[13] N. Brahimi, S. Dauzere-Peres, N. Najid, and A. Nordli, "Single item lot sizing problems," *Eur. J. Oper. Res.*, vol. 168, pp. 1–16, 2006.
[14] G. Bitran and H. Yanasse, "Computational complexity of the capacitated lot size problem," *Manag. Sci.*, vol. 28, pp. 1174–1186, 1982.
[15] A. Akbalik and C. Rapine, "Polynomial time algorithms for the constant capacitated single-item lot sizing problem with stepwise production cost," *Oper. Res. Lett.*, vol. 40, pp. 390–397, 2012.
[16] M. Florian, J. Lenstra, and A. Rinnooy Kan, "Deterministic production planning: Algorithms and complexity," *Manag. Sci.*, vol. 26, pp. 669–679, 1980.
[17] Y. Pochet and L. Wolsey, "Lot-sizing with constant batches: Formulations and valid inequalities," *Math. Oper. Res.*, vol. 18, pp. 767–785, 1993.
[18] C. van Hoesel and A. Wagelmans, "An $O(T^3)$ algorithm for the economic lotsizing problem with constant capacities," *Manag. Sci.*, vol. 42, pp. 142–150, 1996.

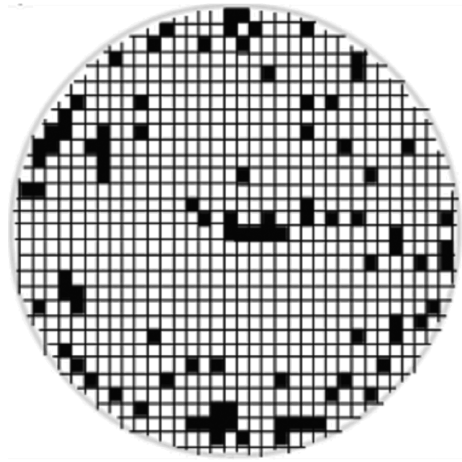


Fig. 1. An example of defective wafer bin maps: □ functional bins; ■ defective bins.

Similarity Searching for Defective Wafer Bin Maps in Semiconductor Manufacturing

Chung-Shou Liao, *Member, IEEE*, Tsung-Jung Hsieh, Yu-Syuan Huang, and Chen-Fu Chien, *Member, IEEE*

***Abstract*—Because high-dimensional wafer bin maps (WBMs) cause various features, it is difficult to search the similarity among WBMs via conventional pattern recognition methods. This study develops a novel morphology-based support vector machine for defective wafer detection. The experimental results demonstrate its usefulness in yield improvements on precision and computation cost.**

***Note to Practitioners*—Semiconductor manufacturing in complicated nanotechnology is facing tough challenge for quick response to yield excursion for shortening time to market and reducing the cost to maintain competitive advantages. Due to the increasing complexity of nanotechnology for wafer fabrication, increasingly high inspection costs and yield loss associated with defective wafers have become a critical concern of semiconductor manufacturers. Focused on real settings of practical industrial experiments, this study provides a novel approach to searching similar WBMs from huge wafer spatial data to quickly identify potential causes for yield enhancement. This approach was validated in real setting in Taiwan and the results showed its practical viability.**

***Index Terms*—Data mining, morphology, semiconductor manufacturing, similarity search, support vector machines, wafer bin maps.**

I. Introduction

With increasingly sophisticated manufacturing processes in the semiconductor industry, the cost to migrate the required techniques adds substantially to semiconductor production costs. Indeed, for the semiconductor industry as wafer production and size continues to grow, the volume of in-line and off-line data required to diagnose yield conditions is growing exponentially. Furthermore, high-volume wafer fabrication facilities typically produce thousands of wafers per week, and many of these wafers are inspected and found to be defective [1], [2], [3]. In this scenario, yield improvement is critically important, as is maintaining competitive processes and low die costs for semiconductor wafers in a fabrication facility.

During the final process of wafer fabrication, Circuit Probe (CP) test will determine whether the corresponding die is good for packaging into chip. Spatial patterns of testing results are WBMs that provide crucial information to identify process failures, as illustrated in Fig. 1. These patterns are formed by marking the defective wafers, so that manufacturing engineers may use the patterns of WBMs as clues to investigate the causes of failures resulting in yield losses. As there are many WBMs to be evaluated, the judgments in semiconductor manufacturing thus far still rely on human labor. As a result, the judgments may be inconsistent owing to human factors (e.g., fatigue) because of the substantial workload.

In particular, not only should the defective dies be detected before package, but assignable causes should also be attributed to reduce yield and profit loss due to scrapped wafers. During the CP test, wafers are inspected by retrieving information about defect patterns [4]. More precisely, WBM patterns can provide information to help better monitor the processes and products. The WBMs, in many cases, contain characteristic patterns, or signatures, which provide insight into the fitness of the manufacturing processes. A bin can be conceptualized as a bucket and viewed by mapping the results of these electrical tests onto a 2-D space.

These defective patterns are usually associated with specific manufacturing problems, and can provide process and product engineers with important clues regarding the identification of causes and their solutions in order to improve yields [5]. Stapper [6] indicated that defects are typically clustered, rather than dispersed randomly over a wafer, and that these clusters become more evident as the wafer size increases.

In the literature, the three proposed approaches to solving the pattern recognition problem are stated as: the statistical approach, the heuristic approach, and the simulation approach [7]. The statistical approach classifies patterns based on an extracted feature set, and an underlying statistical model for generating these patterns. The heuristic approach

Manuscript received May 02, 2013; accepted July 15, 2013. Date of publication August 29, 2013; date of current version June 30, 2014. This paper was recommended for publication by Associate Editor S. Zhou and Editor H. Ding upon evaluation of the reviewers' comments. This work was supported by the Advanced Manufacturing and Service Management Research Center, National Tsing Hua University under Toward World-Class Universities Projects 101N2073E1, 101N2074E1 and the National Science Council of Taiwan under Grant NSC100-2221-E-007-108-MY3, Grant NSC100-2628-E-007-017-MY3, and Grant NSC102-2221-E-007-075-MY3. *(Corresponding author: C.-S. Liao.)*

The authors are with the Department of Industrial Engineering and Engineering Management, National Tsing Hua University, Hsinchu 300, Taiwan (e-mail: csliao@ie.nthu.edu.tw; tsungjung.hsieh@gmail.com; s9934519@m99.nthu.edu.tw; cfchien@mx.nthu.edu.tw).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TASE.2013.2277603

1545-5955 © 2013 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission.
See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.
Authorized licensed use limited to: Hochschule Heilbronn. Downloaded on September 16,2025 at 05:20:29 UTC from IEEE Xplore. Restrictions apply.

utilizes soft computing schemes, such as *genetic algorithms* and *fuzzy sets*, to perform pattern recognition processes. However, genetic systems typically require expensive evaluation processes to achieve an optimal solution [8]. Furthermore, a major limitation of the fuzzy logic controller is that it requires that suitable linguistic control rules are generated, as well as the knowledge and experience of human experts [9]. The simulation approach emulates the computational paradigm of a biological system, subsequently leading to a class of artificial neural systems, termed *neural networks*[7], [10], [11]. However, the main drawback of neural networks is their inability to determine the number of layers and number of neurons per layer [12]. To improve upon this insufficiency of neural networks, *support vector machines* (SVM) were proposed in 1995 [13] and have been widely used for pattern recognition in recent years. Several studies noted that SVM classification was more accurate than previous classification algorithms [14], [15].

Our Contribution: In this work, we propose a morphology-based SVM (MSVM) similarity searching method to generate wafer samples with certain degrees of similarity, as compared to the objective *target wafer maps*. From a practical perspective, the failure patterns of defective wafers are increasingly complex, and possess the characteristics of bin rotation, so traditional pattern recognition methods cannot be adapted to address this problem. Nevertheless, we collaborated with the industry, and provide the MSVM as an alternative for achieving high-precision detection, thereby lowering costs and saving time.

The rest of this paper is organized as follows. Section II describes the problem. Section III presents the central concept of our morphology-based SVM approach. Section IV shows experimental results to demonstrate the usefulness of our approach. Finally, we conclude with some discussions and future work in Sections V and VI.

II. The Investigated Problem

A. Problem Description

The similarity search of defective wafer patterns in this work mainly addresses two problems as follows.

First, the traditional pattern recognition or classification methods can not easily recognize complex patterns due to the increasing variety of defective wafers. In particular, a slight change of bins can form a totally different defective pattern from its original. Furthermore, as the wafer size increases, the WBMs will extend from both x- and y-axes directions, causing the defective portions to show a square growth as well. Thus, the problem of complexity grows quadratically, since the bin maps with higher dimensions may have a great deal of variation, and further generate several more complicated wafer maps, such as in the case of "map group movement," which may have a high degree of similarity to the wafer maps. Consequently, owing to the high-dimensional bin map, it is extremely difficult to capture the variations of each dimension.

Second, the problem is that the wafer samples in stock from fabricators are costly and limited; besides, the types of defective wafer are too numerous to collect. This problem results in inaccurate decisions by such machine learning methods as ART [16], SOM [17], SVM [17], [18], etc., when using insufficient training samples. Because more functional wafers possess more precise designs, it is vital to consider possible changes in small areas of the wafers. In this case, under the high-dimensional WBMs, a small change may form a completely different wafer type, making it difficult to discover differences via traditional pattern recognition techniques and classification methods. Thus, in order to reduce the inspection cost for a large volume of productivity, it is necessary to generate representative wafer samples with the original characteristics.

Based on the above reasons, we construct morphological training samples for the SVM learning procedure in sufficient quantities efficiently, and avoid excessive expense in sample sources. Also, for the high-dimensional bin maps, a preprocessing *scale normalization* developed in collaboration with the industry is incorporated into the procedure of wafer generation, so that the size of wafer bin maps in the training phase is consistent.

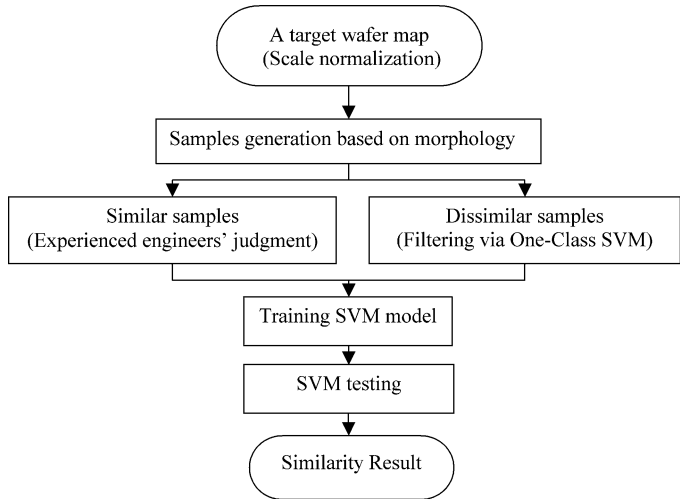


Fig. 2. Flowchart of the proposed MSVM.

B. Problem Resolving Flowchart

The proposed morphology-based support vector machine (MSVM) for similarity searching will generate wafer samples with various certain degrees of similarity, as compared to the objective *target wafer maps*. The implementation flowchart of the MSVM is illustrated in Fig. 2.

First, we obtain a target wafer map as the target that has gone through normalization preprocessing; then, to directly extract more feature information from the target wafer map, the MSVM tool combines with morphology to generate various wafer patterns. These simulated wafer features, based on morphology, provide for several types, including dilation, erosion, opening, closing, position shift, density change and rotation with variations, with respect to the target wafer map, which will be introduced in more detail later.

Next, the generated training samples contain both similar and dissimilar wafer maps; the former is a reference based on domain experts, and the latter is derived through a filtering process, which uses One-Class SVM to exclude possible similar samples, so as to ensure a high quality training phase in SVM. Recent research [18] has presented a recognition system using SVM with a defect cluster index to efficiently and accurately recognize wafer defect patterns. The cluster index was designed to transform the information regarding the proportion of wafer defects into a numerical expression, and used them as the SVM inputs. However, when more complex and high-dimensional wafer patterns are considered, more aspects for pattern extraction for SVM must be included. Based on the morphological training data from the first phase, SVM will categorize the testing samples into groups with respect to the given target wafer. In so doing, the proposed similarity search method employs the SVM trained by morphology generating samples, which can search for defective wafers as well as the corresponding causes of their defects, depending on the process engineers' demands.

III. Methodology

Each wafer may have fabricated from several hundred to several thousand dies on its surface. A typical WBM usually contains a number of dies failed in different functional tests. For visualization and analysis

Authorized licensed use limited to: Hochschule Heilbronn. Downloaded on September 16,2025 at 05:20:29 UTC from IEEE Xplore. Restrictions apply.

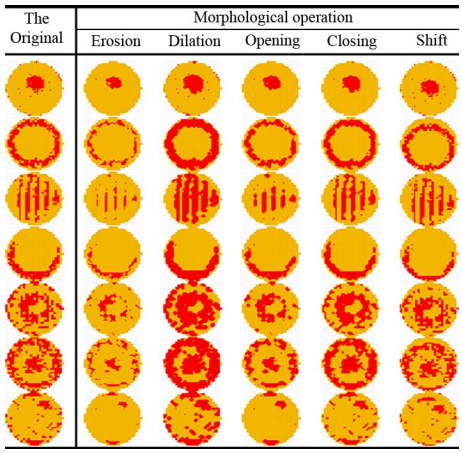


Fig. 3. An illustration of morphology-based samples.

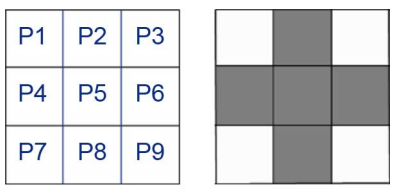


Fig. 4. The structure element in morphology (a) with respect to a wafer bin (b) influential elements.

purposes, WBM is usually transformed into a binary map and binary code, or two different colors are used for representation. This work uses red squares, or 1, to denote defective chips, and yellow squares, or 0, to denote functional chips. Next, we highlight the methods used in the proposed MSVM, including morphological sample generation, One-Class SVM sample filter, and a brief description of the training tool, SVM.

A. Morphology-Based Sample Generation

Morphology was first proposed by Matheron and Serra [19] in 1968. Originally, morphology was employed in mathematics; current morphology is extended to several aspects in image processing. The concept and operations of morphology we used are illustrated in Figs. 3, 5, and 6. In addition to the five common features, we propose two new morphological methods for sample generation in order to adapt to changeable patterns of complex WBMs.

Erosion: Each chip or bin can be viewed as a *structural element* (SE), as shown in Fig. 4(a), which is formed by a 3×3 matrix. In the erosion operation, we focus on a specific region and locate each considered bin at position P5, which denotes whether conserving P5 (the bin) is decided by some elements (bins) directly connecting to P5, i.e., Pi, $i = 2$, 4, 6, and 8. These elements are called *influential elements* (IE), as shown in Fig. 4(b).

The effect of erosion is to make the image structure thin and sparse (see Fig. 3). Equation (1) shows the erosion operation

$$f \Theta SE = P5 \cap (P1 \cap P2 \cap P3 \cap P4 \cap P5 \cap P6 \cap P7 \cap P8 \cap P9) \quad (1)$$

where f denotes the original image, and SE denotes structural elements. Note that if a defective element is surrounded by n defective IE ($1 \leq n \leq 4$), that element becomes functional (set as 0). Each bin is subjected to the erosion operation independently, i.e., all elements must be viewed under original conditions when erosion is applied. The parameter n can be set by the user in order to apply the effects of slight or heavy erosion.

Dilation: The dilation operation is the opposite of the erosion operation, and is described as (2)

$$f \oplus SE = P5 \cup (P1 \cup P2 \cup P3 \cup P4 \cup P6 \cup P7 \cup P8 \cup P9). \quad (2)$$

Similarly, f is the original image, and SE refers to the structural elements. An illustration is shown in Fig. 3. According to (2), as long as a defective element is surrounded by at least one defective IE, the whole IE becomes defective (set as 1). Again, each chip performs the operation independently; that is, the defective IEs must be viewed under original conditions when the dilation operation is initiated.

Opening: The opening operation is designed to make the original pattern thinner, connecting or noncontinuous; also, the image could have a smoother contour (see Fig. 3). Opening is realized by first performing erosion, and then dilation. The opening operation is shown in (3)

$$f \circ SE = (f \Theta SE) \oplus SE. \quad (3)$$

Closing: The closing operation is opposite to opening. The effect of closing is to fill up or thicken the thinner portions or portions that appear to be noncontinuous. The corresponding column in Fig. 3 shows several connected portions become thicker and are connected after the closing operation. Closing can be realized by performing dilation first, and then erosion. This is shown in (4)

$$f \bullet SE = (f \oplus SE) \Theta SE. \quad (4)$$

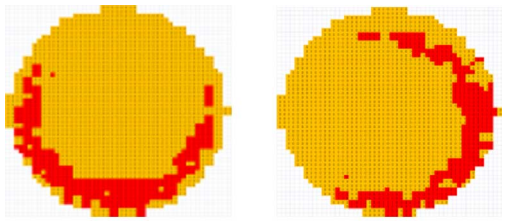
Shift: In order to generate the effect of pattern movement, the *shift* operation is considered for sample generation. According to the observation of domain experts, the patterns of WBMs may exhibit group movement, and this movement will also have high degree of similarity to the original (target) wafer map (i.e., the target wafer map); SVM is likely to have an incorrect judgment because it is based on the opposite position of two extremely different classes in the sample space, and then separates the sample points. Therefore, pattern shift is an essential method of sample generation in the context of using SVM in WBM. In this work, the shift direction is designed for eight positions: right, left, upper, down, upper right, lower right, upper left, and lower left. The shift size is also an adjustable parameter. In this case, the WBMs can move with a preset direction and range. Fig. 3 shows that the inner chips shift down for three steps (bins).

In addition, two other morphology-based operations are suggested by domain experts as follows.

Rotation: Due to a lack of constant positions or directions for wafer placement during the detection process, it is essential to determine the problem with the training error caused by incorrect corresponding placement. In this case, we use the principle of polar coordinates to generate the training samples with various rotation angles, where all bins rotate around the center in a preset counterclockwise angle within 0° to 360°. Fig. 5 shows that each chip rotates 90° counterclockwise around the center.

Density change: If we could dictate that one or more specific WBM areas will have different density, it would be convenient to modify the generated samples with many forms. More precisely, we cut the wafer into an $n \times n$ grid (here, $n = 9$, depending on users), and the selected bins in a specific grid have the following three options for density change: increase, decrease and random. For example, suppose that n equals 9, the degree of density change is 0.3 (this value is within 0 to 1), and we intend to increase/decrease/randomize the density of the

Authorized licensed use limited to: Hochschule Heilbronn. Downloaded on September 16,2025 at 05:20:29 UTC from IEEE Xplore. Restrictions apply.



(a) (b)

Fig. 5. An illustration of Rotation (a) The original wafer (b) 90° counterclockwise rotation.

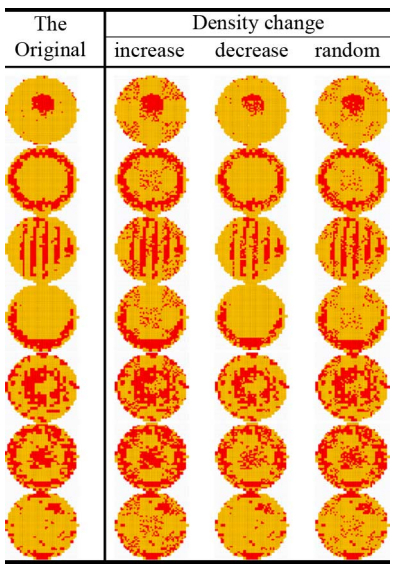


Fig. 6. An illustration of density change.

center grid (see Fig. 6), then 30% of bins in that grid will set 1/0/ 0–1 normal distribution.

We briefly summarize that the above seven morphological operations can be divided into two groups. The first group consists of the common existing methods, *erosion, dilation, opening, closing* and *shift*; the remainder includes our novel methods, *rotation* and *density change*.

B. Sample Filtering Utilizing One-Class SVM

One-class classification techniques are particularly useful in cases of *two-class* learning problems, whereby one of the classes, referred to as the *target class*, is well-sampled, whereas the other one, referred to as the *outlier class*, is severely undersampled [20]. The small number of examples from the *outlier* class may result from the fact that samples are too difficult to obtain from this class. Therefore, the goal of one-class classification is to construct a decision surface (function) around the examples from the target class in order to distinguish between the *target objects* and the *outliers* [21].

Indeed, similar training data were generated based on the inputs of domain experts. In addition, the similar part can be distinguished from the systematic defect patterns (SDP) by using a one-class classification technique with recognized similar samples, where SDP is a set of the specific morphology-based samples provided by the company. The one-class SVM can be applied to obtain the dissimilar training samples that are filtered from SDP. In this work the LIBSVM (version 3.11) was utilized [22] since it is an integrated tool for support vector classification and regression, which can handle the problem of one-class classification. We used the standard parameters of LIBSVM, chose the number of features (bins) and, by trial-and-error, selected the appropriate kernel and its appropriate parameters.

C. Support Vector Machine (SVM)

Similar to the one-class SVM, as a typical supervised learning method, the underlying theme of the SVM is to learn from the data. The difference between them is that the (Two-Class) SVM uses two parties of examples for training. Suppose that there is an input space X, an output space Y, and a training data set $TD = \{(x_1, y_1), (x_2, y_2), \ldots, (x_N, y_N)\} \subseteq (X \times Y)^N$, and N is the size of the training data. The output space $Y \in \{-1, +1\}$ determines the learning type, and leads to a binary classification problem. Geometrically, the basic concept behind SVM is to maximize the margin of separation of the hyperplane in a feature space.

SVM's creator, Vapnik [23], showed how training an SVM leads to a QP problem with bound constraints and linear equality constraints. This can be solved by constructing a Lagrangian function, and transforming it into the dual form, as shown in the following equation:

$$\begin{cases} \max L(\alpha) = \sum_{i=1}^{N} \alpha_i - \frac{1}{2} \sum_{i,j=1}^{N} \alpha_i \alpha_j y_i y_j \varphi(x)_i^T \varphi(x)_i \\ \qquad = \sum_{i=1}^{N} \alpha_i - \frac{1}{2} \sum_{i,j=1}^{N} \alpha_i \alpha_j y_i y_j K(x_i, x_j) \\ \text{s.t} \quad \sum_{i=1}^{N} \alpha_i^N \alpha_i y_i = 0, 0 < \alpha_i \leq Q, i = 1, 2, \ldots, N \end{cases} \tag{5}$$

where N stands for the size of training samples, i.e., the number of simulated samples; $x_p (p = 1, 2, \ldots, N)$ is a vector variable with dimension D, which denotes the number of wafer features; y is the response variable indicating whether a sample is similar to the target wafer. In addition, $\varphi(\cdot)$ is a feature map, and Q is an upper bound parameter controlling the tradeoff between margin maximization and tolerable classification errors. Parameters α_i are the so-called Lagrange multipliers, while the kernel function is defined as: $K(x_i, x_j) = \varphi(x_i)^T \varphi(x_j)$. The elegance of using the kernel function is evident in dealing with feature spaces of arbitrary dimensionality without having to explicitly compute the map function. The training was implemented by using the LIBSVM software [22], and the parameters were selected via a forced search process in the LIBSVM, where the parameters settings including N and SVM type: c-svc is automatically read in. In addition, after many trials, the most commonly used kernel function, the radial basis function (RBF), is also utilized here, and the mathematical formula is shown as (6)

$$K_{\mathrm{RBF}}(x_i, x_j) = \exp(\gamma |x_i - x_j|)^2 \tag{6}$$

where γ is the kernel parameter.

IV. Case Studies

To validate our method, we used real data in industry for comparison. The company provided two lots of the target wafer maps, namely, 395-bin and 1742-bin. The effectiveness was tested on these independent wafer maps from the two lots as well. The goal of this experiment was to classify the testing data into two classes, based on the similarity degree of training sample generation with respect to the given target wafer maps. We used the domain experts' judgments (called *Official*) as the measure standard, and compared the classified results with *Official*. The experimental results showed that MSVM achieved satisfactory performance, even considering the complex target maps. Moreover, in terms of the execution time, MSVM saved nearly a quarter of the time taken by the method currently used in the semiconductor industry.

A. The Method Used in the Semiconductor Industry

For the purpose of complete illustration, we briefly introduce the main idea of similar search used in the semiconductor industry, as

Authorized licensed use limited to: Hochschule Heilbronn. Downloaded on September 16,2025 at 05:20:29 UTC from IEEE Xplore. Restrictions apply.

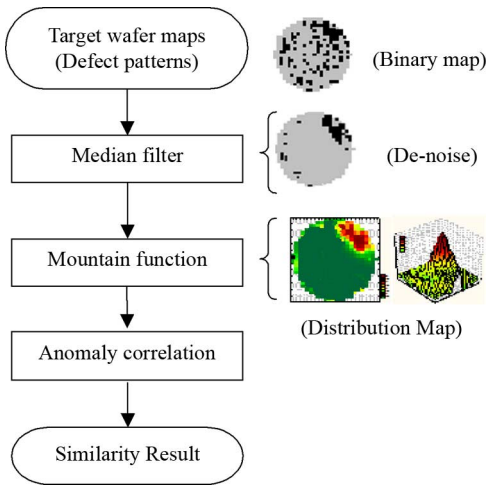


Fig. 7. Workflow of MMA.

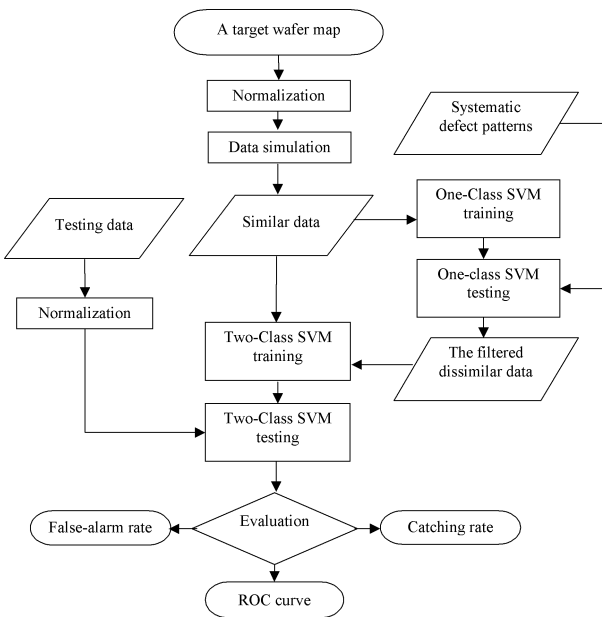


Fig. 8. Experiment framework for WBMs similar searching.

shown in Fig. 7. The workflow in Fig. 7 consists of three core steps: *median filter*[24], *mountain function*[25], and *anomaly correlation* (called MMA for short).

The *median filter* is often applied to noise reduction on an image or signal, which is a nonlinear digital filtering technique used to remove noise. This step is a typical preprocessing stage intended to improve the results of later processing (e.g., edge detection on an image) [24]. Next, the *mountain function* computes the density of the foreground pixels around a given point on an image in order to realize the whole distribution map. The definition and further details can be found in [25]. Finally, the *anomaly correlation* is used for the expression of similarity (7), where T is the number of testing samples, and g_i, t_i mean the *mountain values* of the target and testing wafers, respectively

$$\text{Anomaly correlation} = \frac{\sum_{i=1}^{T}(g_i - \bar{g})(t_i - \bar{g})}{\sqrt{\sum_{i=1}^{T}(g_i - \bar{g})^2 \sum_{i=1}^{T}(t_i - \bar{g})^2}}. \quad (7)$$

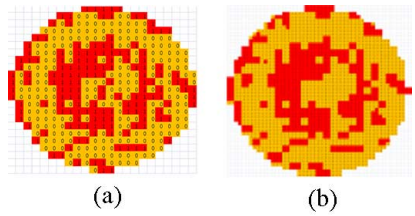


Fig. 9. Scale normalization: (a) original: 395-bin and (b) normalized: 1854-bin.

B. Experiment Framework

Fig. 8 describes the whole picture of the experimental execution. Initially, suppose the collaborator provides some target wafer maps, i.e., the main defect patterns, and we then have to use MSVM to deal with the subsequent wafer lots to be detected. For the source of (testing or target) wafer maps, each lot of wafers may have different scale; in this case, scale normalization is essential preprocessing in order to avoid the problem of size adjustment during the training phase. More precisely, scale normalization not only makes a specific batch of wafers for a similarity search, but also facilitates the subsequent SVM training phase. Currently, the most adaptive size is 1854-bin, since it is convenient for the transformation of both large and small bin-size wafer maps in the industry. As seen from Fig. 9, it is the case that a wafer with a scale of 395-bin transformed into one of 1854-bin. For the process of scale normalization, for example, when the 395-bin is transformed into 1854-bin, the prototype of 395-bin is divided into a grid with 1854 bins, where the chips in 1854-bin are decided to be red (defective chips) only when the red parts of original 395-bin occupy arbitrary area inside the chips of 1854-bin. For smooth processing, this preprocessing is combined with the wafer generation procedure in the industry.

Next, we proceed with the generation of morphological training samples with a procedure of characteristic separation for similarity confirmation (based on domain experts' judgments) and dissimilar samples (One-Class SVM filtering) (see Fig. 10). After the two-class SVM training phase, the normalized testing data are sent for similarity search, in order to classify them into classes 1 (similar to the target) and 2 (dissimilar to the target).

To evaluate the classification performance, the *Receiver Operating Characteristic* (ROC) curve is used to observe the relevance of the catching rate (true positive rate, X axis in ROC), and the false-alarm rate (false positive rate, Y axis in ROC). The formal definitions are presented in Fig. 11 and eqs. (8) and (9). An ideal classifier is required to produce an ROC curve above the diagonal line; furthermore, the area below the ROC curve serves to judge the degree of performance: the larger the better. In addition, from a practical viewpoint in the industry, the catching rate is anticipated to be more than 0.8, and the false-alarm rate under 0.1

$$\text{Catching rate} = \frac{\text{TP}}{(\text{TP} + \text{FN})} \quad (8)$$

$$\text{False} - \text{alarm rate} = \frac{\text{FP}}{(\text{FP} + \text{TN})}. \quad (9)$$

C. Results

This work was accomplished in collaboration with a semiconductor firm in Taiwan, assisting them in dealing with defective wafer detection. The wafer data from the firm has been divided into two main types: simple and complex patterns. The two types contain five and four patterns, respectively. Fig. 12 draws the target wafer maps of those patterns and several corresponding training samples. As for the number of

Authorized licensed use limited to: Hochschule Heilbronn. Downloaded on September 16,2025 at 05:20:29 UTC from IEEE Xplore. Restrictions apply.

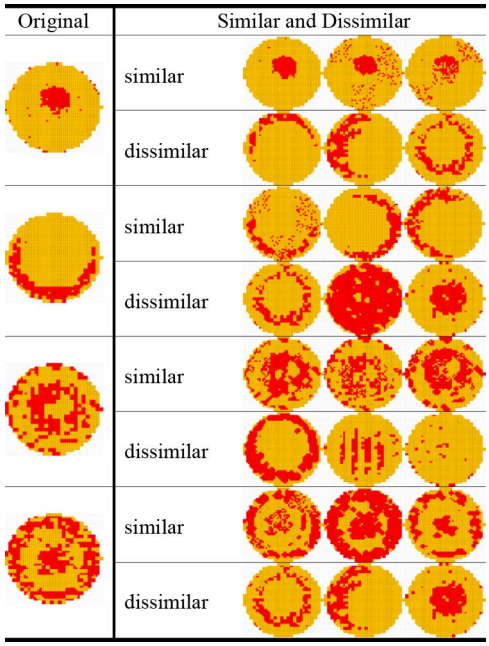


Fig. 10. An illustration of similar and dissimilar samples.

Items	# classifying into Positive	# classifying into Negative
Positive data	True Positive (TP)	False Negative (FN)
Negative data	False Positive (FP)	True Negative (TN)

Fig. 11. Confusion matrix.

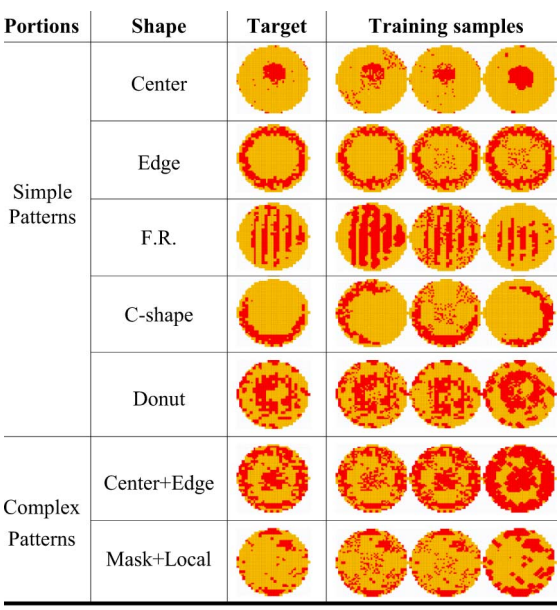


Fig. 12. An illustrations of created samples.

training samples, for each pattern the similar part has 234 wafers, and the dissimilar part has nearly 234, because of the filtering process.

1) Simple Patterns: Table I presents the results of simple patterns. The amount of testing data depends on the stock situation, and we also provided the *official* results, as judged by the domain experts. The findings are given as follows.

TABLE I
RESULTS OF SIMPLE-PATTERNS

Simple-Shape	# Testing	# Official	#MSVM	Catching rate	False-alarm rate
Center	444	52	57	48/52=0.92	9/392=0.02
Edge	444	69	81	62/69=0.89	19/375=0.05
F.R.	404	50	54	36/50=0.72	18/354=0.05
C-shape	404	121	143	117/121=0.96	26/283=0.09
Donut	278	15	25	11/15=0.73	14/263=0.05
Total	1974	307	360	274/307=0.89	86/1667=0.05

TABLE II
RESULTS OF COMPLEX-PATTERNS

Complex-Shape	# Testing	# Official	#MSVM	Catching rate	False-alarm rate
Center+Edge	492	36	46	32/36=0.88	14/456=0.03
Mask+Local	650	57	85	56/57=0.98	29/593=0.05
Total	1142	93	131	88/93=0.95	43/1049=0.04

TABLE III
RESULTS OF AVERAGE CPU TIME COST

Panel *A*: MSVM CPU cost		
Items		Time expense
Morphology-based sample simulation	Size, Density, Shift	12s
	Rotation	36s
Training Phase		2m 3s
Testing Phase		2s
Total		2m 53s
Panel *B*: MMA CPU cost		
Items		Time expense
Median Filter		>5s
Mountain Function		10m 36s
Anomaly correlation		>3s
Total		>10 m 44s

For the *Center* case, the MSVM reaches excellent performance since it is easier to identify the pattern. *Edge* and *C-shape* are highly similar, with the only difference being the circular gap. Once the gap is smaller than the semicircle, MSVM may generate a slight false-alarm due to the misjudgment resulting from the higher similarity to *Edge*. Moreover, for the *Edge* and *C-shape* cases, the rotation feature needs to be considered because of the placement problem mentioned in *rotation*, Section III. In general, though, both have satisfactory performance, especially *C-shape*.

The *Donut* is scarce in stock, and thus has a slightly low catching rate owing to the small number of similar samples compared with other patterns. The last simple pattern is $F.R$. Because the strips have various degrees of coarseness and arbitrary distributions, $F.R$, has many characteristics of other simple patterns. We believe that precision could be increased by offering specific definitions of WBM patterns or improved training machines. Thus, $F.R$, presents a challenging work for future endeavors.

The comparison of simple patterns between MSVM and MMA can be referred to the ROC curves (see Fig. 13(a) to (e)) in terms of the area below the curve. Because the MMA cannot perform the experiment to detect all the patterns with its current approach, some simple cases (*C-Shape* and *Donut*) only present the results of MSVM. The main limitation of MMA is that it lacks the rotation features in support of the consideration of different angle aspects that *C-Shape* and *Donut* require. As seen from Fig. 13(a) to (e), *Edge*, MSVM, and MMA have ideal and close performance. For the remaining simple cases, however, MSVM obviously achieves a higher catching rate than does MMA under all false-alarm rates; MSVM has a lower false-alarm rate under all catching rates.

Authorized licensed use limited to: Hochschule Heilbronn. Downloaded on September 16,2025 at 05:20:29 UTC from IEEE Xplore. Restrictions apply.

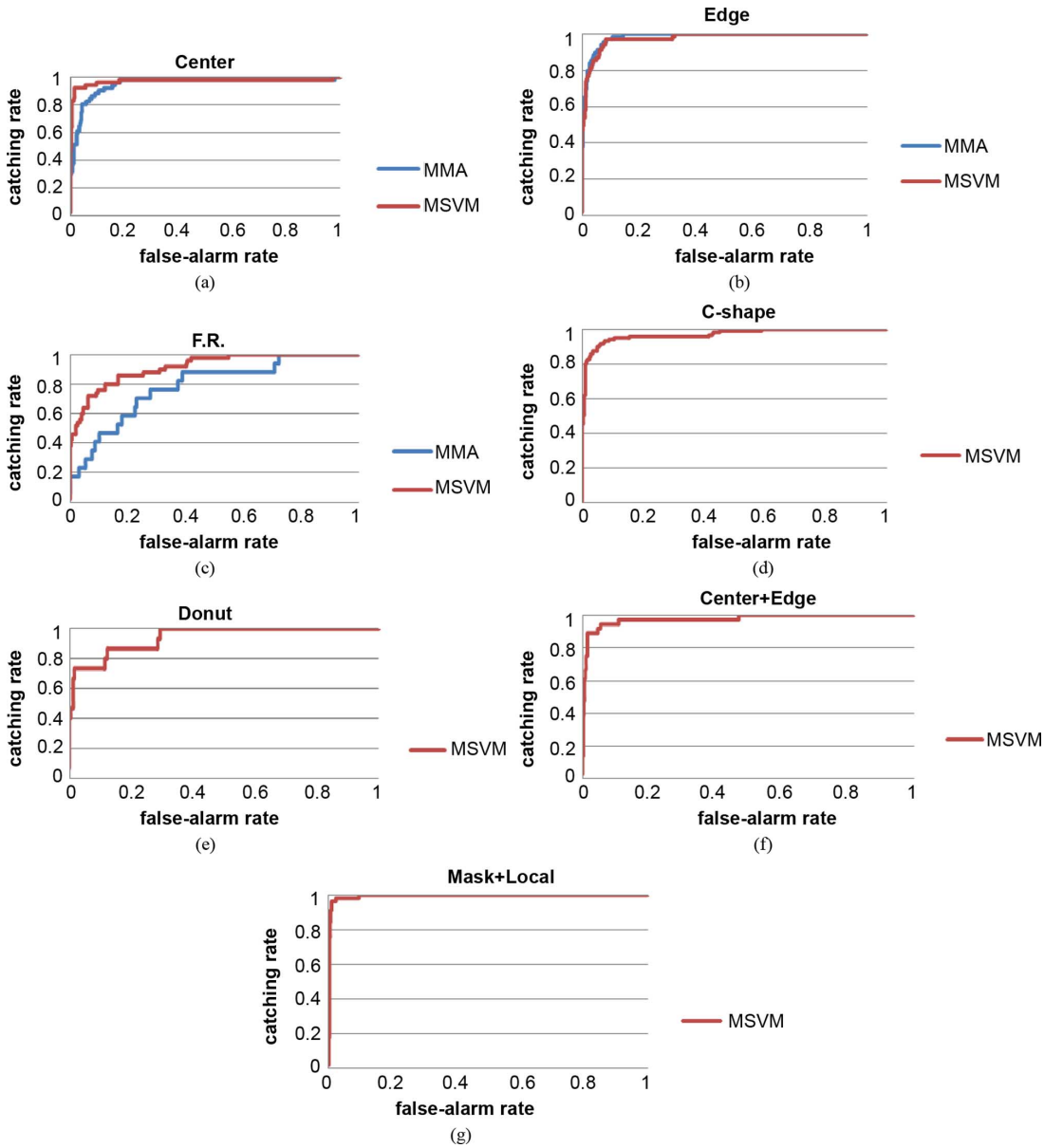


Fig. 13. The performances in all testing patterns. (a)–(e) Simple patterns. (f) and (g) Complex patterns.

2) Complex Patterns: With constantly developing manufacturing technology, engineers have to deal with ever larger scales and more diverse WBMs. Moreover, defective wafer maps may generate complex patterns. We tested our method, MSVM on the two complex patterns that have not been successfully detected by MMA or any other machine learning-based methods. The experimental results are presented in Table II and Fig. 13(f) and (g), where they show ROC curves with large areas under the curves, while the catching rate increases to 1 when the false-alarm value is around 0.05. All told, MSVM, on average, obtains a high catching rate of nearly 0.95, and a low false-alarm rate of below 0.05.

D. Result Summary

In the literature, there are few methods contributing to wafer detection using similarity search in the semiconductor industry. Similarity searching must consider variations of many features. This problem is very different from those more commonly related to pattern recognition. In addition, significant CPU time would be expended if those pattern recognition approaches such as MMA were to take all of the features into consideration. As a result, MMA can only handle fewer simple-pattern cases.

Table III presents the average execution time in each step of MMA and MSVM. MMA and MSVM were performed on computer facilities: an Intel E7500, 2.93 GHz PC with 1.96 GB memory, and an Intel (R) Core(TM) i5, 3.20 GHz PC with 4 GB memory, respectively. It can be observed that MSVM saves nearly a quarter of the time taken by MMA; more importantly, MSVM has a better wafer detection performance.

V. Discussions

In this study, the similarity measurement for evaluating the performance of the proposed approach was only compared with MMA because of the confidential reason. There are some spatial characteristics in wafer maps, which have been discussed in the literature [26], [27]. These characteristics such as correlogram may capture defective patterns better than the two-dimension wafer map data [26]. More precisely, spatial statistic can deal with bin-level wafer maps by comparing

Authorized licensed use limited to: Hochschule Heilbronn. Downloaded on September 16,2025 at 05:20:29 UTC from IEEE Xplore. Restrictions apply.

the number of functional bins around a defective bin with the number of defective bins around a functional bin. Although the spatial statistics performed well in previous work, there are several critical issues in similarity measurement when considering complicated wafer features such as rotation and shifts. For example, in terms of rotation, wafers are actually not a perfectly round shape, so some bins need to be inserted or ignored concerning the rotation bins. In this case, the spatial statistic measurement cannot provide an effective similarity measurement because of the diversity of wafer maps. The future studies will incorporate the comments and judgment from the experienced engineers into the tool, and focus on a more general similarity measurement, especially for the complex shapes.

VI. Conclusion

This work has proposed a novel approach for combining a supervised SVM classifier with a morphology-based sample simulation for similarity searching of binary WBM defect patterns for yield enhancement. Owing to the increasing size and variations of defective wafer patterns, conventional pattern recognition or classification methods have difficulty in determining the pattern types. According to the original wafer characteristics of the target wafer maps, this study generated wafer samples through morphology-inspired conceptions. This proposed morphology-based SVM (MSVM) extends to more complex defect patterns (shapes). More precisely, instead of traditional methods, the proposed MSVM has been employed to engage in similarity search through a morphological learning process, which is built in scale normalization, samples filtering process, and seven types of similar samples generation approach. The experimental results showed that MSVM not only has great performance in terms of ROC curve (catching and false-alarm rates), but has faster execution time.

References

[1] Q. Zhou, L. Zeng, and S. Zhou, "Statistical detection of defect patterns using Hough transform," *IEEE Trans. Semicond. Manuf.*, vol. 23, no. (3), pp. 370–380, Aug. 2010.

[2] C. F. Chien, J.-Z Wu, and C.-C. Wu, "A two-stage stochastic programming approach for new tape-out allocation decisions for demand fulfillment planning in semiconductor manufacturing," *Flexible Services Manuf. J.*, vol. 25, no. 3, pp. 286–309, 2013.

[3] C. Chien, W. Wang, and J. Cheng, "Data mining for yield enhancement in semiconductor manufacturing and an empirical study," *Expert Syst. With Applicat.*, vol. 33, no. 1, pp. 1–7, 2007.

[4] J. G. Shanthikumar, S. Ding, and M. T. Cheng, "Queueing theory for semiconductor manufacturing systems: A survey and open problems," *IEEE Trans. Autom. Sci. Eng.*, vol. 4, no. 4, pp. 513–522, Oct. 2007.

[5] R. C. Leachman and S. Ding, "Excursion yield loss and cycle time reduction in semiconductor manufacturing," *IEEE Trans. Autom. Sci. Eng.*, vol. 8, no. 1, pp. 112–117, Jan. 2011.

[6] C. H. Stapper, "The effects of wafer to wafer defect density variations on integrated circuit defect and fault distributions," *IBM J. Res. Develop.*, vol. 29, no. 1, pp. 87–97, 1985.

[7] A. K. Jain, R. P. W. Duin, and J. Mao, "Statistical pattern recognition: A review," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 22, no. 1, pp. 4–37, Jan. 2000.

[8] B. Bhanu, S. Lee, and J. Ming, "Adaptive image segmentation using a genetic algorithm," *IEEE Trans. Syst. Man Cybern.*, vol. 25, no. 12, pp. 1543–1567, Dec. 1995.

[9] P. Farzin, N. Sulaiman, S. Roosta, M. H. Marhaban, and R. Ramli, "Design a new sliding mode adaptive hybrid fuzzy controller," *J. Adv. Sci. Eng. Res*, vol. 1, pp. 115–123, 2011.

[10] H. Kim, K. Lee, B. Jeon, and C. Song, "Quick wafer alignment using feedforward neural networks," *IEEE Trans. Autom. Sci. Eng.*, vol. 7, no. 2, pp. 377–382, Apr. 2010.

[11] C.-J Kuo, C.-F Chein, and J.-D Chen, "Manufacturing intelligence to exploit the value of production and tool data to reduce cycle time," *IEEE Trans. Autom. Sci. Eng.*, vol. 8, no. 1, pp. 103–111, Jan. 2011.

[12] J. Wilfredo, Puma-Villanueva, E. P. dos Santos, and F. J. V. Zuben, "A constructive algorithm to synthesize arbitrarily connected feedforward neural networks," *Neurocomputing*, vol. 75, pp. 14–32, 2012.

[13] C. Cortes and V. Vapnik, "Support vector networks," *Mach. Learning*, vol. 20, no. 3, pp. 273–297, 1995.

[14] M. Farhan, G. Kassem, M. Abdullah, and S. Akbar, "Support vector machine classifier for pattern recognition," in *Proc. 1st Int. Conf. Informat. Comput. Intell. (ICI)*, 2011, pp. 272–277.

[15] X. Ji, Y. Li, Z. Wang, F. Wang, and Q. Liu, "Partial discharge pattern recognition of XLPE cable connector based on support vector machine," in *Proc. Int. Conf. Elect. Control Eng. (ICECE)*, 2011, pp. 2422–2425.

[16] S- C. Hsu and C.-F Chien, "Hybrid data mining approach for pattern extraction from wafer bin map to improve yield in semiconductor manufacturing," *Int. J. Prod. Economics*, vol. 107, pp. 88–103, 2007.

[17] T.-S Li and C.-L. Huang, "Defect spatial pattern recognition using a hybrid SOM-SVM approach in semiconductor manufacturing," *Expert Syst. Appl.*, vol. 36, pp. 374–385, 2009.

[18] L.-C Chao and L.-I. Tong, "Wafer defect pattern recognition by multiclass support vector machines by using a novel defect cluster index," *Expert Syst. Appl.*, vol. 36, pp. 10158–10167, 2009.

[19] G. Matheron and J. Serra, "History of Mathematical Morphology," 1968 [Online]. Available: http://cmm.ensmp.fr/~serra/pdf/birth_of_mm.pdf

[20] R. Perdisci, G. Gu, and W. Lee, "Using an ensemble of one-class SVM classifiers to harden payload-based anomaly detection systems," in *Proc. IEEE 6th Int. Conf. Data Mining Location*, Hong Kong, 2006, pp. 488–498.

[21] I. F. de Viana, P. J. Abad, J. L. Alvarez, and J. L. Arjona, "Toward one class classifier techniques applied to verifier information," in *Proc. 6th Iberian Conf. Inf. Syst. Technol. (CISTI)*, 2011, pp. 1–7.

[22] LIBSVM [Online]. Available: http://www.csie.ntu.edu.tw~cjlin/libsvm/

[23] V. N. Vapnik, *The Nature of Statistical Learning Theory*. New York, NY, USA: Springer-Verlag, 1995.

[24] J. Principe, N. Euliano, and W. Lefebvre, *Neural and Adaptive Systems: Fundamentals Through Simulations*. New York, NY, USA: Wiley, 2000.

[25] R. R. Yager and D. P. Filev, "Approximate clustering via the mountain method," *IEEE Trans. Syst. Man Cybern.*, vol. 24, pp. 1279–1284, Aug. 1994.

[26] Y.-S. Jeong, S.-J. Kim, and M. K. Jeong, "Automatic identification of defect patterns in semiconductor wafer maps using spatial correlogram and dynamic time warping," *IEEE Trans. Semicond. Manuf.*, vol. 21, no. 4, pp. 625–637, Nov. 2008.

[27] T. Yuan, W. Kuo, and S. J. Bae, "Detection of spatial defect patterns generated in semiconductor fabrication processes," *IEEE Trans. Semicond. Manuf.*, vol. 24, no. 3, pp. 392–403, Aug. 2011.

Authorized licensed use limited to: Hochschule Heilbronn. Downloaded on September 16,2025 at 05:20:29 UTC from IEEE Xplore. Restrictions apply.