposed | WM811K | |
| Average accuracy | 0.894 | 0.865 | <0.001 |
| Macro-precision | 0.551 | 0.557 | 0.556 |
| Macro-recall | 0.487 | 0.336 | <0.001 |
| Macro-F-score | 0.516 | 0.417 | <0.001 |

**Table 9**
Average recall of experiments.

| Proposed | Average Recall | WM811K | Average Recall |
|---|---|---|---|
| *Cluster–Small–Edge* | 0.785 | *None* | 0.901 |
| *Cluster–All–Center* | 0.760 | *Edge-Ring* | 0.726 |
| *Ring–All–Edge* | 0.681 | *Edge-Loc* | 0.343 |
| *Random –Infrequent–Scattered* | 0.562 | *Scratch* | 0.036 |
| *Scratch–Long–All* | 0.344 | *Center* | 0.005 |
| *Cluster–Small–Others* | 0.173 | *Location* | 0.004 |
| *Scratch–Short–All* | 0.105 | | |

**Table 10**
Cumulated classification result of *Center* and *Location*.

| Prediction | Actual | |
|---|---|---|
| | *Center* | *Location* |
| *Center* | 6 | 0 |
| *Location* | 5 | 13 |

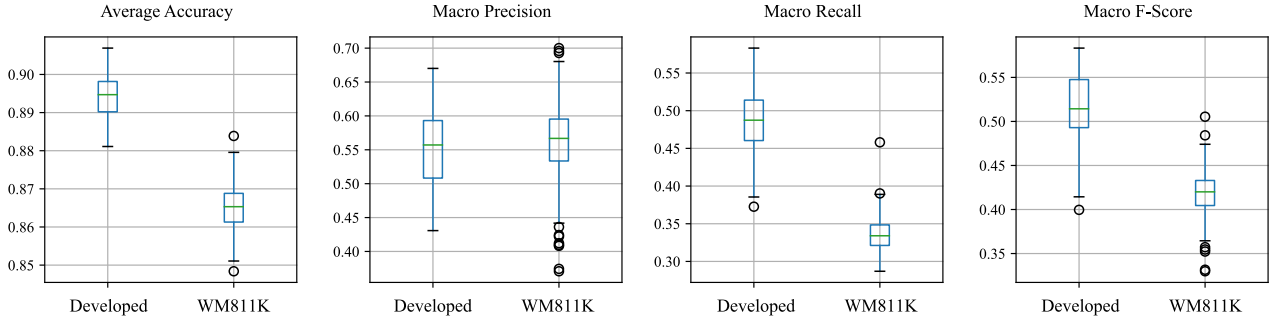



**Fig. 9.** Box plot of the experimental results.

addition of new defect patterns to accommodate emerging patterns (Section 4.2). This extendibility is attributed to the flexibility provided by the spatial dimension-based structure of the proposed taxonomy. This characteristic enables practitioners to continually utilize the proposed taxonomy with some modifications in response to newly emerging patterns, such as those arising from technological advancements.

## 5. Conclusion

The concept of spatial dimension-based defect pattern taxonomy and its development procedure are proposed. The concept was applied to a NAND flash memory semiconductor manufacturer for two years. Results indicate that the proposed taxonomy possesses the three desirable properties of a general taxonomy (i.e., robustness, comprehensiveness, and extendibility). Also, the domain expertise is integrated through the engineers' involvement in the proposed taxonomy development procedure.

The main contribution of this study is the development of a taxonomy possessing the three desirable properties through a spatial dimension-based structure. This structure effectively addresses limitations found in existing taxonomies, resulting in improved robustness, comprehensiveness, and extendibility.

The managerial implications of this study are twofold. Firstly, the classification accuracy of the defect pattern classification system would be improved. The proposed taxonomy possesses robustness and comprehensiveness, allowing for the coherent classification of the WBM dataset. Consequently, this coherently classified dataset contributes to improved accuracy in the defect pattern classification system. Secondly, the proposed taxonomy is considered practical for wafer fabrication diagnostics because it has integrated domain expertise in its development procedure. It encompasses a concise set of defect patterns that are meaningful for diagnosing wafer fabrication failures, thus making the classified results more applicable for diagnostics. These managerial implications remain valid for newly-emerged defect patterns due to the extendibility of the proposed taxonomy.

Three issues for future research are suggested. The first issue resides in the lack of case studies that validate the general applicability of the concept of spatial dimension-based taxonomy. The desirable properties were demonstrated for one wafer fabrication process in this study. Case studies on other manufacturing circumstances should be conducted in future research to validate the general applicability of the concept and development procedure. The second issue resides in the extension of dimensions. The mixed-type defect patterns occur when several defect patterns appear simultaneously in a wafer. Mixed-type defect patterns frequently occur due to new technology adoption [5,15]. To classify mixed-type defect patterns, dimensions should be extended. The number [48], overlap [66], and homogeneity [67] of defect patterns may be added to the dimension. The third issue resides in improving the taxonomy development procedure to be more intelligent. The active learning framework can be adopted to make the engineers' involvement efficient. For instance, inspection of clustered defects [68] can reduce the number of WBMs to be inspected in Stage 1 of the taxonomy development procedure. Manual identification of the newly-emerged defect patterns can be replaced by the classification uncertainty analysis [53] for prompt extension of the existing taxonomy.

## CRediT authorship contribution statement

**Seung-Hyun Choi:** Writing – original draft, Validation, Methodology, Conceptualization. **Dong-Hee Lee:** Writing – review & editing, Methodology. **Eun-Su Kim:** Methodology. **Young-Mok Bae:** Funding acquisition, Data curation. **Young-Chan Oh:** Funding acquisition, Data curation. **Kwang-Jae Kim:** Writing – review & editing, Supervision.

## 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

The data that has been used is confidential.

## Acknowledgment

This work was supported by the National Research Foundation of Korea (NRF) grant funded by the Korea government (MSIT) [NRF-2019R1A2C1007834] and by SK hynix Inc.

## References

[1] S. Park, J. Jang, C.O. Kim, Discriminative feature learning and cluster-based defect label reconstruction for reducing uncertainty in wafer bin map labels, J. Intell. Manuf 32 (2021) 251–263, https://doi.org/10.1007/s10845-020-01571-4.
[2] J.-H. Kim, H.-S. Kim, J.-S. Park, K.-H. Mo, P.-S. Kang, Bin2Vec: a better wafer bin map coloring scheme for comprehensible visualization and effective bad wafer classification, Appl. Sci. (Switzerland) 9 (2019) 597, https://doi.org/10.3390/app9030597.
[3] T.-S. Li, C.-L. Huang, Defect spatial pattern recognition using a hybrid SOM–SVM approach in semiconductor manufacturing, Expert. Syst. Appl 36 (2009) 374–385, https://doi.org/10.1016/j.eswa.2007.09.023.
[4] S.P. Cunningham, S. MacKinnon, Statistical methods for visual defect metrology, IEEE. Trans. Semicond. Manuf. 11 (1998) 48–53, https://doi.org/10.1109/66.661284.
[5] K. Kyeong, H. Kim, Classification of mixed-type defect patterns in wafer bin maps using convolutional neural networks, IEEE. Trans. Semicond. Manuf. 31 (2018) 395–402, https://doi.org/10.1109/TSM.2018.2841416.
[6] W. Taam, M. Hamada, Detecting spatial effects from factorial experiments: an application from integrated-circuit manufacturing, Technometrics 35 (1993) 149–160, https://doi.org/10.1080/00401706.1993.10485037.
[7] K. Radigan, B. Sheumaker, N. Heller, Using full wafer defect maps as process signatures to monitor and control yield, in: Proceedings IEEE/SEMI international semiconductor Manufacturing science symposium, IEEE, 1991, pp. 129–135, https://doi.org/10.1109/ISMSS.1991.146281.
[8] M.H. Hansen, V.N. Nair, D.J. Friedmand, D. Friedman, Process improvement through the analysis of spatially clustered defects on wafer maps, 1999.
[9] K. Nakata, R. Orihara, Y. Mizuoka, K. Takagi, A comprehensive big-data-based monitoring system for yield enhancement in semiconductor manufacturing, IEEE. Trans. Semicond. Manuf. 30 (2017) 339–344, https://doi.org/10.1109/TSM.2017.2753251.
[10] K. Taha, CDID: a system for identifying the root cause of a defect in semiconductor wafer fabrication, IEEE. Trans. Semicond. Manuf. 31 (2018) 221–231, https://doi.org/10.1109/TSM.2018.2808703.
[11] A. Drozda-Freeman, M. McIntyre, M. Retersdorf, C. Wooten, X. Song, A. Hesse, The Application and Use of an Automated Spatial Pattern Recognition (SPR) System in the Identification and Solving of Yield Issues in Semiconductor Manufacturing, in: Intergovernmental Panel on Climate Change (Ed.), 2007 IEEE/SEMI Advanced Semiconductor Manufacturing Conference, IEEE, Cambridge, 2007: pp. 302–305, https://doi.org/10.1109/ASMC.2007.375121.
[12] Y. Kong, D. Ni, Semi-supervised classification of wafer map based on ladder network, in: IEEE International Conference on Solid-State and Integrated Circuit Technology (ICSICT), 2018, pp. 1–4, https://doi.org/10.1109/ICSICT.2018.8564982.
[13] T. Nakazawa, D.V. Kulkarni, Anomaly detection and segmentation for wafer defect patterns using deep convolutional encoder–decoder neural network architectures in semiconductor Manufacturing, IEEE. Trans. Semicond. Manuf. 32 (2019) 250–256, https://doi.org/10.1109/TSM.2019.2897690.
[14] C.Y. Hsu, J.C. Chien, Ensemble convolutional neural networks with weighted majority for wafer bin map pattern classification, J. Intell. Manuf 33 (2022) 831–844, https://doi.org/10.1007/s10845-020-01687-7.
[15] T.S. Kim, J.W. Lee, W.K. Lee, S.Y. Sohn, Novel method for detection of mixed-type defect patterns in wafer maps based on a single shot detector algorithm, J. Intell. Manuf (2021), https://doi.org/10.1007/s10845-021-01755-6.
[16] C.H. Jin, H.J. Kim, Y. Piao, M. Li, M. Piao, Wafer map defect pattern classification based on convolutional neural network features and error-correcting output codes, J. Intell. Manuf 31 (2020) 1861–1875, https://doi.org/10.1007/s10845-020-01540-x.
[17] D.F. Nettleton, A. Orriols-Puig, A. Fornells, A study of the effect of different types of noise on the precision of supervised learning techniques, Artif. Intell. Rev 33 (2010) 275–306, https://doi.org/10.1007/s10462-010-9156-z.
[18] N. Omri, Z. Al Masry, N. Mairot, S. Giampiccolo, N. Zerhouni, Industrial data management strategy towards an SME-oriented PHM, J. Manuf. Syst 56 (2020) 23–36, https://doi.org/10.1016/j.jmsy.2020.04.002.

[19] P. Wang, M. Luo, A digital twin-based big data virtual and real fusion learning reference framework supported by industrial internet towards smart manufacturing, J. Manuf. Syst 58 (2021) 16–32, https://doi.org/10.1016/j.jmsy.2020.11.012.

[20] M.J. Wu, J.S. Jang, J.L. Chen, Wafer map failure pattern recognition and similarity ranking for large-scale data sets, IEEE. Trans. Semicond. Manuf. 28 (2015) 1–12, https://doi.org/10.1109/TSM.2014.2364237.

[21] Y. Wang, D. Ni, A deep learning analysis framework for complex wafer bin map classification, IEEE. Trans. Semicond. Manuf. 36 (2023) 367–377, https://doi.org/10.1109/TSM.2023.3269230.

[22] K.S.M. Li, N.C.Y. Tsai, K.C.C. Cheng, X.H. Jiang, P.Y.Y. Liao, S.J. Wang, A.Y. A. Huang, L. Chou, C.S. Lee, TestDNA: novel wafer defect signature for diagnosis and pattern recognition, IEEE. Trans. Semicond. Manuf. 33 (2020) 383–390, https://doi.org/10.1109/TSM.2020.2992927.

[23] Y. Wang, D. Ni, Multi-bin wafer maps defect patterns classification, Proceedings - 2019 IEEE International Conference on Smart Manufacturing, Industrial and Logistics Engineering, SMILE 2019 (2019) 48–52, https://doi.org/10.1109/SMILE45626.2019.8965299.

[24] C.Y. Hsu, W.J. Chen, J.C. Chien, Similarity matching of wafer bin maps for manufacturing intelligence to empower industry 3.5 for semiconductor manufacturing, Comput. Ind. Eng 142 (2020) 106358, https://doi.org/10.1016/j.cie.2020.106358.

[25] C.-W.-W. Liu, C.-F.-F. Chien, An intelligent system for wafer bin map defect diagnosis: an empirical study for semiconductor manufacturing, Eng. Appl. Artif. Intell 26 (2013) 1479–1486, https://doi.org/10.1016/j.engappai.2012.11.009.

[26] R.C. Nickerson, U. Varshney, J. Muntermann, A method for taxonomy development and its application in information systems, Eur. J. Inf. Syst. 22 (2013) 336–359, https://doi.org/10.1057/ejis.2012.26.

[27] J. Jang, G.T. Lee, Decision fusion approach for detecting unknown wafer bin map patterns based on a deep multitask learning model, Expert. Syst. Appl 215 (2023) 119363, https://doi.org/10.1016/j.eswa.2022.119363.

[28] M.O. Gökalp, E. Gökalp, K. Kayabay, A. Koçyiğit, P.E. Eren, Data-driven manufacturing: an assessment model for data science maturity, J. Manuf. Syst 60 (2021) 527–546, https://doi.org/10.1016/j.jmsy.2021.07.011.

[29] C.-H. Wang, W. Kuo, H. Bensmail, Detection and classification of defect patterns on semiconductor wafers, IIE. Trans. 38 (2006) 1059–1068, https://doi.org/10.1080/07408170600733236.

[30] H.-W. Hsieh, F.-L. Chen, Recognition of defect spatial patterns in semiconductor fabrication, Int. J. Prod. Res 42 (2004) 4153–4172, https://doi.org/10.1080/00207540410001716507.

[31] J.Y. Hwang, W. Kuo, Model-based clustering for integrated circuit yield enhancement, Eur. J. Oper. Res 178 (2007) 143–153, https://doi.org/10.1016/j.ejor.2005.11.032.

[32] T. Yuan, W. Kuo, A model-based clustering approach to the recognition of the spatial defect patterns produced during semiconductor fabrication, IIE. Trans. 40 (2007) 93–101, https://doi.org/10.1080/07408170701592556.

[33] C.-H. Wang, Separation of composite defect patterns on wafer bin map using support vector clustering, Expert. Syst. Appl 36 (2009) 2554–2561, https://doi.org/10.1016/j.eswa.2008.01.057.

[34] Q. Zhou, L. Zeng, S. Zhou, Statistical detection of defect patterns using Hough transform, IEEE. Trans. Semicond. Manuf. 23 (2010) 370–380, https://doi.org/10.1109/TSM.2010.2048959.

[35] T. Yuan, W. Kuo, S.J. Bae, Detection of spatial defect patterns generated in semiconductor fabrication processes, IEEE. Trans. Semicond. Manuf. 24 (2011) 392–403, https://doi.org/10.1109/TSM.2011.2154870.

[36] G.-H. Choi, S.-H. Kim, C.-H. Ha, S.-J. Bae, Multi-step ART1 algorithm for recognition of defect patterns on semiconductor wafers, Int. J. Prod. Res 50 (2012) 3274–3287, https://doi.org/10.1080/00207543.2011.574502.

[37] K. Maksim, B. Kirill, Z. Eduard, G. Nikita, B. Aleksandr, L. Arina, S. Vladislav, M. Daniil, K. Nikolay, Classification of wafer maps defect based on deep learning methods with small amount of data, in: 2019 International Conference on Engineering and Telecommunication (EnT)IEEE, 2019, pp. 1–5, https://doi.org/10.1109/EnT47717.2019.9030550.

[38] F.-L. Chen, S.-F. Liu, A neural-network approach to recognize defect spatial pattern in semiconductor fabrication, IEEE. Trans. Semicond. Manuf. 13 (2000) 366–373, https://doi.org/10.1109/66.857947.

[39] S.J. Bae, J.Y. Hwang, W. Kuo, Yield prediction via spatial modeling of clustered defect counts across a wafer map, IIE. Trans. 39 (2007) 1073–1083, https://doi.org/10.1080/07408170701275335.

[40] M. Liukkonen, Y. Hiltunen, Recognition of systematic spatial patterns in silicon wafers based on SOM and K-means, IFAC-PapersOnLine 51 (2018) 439–444, https://doi.org/10.1016/j.ifacol.2018.03.075.

[41] F. Di Palma, G. De Nicolao, G. Miraglia, E. Pasquinetti, F. Piccinini, Unsupervised spatial pattern classification of electrical-wafer-sorting maps in semiconductor manufacturing, Pattern. Recognit. Lett 26 (2005) 1857–1865, https://doi.org/10.1016/j.patrec.2005.03.007.

[42] C.-H. Wang, S.-J. Wang, W.-D. Lee, Automatic identification of spatial defect patterns for semiconductor manufacturing, Int. J. Prod. Res 44 (2006) 5169–5185, https://doi.org/10.1080/02772240600610822.

[43] S.-C. Hsu, C.-F. Chien, Hybrid data mining approach for pattern extraction from wafer bin map to improve yield in semiconductor manufacturing, Int. J. Prod. Econ 107 (2007) 88–103, https://doi.org/10.1016/j.ijpe.2006.05.015.

[44] L.-C. Chao, L.-I. Tong, Wafer defect pattern recognition by multi-class support vector machines by using a novel defect cluster index, Expert. Syst. Appl 36 (2009) 10158–10167, https://doi.org/10.1016/j.eswa.2009.01.003.

[45] M.-P.-L. Ooi, H.-K. Sok, Y.-C. Kuang, S. Demidenko, C. Chan, Defect cluster recognition system for fabricated semiconductor wafers, Eng. Appl. Artif. Intell 26 (2013) 1029–1043, https://doi.org/10.1016/j.engappai.2012.03.016.

[46] C.-S. Liao, T.-J. Hsieh, Y.-S. Huang, C.-F. Chien, Similarity searching for defective wafer bin maps in semiconductor manufacturing, IEEE. Trans. Autom. Sci. Eng. 11 (2014) 953–960, https://doi.org/10.1109/TASE.2013.2277603.

[47] G. Tello, O.Y. Al-Jarrah, P.D. Yoo, Y. Al-Hammadi, S. Muhaidat, U. Lee, Deep-structured machine learning model for the recognition of mixed-defect patterns in semiconductor fabrication processes, IEEE. Trans. Semicond. Manuf. 31 (2018) 315–322, https://doi.org/10.1109/TSM.2018.2825482.

[48] C.-F. Chien, S.-C. Hsu, Y.-J. Chen, A system for online detection and classification of wafer bin map defect patterns for manufacturing intelligence, Int. J. Prod. Res 51 (2013) 2324–2338, https://doi.org/10.1080/00207543.2012.737943.

[49] M.-P.-L. Ooi, E.K.J. Joo, Y.C. Kuang, S. Demidenko, L. Kleeman, C.W.K. Chan, Getting more from the semiconductor test: data mining with defect-cluster extraction, IEEE. Trans. Instrum. Meas 60 (2011) 3300–3317, https://doi.org/10.1109/TIM.2011.2122430.

[50] J.W. Cheng, M.P.-L. Ooi, C. Chan, Y.C. Kuang, S. Demidenko, Evaluating the Performance of Different Classification Algorithms for Fabricated Semiconductor Wafers, in: 2010 Fifth IEEE International Symposium on Electronic Design, Test & ApplicationsIEEE, 2010, pp. 360–366, https://doi.org/10.1109/DELTA.2010.69.

[51] N.G. Shankar, Z.W. Zhong, Defect detection on semiconductor wafer surfaces, Microelectron. Eng 77 (2005) 337–346, https://doi.org/10.1016/j.mee.2004.12.003.

[52] A.I. Mirza, G. O'Donoghue, A.W. Drake, S.C. Graves, Spatial yield modeling for semiconductor wafers, in: Proceedings of SEMI Advanced Semiconductor Manufacturing Conference and WorkshopIEEE, 1995, pp. 276–281, https://doi.org/10.1109/ASMC.1995.484386.

[53] J. Shim, S. Kang, S. Cho, Active learning of convolutional neural network for cost-effective wafer map pattern classification, IEEE. Trans. Semicond. Manuf. 33 (2020) 258–266, https://doi.org/10.1109/TSM.2020.2974867.

[54] M. Saqlain, B. Jargalsaikhan, J.Y. Lee, A voting ensemble classifier for wafer map defect patterns identification in semiconductor manufacturing, IEEE. Trans. Semicond. Manuf. 32 (2019) 171–182, https://doi.org/10.1109/TSM.2019.2904306.

[55] R. di Bella, D. Carrera, B. Rossi, P. Fragneto, G. Boracchi, Wafer defect map classification using Sparse convolutional networks, in: International Conference on Image Analysis and Processing, Springer International Publishing, 2019, pp. 125–136, https://doi.org/10.1007/978-3-030-30645-8_12.

[56] M. Piao, C.H. Jin, J.Y. Lee, J.-Y. Byun, Decision tree ensemble-based wafer map failure pattern recognition based on radon transform-based features, IEEE. Trans. Semicond. Manuf. 31 (2018) 250–257, https://doi.org/10.1109/TSM.2018.2806931.

[57] J. Yu, X. Zheng, J. Liu, Stacked convolutional sparse denoising auto-encoder for identification of defect patterns in semiconductor wafer map, Comput. Ind 109 (2019) 121–133, https://doi.org/10.1016/j.compind.2019.04.015.

[58] U. Batool, M.I. Shapiai, H. Fauzi, J.X. Fong, Convolutional neural network for imbalanced data classification of silicon wafer defects, in: 2020 16th IEEE International Colloquium on Signal Processing & Its Applications (CSPA)IEEE, 2020, pp. 230–235, https://doi.org/10.1109/CSPA48992.2020.9068669.

[59] J. Cohen, A coefficient of agreement for nominal scales, Educ. Psychol. Meas 20 (1960) 37–46, https://doi.org/10.1177/001316446002000104.

[60] F. Cabitza, A. Campagner, D. Albano, A. Aliprandi, A. Bruno, V. Chianca, A. Corazza, F. Di Pietto, A. Gambino, S. Gitto, C. Messina, D. Orlandi, L. Pedone, M. Zappia, L.M. Sconfienza, The elephant in the machine: proposing a new metric of data reliability and its application to a medical case to assess classification reliability, Appl. Sci. 10 (2020) 4014, https://doi.org/10.3390/app10114014.

[61] T. Ishida, I. Nitta, D. Fukuda, Y. Kanazawa, Deep Learning-Based Wafer-Map Failure Pattern Recognition Framework, in: 20th International Symposium on Quality Electronic Design (ISQED)IEEE, 2019, pp. 291–297, https://doi.org/10.1109/ISQED.2019.8697407.

[62] A. Krizhevsky, I. Sutskever, G.E. Hinton, ImageNet classification with deep convolutional neural networks, Commun. ACM 60 (2017) 84–90.

[63] D.P. Kingma, J. Ba, Adam: A Method for Stochastic Optimization, (2014). https://arxiv.org/abs/1412.6980 (accessed May 10, 2021).

[64] M. Sokolova, G. Lapalme, A systematic analysis of performance measures for classification tasks, Inf. Process. Manag 45 (2009) 427–437, https://doi.org/10.1016/j.ipm.2009.03.002.

[65] A. Campagner, D. Ciucci, C.-M. Svensson, M.T. Figge, F. Cabitza, Ground truthing from multi-rater labeling with three-way decision and possibility theory, Inf. Sci. (n. y) 545 (2021) 771–790, https://doi.org/10.1016/j.ins.2020.09.049.

[66] Y. Kong, D. Ni, Recognition and Location of Mixed-type Patterns in Wafer Bin Maps, in: 2019 IEEE International Conference on Smart Manufacturing, Industrial & Logistics Engineering (SMILE)IEEE, 2019, pp. 4–8, https://doi.org/10.1109/SMILE45626.2019.8965309.

[67] C.-H. Wang, Recognition of semiconductor defect patterns using spatial filtering and spectral clustering, Expert. Syst. Appl 34 (2008) 1914–1923, https://doi.org/10.1016/j.eswa.2007.02.014.

[68] J. Shim, S. Kang, S. Cho, Active cluster annotation for wafer map pattern classification in semiconductor manufacturing, Expert. Syst. Appl 183 (2021) 115429, https://doi.org/10.1016/j.eswa.2021.115429.