Expert Systems With Applications 260 (2025) 125378

Contents lists available at ScienceDirect

Expert Systems With Applications

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

A framework for detecting unknown defect patterns on wafer bin maps using active learning

Jin-Su Shin [a,b], Min-Joo Kim [a,b], Dong-Hee Lee [c,*]

[a] Department of Semiconductor and Display Engineering, Sungkyunkwan University, 2066, Seobu-ro, Suwon-si, Gyunggi-do 16419, Republic of Korea
[b] Memory Division, Samsung Electronics Co, Ltd., 1-1 Samsungjeonja-ro, Hwaseong-si, Gyeonggi-do 18448, Republic of Korea
[c] Department of Industrial Engineering, Sungkyunkwan University, 2066, Seobu-ro, Suwon-si, Gyunggi-do 16419, Republic of Korea

ARTICLE INFO

Keywords:
Wafer bin map classification
Active learning
Real field data
Unknown defect patterns on wafer bin map
Convolutional neural network
Human in the loop

ABSTRACT

In a semiconductor manufacturing process, it is important to detect and classify defect patterns in Wafer Bin Maps (WBMs) and identify the root cause of these defects for tight quality control. Recently, various deep learning methods have been applied, but these methods suffer from poor classification, limited data labelling, and the inability to detect and learn new defect patterns. Moreover, the methods prioritize improving the accuracy and speed of classification models over detecting and classifying unknown defect patterns. Against this background, we developed an abnormal pattern detector based on a One-Class Support Vector Machine (SVM) that classifies whether defect patterns are known or unknown. For the known patterns, we used transfer learning based on a ResNet50 classifier pre-trained with ImageNet1K data to further classify the defect pattern in the WBM. For the unknown patterns, clustering is performed using Density-Based Spatial Clustering Application with Noise (DBSCAN) to assign new labels, and the classifier is updated through active learning. This enables the detection of unknown patterns and effectively updates the abnormal detector and pre-trained classifier even during the use of the classifier. Experiment results from the WM-811K dataset verify that the proposed method can detect unknown patterns while maintaining excellent classification performance for known patterns. Moreover, it can continuously maintain the high classification performance of the detector and pre-trained classifier through the active learning. Also, applicability in real semiconductor manufacturing environments was demonstrated using real industrial data with an unknown pattern ("Eye Defect Pattern"), not included in the WM-811K dataset.

1. Introduction

The semiconductor manufacturing industry remains a key driving force behind modern technological development and continues to bring revolutionary changes to various aspects of our daily lives. To increase the processing power and functionality of chips, semiconductor research aims to integrate many devices in a single chip. Additionally, modern semiconductor manufacturing facilities consistently adopt state-of-the-art instrumentation and processes capable of producing devices in nanoscale. Consequently, aggressive miniaturization of semiconductor circuits poses significant manufacturing challenges. Nanoscale semiconductor fabrication processes increasingly influence the manufacturing cost of wafers by introducing new defects (Zhu et al., 2022). Moreover, when previously unknown, mixed, and complex defects occur, process yields decrease, and the stability of the manufacturing process weakens (Kim & Behdinan, 2023). Therefore, the importance of technology to early detect unknown defects and predict potential defects is becoming more pronounced (Batool et al., 2021).

Wafer bin maps (WBM) are data sets generated from the results of electrical die testing and sorting. WBMs provide visual information about the spatial distribution of failed chips. As shown in Table 1, the type of WBM defect pattern is closely related to the root cause of the defects. Therefore, accurately classifying and labeling defect patterns in WBMs can help pinpoint the specific fabrication process causing the defects. However, accurate classification and labeling of unknown defect patterns in WBMs still rely heavily on the experience and intuition of process engineers, as well as human vision (Hsu et al., 2020). Various image classification technologies utilizing pattern recognition and deep learning are actively being investigated to efficiently classify defect patterns in WBMs (Chen et al., 2022). Most of these studies rely on

* Corresponding author.
E-mail addresses: jinsushin@g.skku.edu (J.-S. Shin), havefun2@g.skku.edu (M.-J. Kim), dhee@skku.edu (D.-H. Lee).

https://doi.org/10.1016/j.eswa.2024.125378
Received 6 April 2024; Received in revised form 24 August 2024; Accepted 9 September 2024
Available online 11 September 2024

Table 1
Various WBM defect patterns and corresponding causes. (WM-811K Class).

Defect Name	WBM Image	Defect Pattern Description	Source of defects
Center		Defects concentrated in the center of the wafer	Irregular radio frequency (RF) operation (Hansen & Thyregodb, 1998; Hsu et al., 2020)or unusual liquid flow
Donut		Defect pattern shaped like a ring with an opening in the middle	Accumulation of residues resistant to removal during photoresist cleaning
Edge-Loc		Defects clustered on the edge of the wafer	Irregular temperature annealing (Hansen et al., 1997) or impurities in the load lock valve
Edge-Ring		Defects located along the edge of the wafer	Anomalous temperature regulation during rapid thermal process (RTP) (Tello et al., 2018) or cold process, and insufficient heating due to wafer thickness variations
Loc		Defects clustered on the inner region of the wafer	Vacuum pressure differences due to slit valve leaks or poor pump operation, and vibration of internal parts Hansen & Thyregodb, 1998)
Near-Full		Entire wafer is covered with defects	Wafer surface photoresist (PR) rupture due to electron overcharge or abnormalities in the plasma ion beam implant process
Scratch		A pattern of defects where defects are continuously connected within the wafer	Scratches on the wafer surface by transfer robots during the wafer handling sequence, surface damage by humans, or scratches caused by polishing during chemical–mechanical polishing (CMP) (Kim et al., 2021)
Random		Defects occur sporadically	Abnormal vacuum or gas or fluid eruption

supervised learning (Cha & Jeong, 2022; Chen et al., 2022; Kim et al., 2021; Kim et al., 2023; Nag et al., 2022; Nakazawa & Kulkarni, 2018; Shinde et al., 2022; Xu et al., 2022; Yu et al., 2019; Bae & Kang, 2023; Yu et al., 2021), which assumes knowledge of all correct labels in advance. These studies have mainly focused on improving the performance of classifiers using state-of-the-art deep learning techniques or downscaling the model weights and training time (Hsu et al., 2020; Yu et al., 2021). To overcome the limitations of such strategies, studies on defect pattern identification based on unsupervised learning are being pursued (Wu et al., 2015; Lee et al., 2021) assuming no labels for all data or utilizing semi-supervised learning in situations where labeling is limited and data is scarce (Lee & Kim, 2020; Shim et al., 2020; Yu & Liu, 2021; Kang & Kang, 2023).

Generally, supervised learning with correctly labeled data can achieve high classification performance. However, the classification algorithm used in the final dense layer of the classifier, such as SoftMax functions, relies on statistical methods to infer the correct answers based on previously given labels. Therefore, supervised learning is unable to distinguish undefined defects. On the other hand, unsupervised learning allows the verification of undefined defects without labels. However, process engineers must manually check and verify numerous samples to ensure classification accuracy, and classification performance may be lower than that for supervised learning. Semi-supervised learning, frequently applied in WBM classification to enhance labeling efficiency and address data scarcity issues, depends on similarities with existing labels. Assigning an accurate label when encountering unknown defect patterns is challenging. Consequently, this decreases label assignment accuracy or reduces the ability to detect unknown defect patterns.

In this study we propose a comprehensive detection and classification framework that has high classification accuracy for known defect patterns and the ability to correctly detect unknown defect patterns. To realize this, we introduced new types of defects that could potentially occur in the test data (i.e., unknown defect patterns) and use an abnormal detector based on a One-Class Support Vector Machine (SVM) to classify whether the defect patterns are known or not. For samples classified by known defect patterns, we used transfer learning based on a ResNet50 classifier pre-trained with ImageNet1K data to create a WBM defect pattern classifier, and then the defect pattern of the WBM is classified using the classifier. For samples classified as having unknown defect patterns, clustering was performed using Density-Based Spatial Clustering Application with Noise (DBSCAN), then, new labels were assigned, and the model was updated through an active learning process. Active learning can continuously maintain the high classification performance of the detector and pre-trained classifier. To verify the effectiveness of the proposed method, its performance was validated using actual semiconductor manufacturing industry data and an open WBM dataset called WM-811K.

The remaining part of this paper is structured as follows. Section 2 reviews current trends in WBM classification. Section 3 describes the proposed method followed by an explanation of the architectures used in each module. In Section 4, the effectiveness of the proposed method is demonstrated using the open dataset WM-811K. Section 5 verifies the practical applicability of the proposed method in real semiconductor manufacturing sites. This is achieved by utilizing a classifier trained solely on the WM-811K dataset to validate real industrial data containing unknown defect patterns ("Eye Defect Pattern") not included in the WM-811K. Finally, Section 6 concludes the paper and briefly outlines future research directions.

2. Related works

Recent studies on WBM defect pattern detection are predominantly based on machine learning and deep learning technologies. The algorithms and datasets used in this study, as well as the classification accuracy metrics for each model, are summarized in Table 2. The learning algorithms can be broadly classified into three categories: supervised learning, unsupervised learning, and semi-supervised learning. This section summarizes the limitations of existing studies in detecting unknown defect patterns.

2.1. Supervised learning

Supervised learning trains data based on predefined and accurate labels, which incurs high costs. Supervised learning typically achieves high accuracy because it trains the model based on the correct labels in the data. However, it also has the disadvantage of not being able to predict unlabeled data that is not predefined. As shown in Table 2, a large number of WBM studies are based on supervised learning. These studies applied state-of-the-art deep learning algorithms to achieve high accuracy of defect pattern classification in WBMs with lower computational requirements and faster classification speed. For example, Yu et al. (2019) and Nakazawa and Kulkarni (2018) presented classification methods using CNN, the simplest deep learning model, to classify WBM defect patterns. Chen et al. (2022) sought to improve classification performance by compensating for the imperfect selection of a single DCNN model using two DCNN models in parallel to make the final decision. Cha and Jeong (2022) proposed a classifier based on U-NET with a Convolutional Block Attention Module (CBAM) module to improve the classification performance. Nag et al. (2022) and Xu et al. (2022) improved the existing CBAM module to address unbalanced datasets and improve the classification performance. Nag et al. (2022) proposed WaferSegClassNet (WSCN), which combines classification and segmentation of WBM defect patterns to maximize classification performance.

Table 2
WBM classification methods and data sets used in existing studies.

Category	Reference	Algorithm	Data Set	Model accuracy
Supervised	Chen et al. (2022)	Dual DCNN (Deep Convolutional Neural Network)	WM-811K	Micro Acc: 0.98
	Nakazawa and Kulkarni (2018)	CNN	Simulated Wafer Map (Total 22 Mixed Patterns)	Macro Acc: 0.94
	Yu et al. (2019)	CNN/Similarity Rank (PCA)	WM-811K	MacroAcc: 0.93
	Xu et al. (2022)	ResNet 18(With Improved CBAM)	WM-811K	Macro Acc: 0.95
	Kim et al. (2021)	ResNet50	1238 Wafers from industry	Macro Acc: 0.92
	Nag et al. (2022)	WSCN (WaferSegClassNet)	WM38 (Mixed Type)	Macro Acc: 0.98
	Shinde et al. (2022)	YOLO V4	WM-811K	Macro Acc: 0.96
	Yu et al. (2021)	DenseNet-GCForest	WM-811K	Macro Acc: 0.96
	Bae and Kang (2023)	Modified VGG16	WM-811K & WM38 (Mixed)	–
	Cha and Jeong (2022)	U-Net (With CBAM)	WM-811K & WM38 (Mixed)	Macro Acc: 0.96
Unsupervised	Wu et al. (2015)	SVM/Similarity Ranking	WM-811K	Macro Acc: 0.83 Micro Acc: 0.95
	Lee et al. (2021)	Clustering DPGMM/Similarity Rank	WM-811K, Industry data	–
	Wang and Wang (2023)	Tensor voting/Similarity Rank	WM38 (Mixed Type)	–
Semi-Supervised	Shim et al. (2020)	CNN with Active Learning	WM-811K	–
	Yu and Liu (2021)	PCACAE (PCA Convolutional Auto Encoder)	WM-811K	Macro Acc: 0.94 Micro Acc: 0.97
	Kang and Kang (2023)	SSRL (Semi-Supervised Rotation-invariant Representation Learning)	WM-811K	–
	Lee and Kim (2020)	SS-CDGMM (Semi-Supervised Convolutional Deep Generative Model with Multiple Discriminative Networks)	Circle, Scratch, Partial Ring, Local Zone, and Mixed Pattern	EMR Acc: 0.94
Detect Out-of-Distribution	Kim et al. (2020)	VGGNet	11,789 13 pattern in-distribution 1899 out of distribution	Macro Acc: 0.97

They also investigated supervised learning-based segmentation for WBM classification. Shinde et al. (2022) demonstrated that it is possible to use object detection to predict the location of defects and exploit that information. They proposed a network with a backbone based on the YouOnlyLookOnce (YOLO) algorithm. In addition, Kim et al. (2021) and Bae and Kang (2023) proposed model learning using various data augmentation methods and contrastive learning techniques to solve the problem of imbalance between the labeled data and unlabeled data used for learning.

The main objective of these studies was to enhance both learning performance and classification accuracy for WBM defect patterns using supervised learning-based models. However, training with supervised learning requires correct answer labels and balanced datasets. Due to the nature of semiconductor manufacturing that numerous wafers are manufactured and tested every day, unbalanced and incomplete datasets are created due to random sampling quality inspection. This makes it very challenging to extract accurate labels and appropriate defect data. Additionally, unknown defect patterns that are not previously defined as defect patterns are not detected. Consequently, pre-trained classifiers need to be continuously retrained, increasing losses and costs. Therefore, most supervised learning-based methods cannot be applied directly to actual semiconductor manufacturing conditions.

2.2. *Unsupervised learning*

Unsupervised learning utilizes techniques such as clustering and dimensionality reduction to learn relationships between patterns or structures extracted from data, under the assumption that accurate labels do not exist for all data. It does not require precise labels for all cases and reflects the nature of domains that acquiring data for all defect patterns are challenging. For this reason, research on WBM defect classification using unsupervised learning mainly focuses on similarity comparison based on the relationships between defect patterns in WBM, as shown in Table 2.

Wu et al. (2015) employed a Support Vector Machine (SVM) to initially classify defect patterns. The Euclidean distance between each wafer map was then utilized to provide a similarity ranking, assisting process engineers in making accurate judgments about defects. Similarly, Lee et al. (2021) proposed a method for judging defects by projecting each WBM into polar coordinates, calculating the weight vector value of each WBM, and then comparing the similarities between the weight vectors of the clustered data and the weight vectors of the predefined defect types using the Dirichlet Process Gaussian Mixture Model (DPGMM) technique to present the top N rankings. Additionally, Wang and Wang (2023) compared the degree of similarity between wafer maps using tensor voting techniques.

These unsupervised learning techniques enable visualization of unknown defect patterns that can potentially occur. However, because

Table 3
Comparison of active learning techniques used in WBM study.

Reference	Active learning type	Active learning phase	Assumption about data	Number of iterations	Objective
Kong and Ni, (2020) Shim et al., (2020) Manivannan, (2024)	Pool Based	Train Phase (Offline)	Closed Set (Known Patterns only)	Until all WBM is under the entropy threshold Maximum 100 iterations Until the pre-specified labeling budget	Reduce labeling costs & increase the efficiency of labeling.
Proposed Method	**Stream Based**	**Test Phase (Online)**	**Open Set (Known & Unknown Patterns)**	**Single Pass or None Pass**	**Detect unknown patterns & prevent misclassification**

these unsupervised learning-based methods are trained without accurate labels for the defect patterns, they generally lack classification accuracy and speed compared to supervised learning. Another drawback is that engineers have difficulty in accurately assessing the quantitative performance of their models and must manually inspect large amounts of data to interpret classification results for a single wafer map image. New strategies are needed to address the limited accuracy of classification results, the long detection time, and the high cost of identifying the occurrence of unknown defect patterns.

2.3. Semi-supervised learning

Semi-supervised learning is a machine learning strategy that is halfway between supervised and unsupervised learning. This method is mainly useful when labeled data is limited. Because it is difficult to acquire and label the WBM data, this method can be a good alternative for WBM classification. Some WBM classification methods based on semi-supervised learning reveal that model training can proceed effectively even in situations with insufficient labels. This can be achieved by utilizing efficient data labeling methods or by analyzing the similarities between existing labels to improve predictions for unlabeled data.

Shim et al. (2020) and Kang and Kang (2023) proposed an active learning technique and a semi-supervised rotation-invariant representation learning method, respectively, to effectively collect sufficient labels from training data and train a CNN model with high accuracy. Yu and Liu (2021) proposed a Principal Component Analysis-based Convolutional Auto Encoder (PCACAE) to overcome insufficient training data and unbalanced data. Lee and Kim (2020) assumed a total of 16 types of mixed defects based on four predefined types of single defects such as circle, scratch, partial ring, and local zone, and converted the single class classification problem into a multi-label class problem based on their similarity to single defects. They then proposed a learning model for unlabeled mixed-type defect data. Kim et al. (2020) tried to detect out-of-distribution (ODD) defects using VGGNet. However, their approach required data on both In-Distribution (IND) and ODD defects of interest that must be trained together.

Existing semi-supervised learning studies have proposed efficient classifier learning techniques utilizing limited data labels. However, they cannot identify completely new types of defect patterns unrelated to existing defect patterns. This is because they are dependent on similarities with existing defect types or assume the type and distribution of defect patterns that could occur in advance prior to learning. Furthermore, the generation of new defect labels can lead to inaccurate labeling. Ultimately, a method for implementing generalized models for unknown defect patterns and applying them to real-world semiconductor wafer fabrication is still needed.

2.4. Active-learning

Active learning is a subfield of machine learning that involves selectively labeling the most informative data points during the model's training process. This technique is particularly useful when labeled data is scarce or labeling costs are high. Active learning can be categorized into pool-based and stream-based approaches based on how data is processed and sampled (Cacciarelli & Kulahci, 2024). The pool-based approach involves selecting the most informative instances from a large pool of unlabeled data to request labeling. This method allows the algorithm to evaluate all samples simultaneously, enabling optimal sample selection considering the overall data distribution. However, it can be computationally expensive due to the need to process large amounts of data and may not be suitable for online or streaming data environments. Additionally, it can only train the model within the known classes and the scope of the pre-acquired data. On the other hand, stream-based active learning evaluates each instance individually as data continuously flows in, deciding immediately whether to query for labels. This approach allows for real-time data processing, is memory efficient, and can handle new classes. However, the quality of selection might be lower since it relies on sequentially incoming data.

In semiconductor manufacturing processes, active learning techniques are gaining attention to address issues such as imbalanced data distribution, high labeling costs, and large-scale data processing. These techniques are also being applied to WBM defect pattern analysis research. Table 3 presents the trends in study applying active learning to defect pattern analysis in WBM. Existing studies have applied pool-based active learning to efficiently label training data (Kong & Ni, 2020; Manivannan, 2024; Shim et al., 2020). Since these studies utilize only training datasets in an offline environment, they assume a closed-set data scenario and improve model performance through multiple iterations of training. However, the limitation of these studies is the inability to detect and respond to new defect patterns do not present in the training data. Given the increasing occurrence of new defects due to the miniaturization of semiconductor processes and the introduction of new processes and equipment, this study employs a stream-based active learning approach to update the model regarding these unknown defect patterns. To compensate for the drawbacks of stream-based active learning, this study introduces an online active learning technique using a window-based buffer that can accumulate data over a certain period and amount. This approach effectively processes the buffered data even when unknown defect pattern samples do not occur or only a minimal amount of data exists. It determines whether to update the model, deferring the update if no unknown patterns occur, thereby efficiently labeling detected unknown defect patterns and updating the model to prevent misclassification by the classifier.

3. Proposed method

Existing WBM defect classification methods have mainly adopted deep learning approaches, but they are severely limited by the lack of pre-labeled data and inefficient labeling. Moreover, most methods tend to assume known defect patterns, which can limit their ability to detect unknown defect patterns in actual semiconductor manufacturing processes and significantly reduce the performance of the pre-trained classifier. These problems further limit the practical application of existing research as they require continuous retraining of pre-trained classifiers, as well as additional labeling and data collection for unknown defects. Therefore, this study proposes a comprehensive

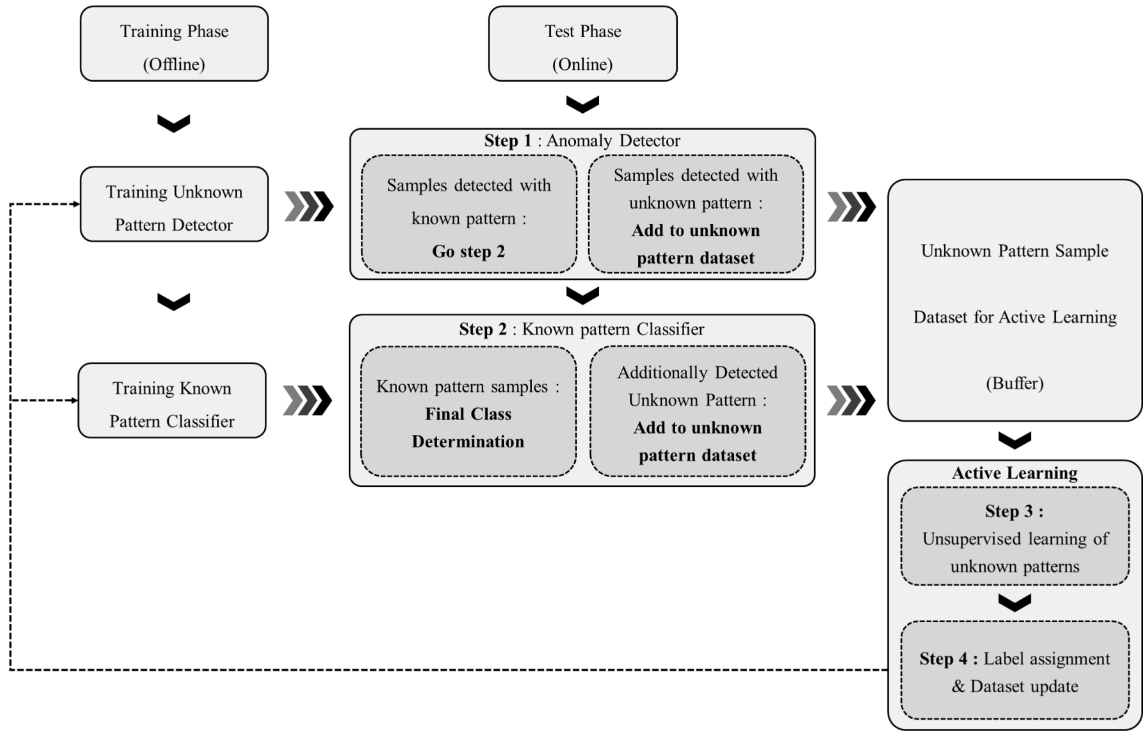

Fig. 1. Framework of the proposed method.

framework, shown in Fig. 1, that can resolve the aforementioned issues related to unknown defect patterns. The proposed framework consists of four steps, as explained in detail below.

3.1. Step 1: detect unknown defect patterns

Step 1 involves using OSR (Open Set Recognition) and anomaly detection approaches to detect previously unknown defect pattern classes using only known WBM pattern data. In this process, the test samples from the actual production environment are subjected to binary classification in an online setting to determine whether they belong to a known pattern type or an unknown new defect pattern type. This approach trains the model offline based solely on known class data, ensuring that when new types of defects emerge in the actual manufacturing environment, they can be effectively detected. If a sample is predicted by the detector to have a known defect pattern, it is considered as a known pattern sample, and the process continues to Step 2. Otherwise, it is considered an unknown pattern sample, and the process proceeds to Step 3. This step only determines whether the sample is as a known or unknown pattern. The type of pattern in the known pattern sample is classified in Step 2. Step 1 ensures that the classifier in Step 2 for known patterns can minimize misclassification, thereby maintaining the performance of the initially trained model.

3.2. Step 2: classification of known pattern samples

A pre-trained classifier trained using supervised learning algorithms is employed to classify defect patterns to build the set of known pattern samples. The classifier is trained using known WBM defect patterns to classify and identify each defect pattern. State-of-the-art supervised learning algorithms such as ResNet, DenseNet, ViT, and EfficientNet are utilized for training and construction of classifiers. Therefore, this step involves selecting the most appropriate algorithm and using it to train the classifier. Additionally, to correct the misclassifications of the anomaly detector, thresholds for each defect patterns are jointly learned. Samples that do not meet the classification thresholds for all defect types are regarded as unknown pattern samples and the process continues to Step 3. In Step 2, samples classified as known patterns in Step 1 are further classified into detailed pattern classes, aiding in the understanding of specific defect causes within the manufacturing process. Steps 1 and 2 complement each other. If only Step 1 exists, it provides information solely on whether samples belong to known patterns or represent unknown defect patterns, making it difficult to trace the detailed causes of defects. Conversely, if only Step 2 exists, there may be an increase in misclassifications by the classifier, potentially degrading overall classification quality. Ultimately, the objectives of Steps 1 and 2 are to effectively detect known and unknown defect patterns using a high-precision supervised learning-based classifier, while accurately identifying detailed defect causes in samples classified as known patterns.

3.3. Step 3: unsupervised learning for unknown pattern samples

An unsupervised learning approach that does not require any pre-trained labels for the unknown pattern samples is used. In this process, the model learns on its own from the unlabeled sample data and discovers new defect patterns. As a result, this step is very helpful in discovering new defect patterns that would have been difficult to find using labeled data alone. Unsupervised learning techniques at this stage can include clustering methods or outlier detection techniques. Clustering helps discover new patterns by grouping data with similar patterns together. Outlier detection can help find new defect patterns by identifying unusual data that deviate from known patterns. DBSCAN, Isolation Forest, KNN, and the Gaussian Mixture Model (GMM) are typical algorithms that can be used for clustering and outlier detection. In Step3, although the stream-based approach in active learning typically involves real-time sample-by-sample judgment in an online environment, the semiconductor manufacturing process involves multiple wafers (up to 25) in a lot and simultaneously generates measurement data from various equipment. In addition, there is a significant imbalance between normal and defect-pattern WBM data, which can limit the ability to add data samples suspected of unknown defect patterns to the

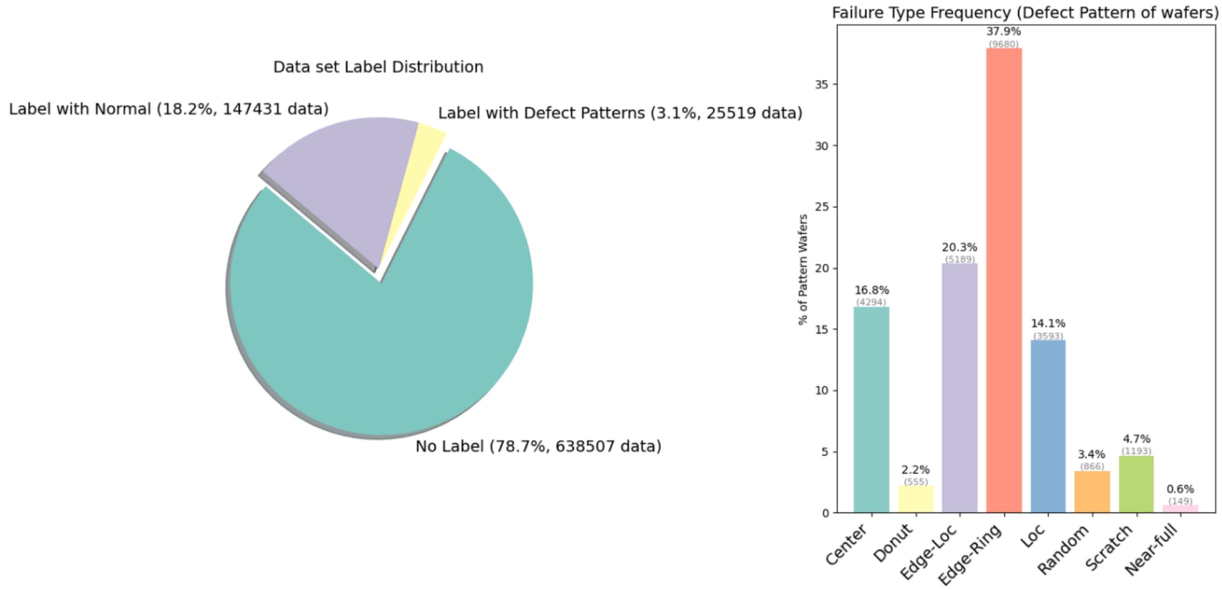

Fig. 2. Label composition of WM-811K data set and distribution of defect pattern types.

dataset. Therefore, one of the techniques used for online active learning is a window-based approach that utilizes the dataset as a buffer. In this approach, data suspected to be unknown defect samples are collected over a certain time period or number of instances, and then unsupervised learning is performed.

3.4. Step 4: model update process based on active learning

In this final step, the anomaly detector of Step 1 and classifier of Step 2 are updated based on the detected unknown defect patterns to improve their performances. To this end, active learning techniques, including Human-in-the-loop (HITL) (Wu et al., 2022), can be employed to continuously improve the model. Active learning technology allows for selective labeling of abnormal samples, enabling rapid recognition of unknown defect patterns and updates to the detector and classifier. HITL, employed in the active learning process, involves human intervention when uncertainty is high, domain knowledge is crucial, and judgment about classification results is challenging. Active learning alone may be sufficient if the unknown defect pattern is similar to an existing known defect pattern or if a pre-trained anomaly detection model can predict the defect type. However, if the defect patterns are entirely different or require additional definition of patterns based on shape, size, and location, the HITL method, relying on engineer domain knowledge, can be effective for updating the models. Through this process, known defect patterns can be classified using a high-precision supervised learning-based classifier. Simultaneously, unknown defect patterns can be detected without the need for pre-training or various assumptions, and the classifier can be updated by labeling the unknown defect patterns using active learning techniques. Ultimately, the proposed framework helps to maintain the high performance of the detector

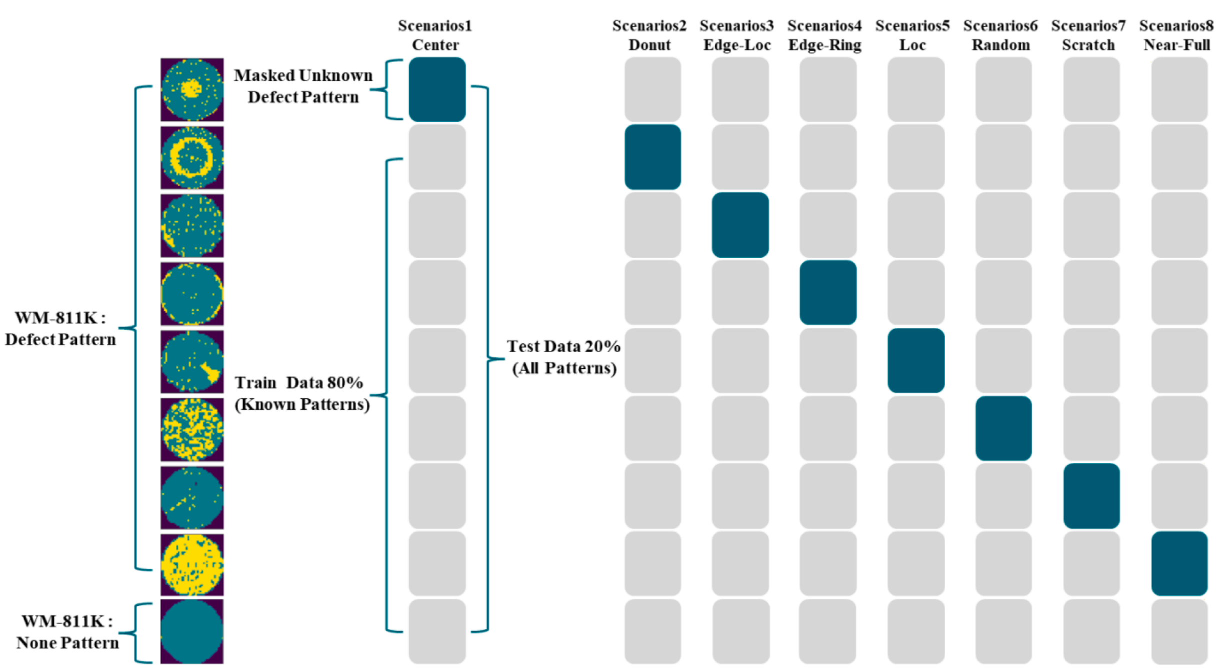

Fig. 3. Creating datasets with unknown defect patterns using the WM-811K dataset.

and classifier and can respond effectively to unknown defect patterns.

4. Case study 1: using WM-811K dataset

In this section, each module used in the proposed framework is described and the practicality and performance of the defect detection method are verified using the open dataset, WM-811K. In Section 4.1, the assumptions and experimental methods for the WM-811K data and unknown defect patterns are introduced. In Sections 4.2–4.5, we discuss the detection performance for such unknown patterns, the clustering results of samples predicted as unknown defect patterns, and the performance of the final classifier updated using the data ultimately labeled through the HITL method. Through this, the performance of the classifier before (Original Classifier) and after (Updated Classifier) using active learning techniques to update the classifier in situations where the test data contains unknown defect patterns is evaluated. Additionally, we compare the performance with a benchmark classifier that was trained on all types of defect patterns in the WM-811K dataset. The experimental results can illustrate how the proposed method performs in a real semiconductor manufacturing environment, especially the efficiency and necessity of the framework to detect unknown defect patterns and update the classifier.

4.1. WM-811K dataset

The WM-811K dataset comprises WBM data measured from semiconductor manufacturing processes and is widely used in numerous studies analyzing WBM defect patterns (Wu et al., 2015). This dataset includes 8 defect patterns (Center, Donut, Edge-Loc, Edge-Ring, Loc, Random, Scratch, Near-Full) and one normal pattern (None). The number of WBM images is 811,457, of which about 21.3 % are labeled, and the remaining 78.7 % are unlabeled as shown in Fig. 2. The data are unevenly distributed across the eight defect patterns. In this study, to simulate the presence of unknown defect patterns, each of the eight defect patterns in the WM-811K dataset was masked as unknown using a "Leave One Out" approach, as shown in Fig. 3. For example, in the first scenario of Fig. 3, Center is marked as unknown while the other patterns are marked as known. This involved excluding defect samples with each pattern from the training and validation datasets while including all defect pattern samples in the test dataset. In this way, a total of 8 scenarios were created, from "Center" to "Near-Full" as shown in Fig. 3. In these scenarios, samples of the "None" pattern, which do not contain any defects, are always assumed to be known patterns and are included in the training dataset in all cases. The effectiveness of the proposed method was then evaluated by examining the average performance changes across these scenarios.

4.2. Step 1: detect unknown defect pattern samples (with One-Class SVM)

In this study, One-Class SVM was selected as an abnormal detector to detect unknown defect patterns. One-Class SVM detects outliers in a dataset by creating a boundary between normal and abnormal data. Widely used for anomaly detection, One-Class SVM has been actively applied in machine learning techniques for WBM defect pattern classification (Baly & Hajj, 2012; Wang & Chen, 2019; Wu et al., 2015). Existing studies have used SVM to determine the presence of defect patterns (i.e., the classification of whether a pattern is normal or defective) in WBM. In this study, all known defects are treated as a single class and it is determined whether a pattern in given sample data belongs to this class, that is, whether it falls within a range of known defect patterns or represents an unknown defect pattern.

To train and evaluate the performance of One-Class SVM, One-Class SVM was trained after excluding each of the eight defect patterns present in the WM-811K dataset as "unknown defect patterns" from the training dataset. The rbf kernel was used as the hyperparameters for training, and a NU parameter value of 0.10 was set using a grid search technique. Afterwards, test data containing excluded defect patterns were used to measure unknown defect detection performance with a One-Class SVM model trained using only previously known types of defects. The results are presented in Table 4. For example, "Center" in Table 4 represents treating the "Center" defect as an unknown pattern, excluding the corresponding defect pattern during training, and showing the classification accuracy for test data containing the Center defect (refer to "Scenarios: Center" in Fig. 3 for details) The results indicate an average classification accuracy of about 85 % for all defect pattern classes.

Table 4
Unknown pattern detection accuracy using One-Class SVM.

Defect pattern	Accuracy (%)
Center	87
Donut	90
Edge-Loc	81
Edge-Ring	72
Loc	83
Random	91
Scratch	88
Near-Full	91
Average (Macro)	**85.4**

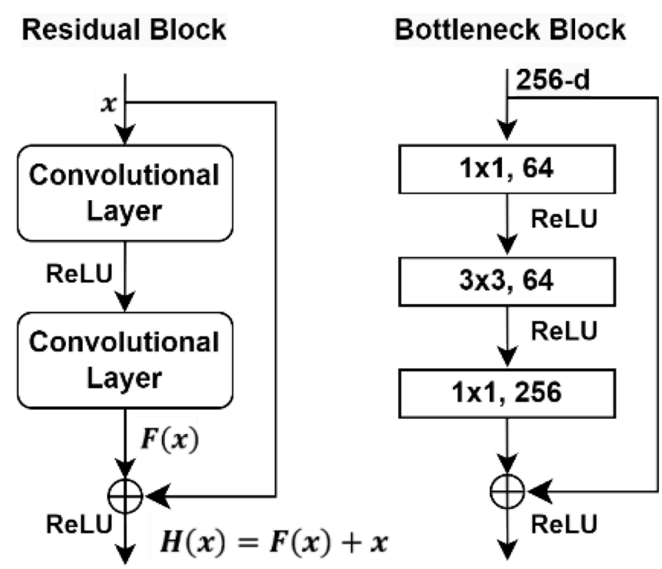

Fig. 4. Left: structure of residual block for ResNet Right: structure of bottleneck block for ResNet50.

4.3. Step 2: classification of known pattern samples (with ResNet50)

ResNet is one of the most popular backbone models used in image classification (He et al., 2015), and various ResNet-based models have been used to classify defect patterns in WBMs (Kim et al., 2021; Xu et al., 2022). Residual learning (skip connections), as shown Fig. 4, was introduced solve gradient vanishing in ResNet. Specifically, ResNet50 further enhances parameter management and prevents overfitting by introducing a bottleneck structure (Fig. 4). As shown in Fig. 4, the bottleneck structure uses a 1 × 1 convolutional layer to reduce the size of the feature map, while utilizing a 3 × 3 convolutional layer to maintain learning ability even in deeper layers. The bottleneck structure can reduce the number of parameters and prevent overfitting at the same time.

In this study, we built a pre-trained classifier using ResNet50 to achieve high classification accuracy for the known pattern samples. In the process of training the pre-trained classifier, severe disparities in the amount of data between defect patterns may cause poor class weight

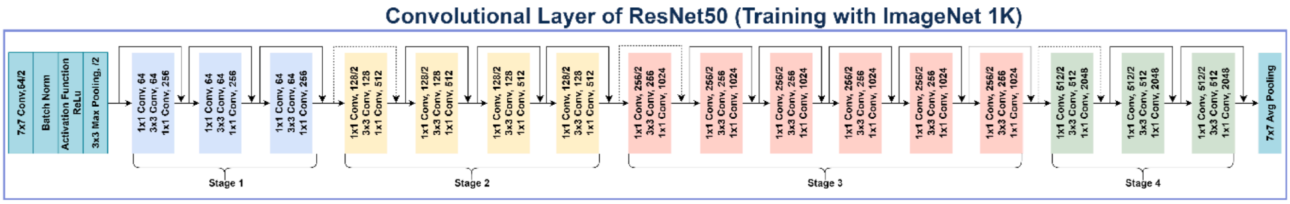

Fig. 5. Convolutional layer of ResNet50 using pre-trained with ImageNet1K.

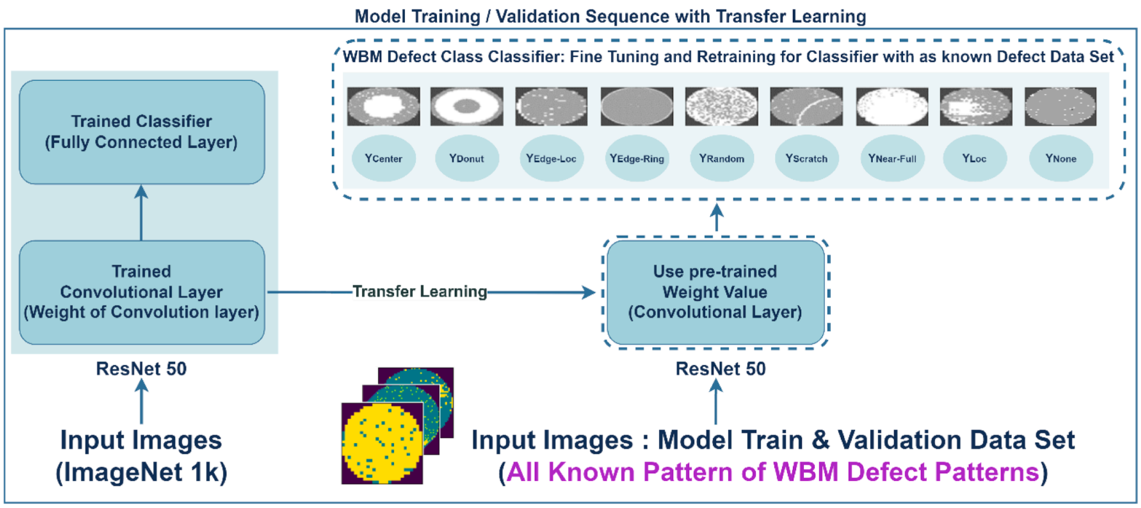

Fig. 6. WBM defect pattern classifier constructed via transfer learning of ResNet50 convolutional layer pre-trained with ImageNet1K.

learning by the classifier (See Fig. 2). To address this problem, this study employed the Transfer Learning technique, using the initial weights of the convolutional layer of the pre-trained ResNet50 on ImageNet1k (refer to the ResNet Convolutional Layer block diagram in Fig. 5). Subsequently, fine-tuning was performed without freezing the weights of the convolutional layers. This allows the model to better adapt to the new WBM image data, which differs from ImageNet 1K, while retaining the general image features. It also enables fine adjustments specific to the WBM defect analysis domain. This approach simultaneously ensures high classification accuracy, adaptability, and efficiency in training time. The process of fine-tuning the model using a training dataset composed solely of known defect patterns to build a classifier for WBM defect analysis can be seen in Fig. 6.

Additionally, we considered the possibility of unknown pattern

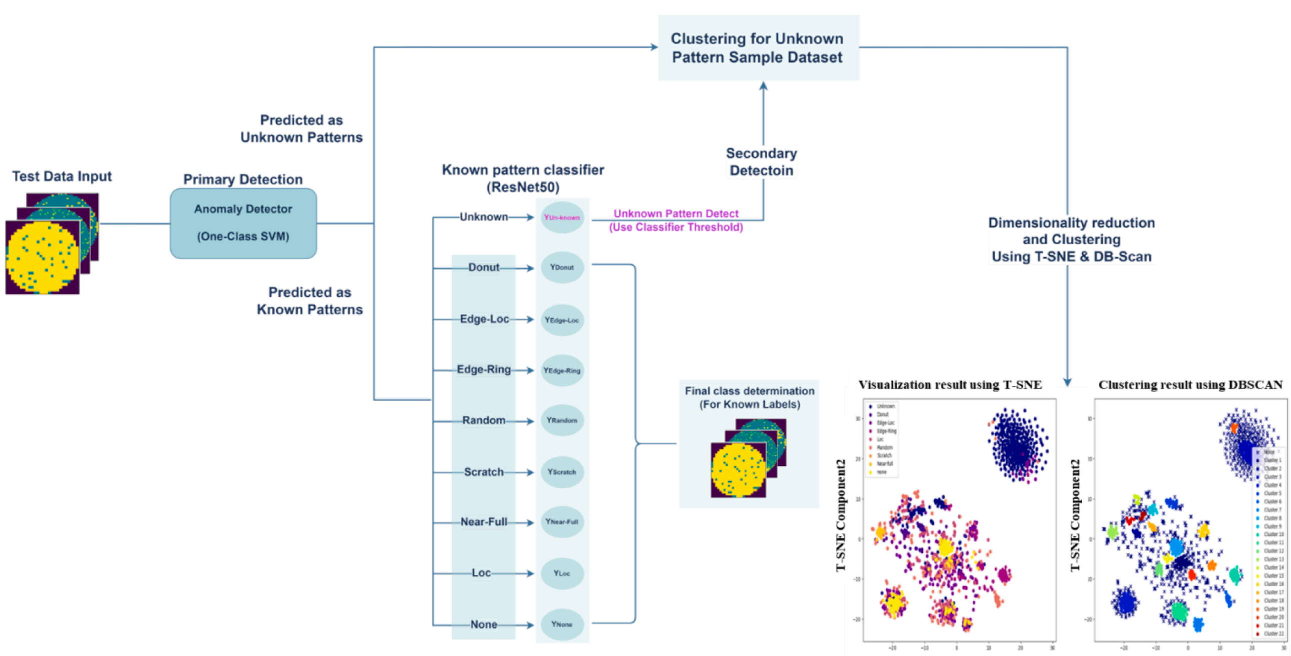

Fig. 7. Multi-step detection process for unknown defect patterns (Center defect pattern is assumed to be an Unknown Class).

Table 5
Performance metrics and confusion matrix index used in the study.

Index Name	Equation
Accuracy	$\frac{TP + TN}{TotalSamples}$
Precision	$\frac{TP}{TP + FP}$
Recall	$\frac{TP}{TP + FN}$
F1 Score	$2 \bullet \frac{Precision \bullet Recall}{Precision + Recall}$
Macro-Accuracy	$\frac{1}{N}\sum_{i=1}^{N} Accuracy_i, N = Total\ number\ of\ classes$
Micro-Accuracy	$\frac{1}{Number\ of\ samples}\sum_{i=1}^{N}(TP_i + TN_i),\ N = Number\ of\ classes$
Cluster Proportion	$\frac{Unknown\ Defect\ Samples\ in\ cluster}{Total\ Number\ of\ sampls\ in\ cluster}$
Performance Gain Ratio	$\frac{Proposed\ Classifier\ Performance - Existing\ Classifier\ Performance}{Existing\ Classifier\ Performance} \bullet 100$

Confusion Matrix Index		Predicted Class	
		Negative (0)	Positive (1)
Actual Class	Negative (0)	TN (True Negative)	FP (False Positive)
	Positive (1)	FN (False Negative)	TP (True Positive)

samples being misclassified as known pattern samples in Step 1. To address this problem, this study introduced an unknown defect pattern class when constructing the multi-class classifier using the ResNet50 model and implemented a multi-stage process for detecting unidentified error patterns using the One-vs-All method. This new class (Unknown Defect Pattern Class) is added to the multi-class classifier, and during the training and validation process of the classifier, the optimal threshold for each known defect pattern class is used to determine whether the sample has an unknown defect pattern The process for determining the optimal threshold for each defect pattern class is detailed in Algorithm 1. In Algorithm 1, we selected precision as the evaluation metric to determine the optimal threshold due to its high relevance in the semiconductor manufacturing environment. In semiconductor manufacturing, the imbalance between normal and abnormal data is significant. If the false positive rate increases, production costs will rise due to the scrapping of wafers and the need for additional precise measurements. Moreover, a higher false positive rate will increase the data labeling costs associated with active learning. Consequently, we chose precision as the performance metric for determining the optimal threshold among various classification performance metrics. This approach aims to effectively manage the unique challenges of semiconductor manufacturing, balancing data operations and labeling costs accordingly. Using the optimal thresholds obtained through this process, test data samples that did not reach the optimal threshold for any class were considered as part of the unknown defect pattern class. Fig. 7 shows the multi-stage detection process for unknown defect patterns on the WM-811K dataset(Scenario 1 in Fig. 3) when a "Center" defect pattern, assumed to be an unknown defect pattern excluded from the classifier's training process, was included in the test data.

In this study, the performance of the classifier was evaluated using indicators such as accuracy, precision, recall, and F1-Score (Table 5). As shown in Fig. 3, it was assumed that each defect pattern in WM-811K is unknown, and thus, training and validation datasets excluding the unknown defect pattern and the test datasets containing the unknown defect pattern were prepared. Using the dataset prepared in this way, the ResNet50 classifier was trained based on transfer learning techniques. Due to the presence of unknown defect patterns in the test dataset, a classifier trained using supervised learning fails to detect them, leading to misclassifications and a subsequent degradation in the performance of pre-trained classifiers. The classification performance is shown in Table 6 (original classifier before applying Step 4). The macro-average accuracy of all classes is approximately 90 % and the average F1-Score is 0.83.

Algorithm 1: Unknown defect pattern detection using the optimal threshold for each class.

 # Function: apply_each_class_optim_threshold
 # Inputs:
 # – Predictions: Model's prediction array $\mathscr{Y}_{pred}$
 # – Actual_labels: Actual class label array $\mathscr{Y}_{true}$
 # – Threshold: Threshold value $\mathscr{E}$
 # – Precision: Precision value $\mathscr{P}$
 # Initialize: $\mathscr{N}$: Numberofknownpatternsclasses , $\mathscr{E}_{range}$: *Thresholdsrangesettingforsearch*
for each known pattern class **do**
  $\mathscr{P}_{best}, \mathscr{E}_{best} \leftarrow 0.0, 0.0$ Set Initial $\mathscr{E}$ Value
  for $\mathscr{E}$ in $\mathscr{E}_{range}$ **do**
    calculate $\mathscr{P}$ with $\mathscr{Y}_{pred}$, $\mathscr{Y}_{true}$, $\mathscr{E}$
    if $\mathscr{P} > \mathscr{P}_{best}$ **then**
      $\mathscr{P}_{best}, \mathscr{E}_{best} \leftarrow \mathscr{P}, \mathscr{E}$
    end if
  end for
    Output: $\mathscr{E}_{optim}[\mathscr{N}] \leftarrow \mathscr{E}_{best}$
end for

Table 6
Classification performance of the original classifier before Step 4.

Unknown patterns	Precision	Recall	F1-Score	Accuracy
Center	0.78	0.80	0.80	0.86
Donut	0.85	0.87	0.86	0.97
Edge-Loc	0.74	0.87	0.80	0.85
Edge-Ring	0.75	0.87	0.79	0.71
Loc	0.74	0.87	0.80	0.89
Random	0.81	0.86	0.83	0.96
Scratch	0.83	0.86	0.85	0.94
Near-Full	0.87	0.87	0.87	0.98
Average (Macro)	**0.80**	**0.86**	**0.83**	**0.90**

Table 7
Cluster proportion measured using each clustering algorithm.

Unknown Defect Cluster Proportion	DBSCAN	KMEANS	Hierarchical Clustering
Center	100 %	100 %	99 %
Donut	100 %	100 %	100 %
Edge-Loc	83 %	31 %	83 %
Edge-Ring	100 %	99 %	100 %
Loc	72 %	99 %	72 %
Random	100 %	22 %	100 %
Scratch	58 %	17 %	33 %
Near-Full	100 %	11 %	52 %
Average (Macro)	89 %	45 %	79 %

4.4. Step 3: unsupervised learning for unknown pattern samples

Generally, as the dimensionality of data increases, the complexity of various computational problems increases due to the so-called 'curse of dimensionality', which is known to make clustering algorithms particularly inefficient. This means that as the number of dimensions in the data increases, it becomes difficult to make meaningful distinctions between similar and dissimilar data (Assent, 2012). WBMs have a shape of 80x80. If clustering is performed on these WBMs without dimensionality reduction, it may be challenging to clearly distinguish between misclassified known pattern samples and unknown pattern samples. Therefore, appropriate dimensionality reduction for WBMs is essential to improve clustering performance.

In this study the unknown pattern samples were subsequently subjected to dimensionality reduction and visualization using t-distributed Stochastic Neighbor Embedding (T-SNE), as shown in Fig. 7. T-SNE is a dimensionality reduction algorithm that maps high-dimensional data to a lower-dimensional space. It maintains the structure and features of the data while transforming similar data points to maintain similar distances in the lower-dimensional space, providing significant advantages for data visualization. The t-distribution has a thicker tail compared to the Gaussian distribution, which preserves distances between data points better with clear cluster boundaries. In semiconductor manufacturing processes, defect data is characterized by significant data imbalance between normal and defective data, as well as among various defect patterns. When unknown defect pattern data occurs very rarely, using t-SNE allows for meaningful statistical analysis even with a small number of data samples. Additionally, the t-distribution can flexibly handle large data deviations. For these reasons, the t-distribution can provide more stable and accurate modeling in semiconductor manufacturing environments where unknown defect pattern samples can appear in various forms. Therefore, in this study, we chose t-SNE to visualize and analyze unknown defect patterns effectively by leveraging its capability to maintain data structure and facilitate clear visualization from high-dimensional datasets.

After mapping the data points to a lower dimension through T-SNE, several clustering algorithms were applied, and the results were compared. The cluster ratio quantifies how well data with the same type of defect is grouped into the same cluster (Table 5). The closer the cluster ratio is to 100 %, the more data with the same defects are clustered in the same cluster. Table 7 shows the performance of each clustering algorithm when samples with unknown patterns among test data are predicted as 'unknown defect patterns' through a multi-step detection process. From the results shown Table 7, defect patterns such as "Center", "Donut", and "Edge-Ring" were effectively clustered regardless of the algorithm used. However, defects with relatively similar characteristics, such as "Edge-Loc", "Loc" and "Scratch" exhibited poorer performance when using the distance-based KMEANS algorithm. In addition, the performance of the hierarchical clustering algorithm was comparable to DBSCAN, but the threshold values need to be specified for each defect pattern. Therefore, DBSCAN was chosen as the final clustering algorithm due to its superior performance and ease of use.

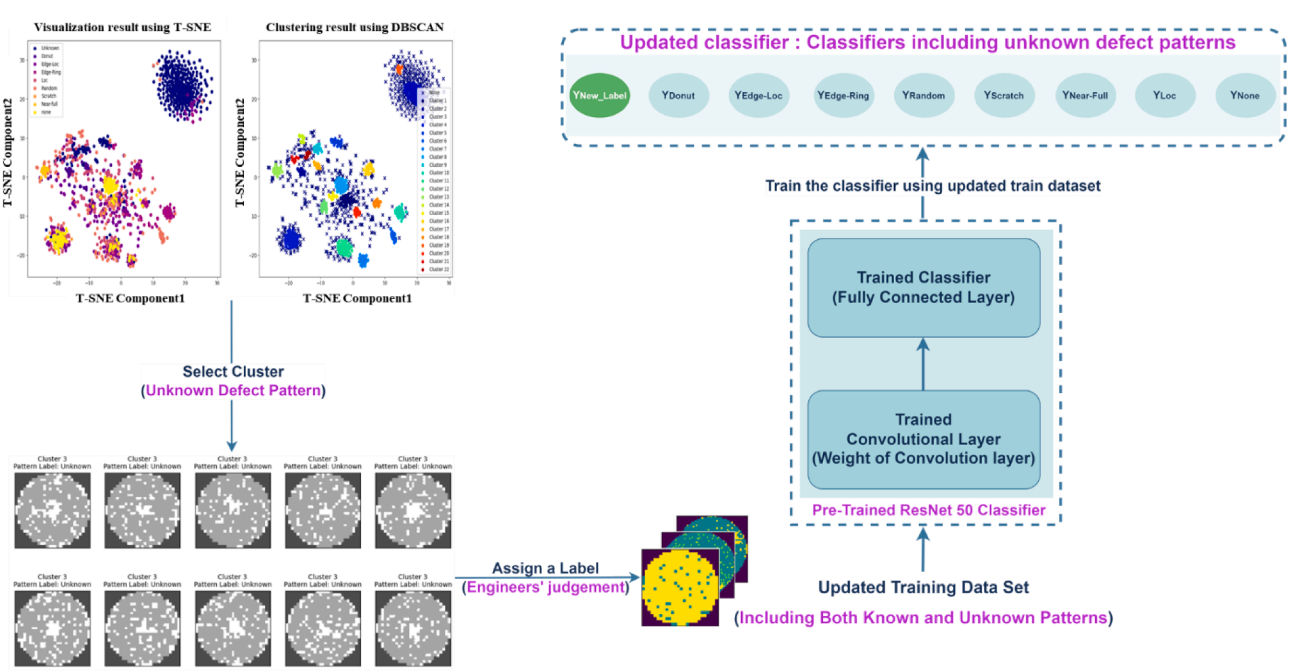

Fig. 8. Process of updating the classifier to an unknown pattern.

Table 8
Classification performance of the updated classifier after applying Step 4.

Unknown patterns	Precision	Recall	F1-Score	Accuracy
Center	0.91	0.92	0.91	0.93
Donut	0.95	0.91	0.97	0.96
Edge-Loc	0.93	0.90	0.91	0.92
Edge-Ring	0.89	0.94	0.91	0.91
Loc	0.89	0.90	0.89	0.92
Random	0.82	0.83	0.80	0.87
Scratch	0.90	0.90	0.90	0.95
Near-Full	0.97	0.94	0.95	0.98
Average (Macro)	0.91	0.91	0.91	0.94
Performance Gain Ratio	+14 %	+6%	+10 %	+4%

Table 9
Performance of the Benchmark classifier trained using all defect classes on WM-811K.

Defect Name	Precision	Recall	F1-Score	Number of Samples
Center	0.92	0.96	0.94	791
Donut	0.92	0.92	0.92	99
Edge-Loc	0.87	0.86	0.86	902
Edge-Ring	0.99	0.98	0.98	1876
Loc	0.80	0.79	0.80	508
Random	0.87	0.93	0.90	127
Scratch	0.82	0.81	0.82	1276
Near-Full	0.93	0.90	0.91	25
None	0.95	0.94	0.95	1950
Average (Macro)	**0.90**	**0.90**	**0.90**	**Total: 7554**

DBSCAN does not require a predefined number of clusters, giving it the flexibility to find clusters based on the distribution of the data. This characteristic of DBSCAN is particularly useful for identifying non-linear clusters and is well suited for semiconductor manufacturing, where the number of defects is not known in advance and the risk of misclassification and noise exposure is high. An average cluster promotion performance of 89 % was achieved using DBSCAN, as shown in Table 7.

4.5. Step 4: Model update process based on active learning

The proposed methodology involves using active learning to update the existing classifier to effectively train unknown defect patterns. In Steps 1 and 2, unknown defect patterns are detected based on their differentiation from known WBM patterns. In Step 3, density-based clustering is used to group unknown data with similar features, making it easier for engineers to consistently review and label the data. This reduces noise and the amount of data that needs to be reviewed, thereby facilitating the HITL (Human-in-the-Loop) process in Step 4 and improving the overall algorithm performance. The final decision on whether an "unknown pattern" exists is made by HITL, performed by a domain knowledgeable semiconductor process engineer. Fig. 8 shows the entire multi-step detection process in a situation where the 'Center' defect pattern among the defect types in the WM-811K data is assumed to be an unknown defect type. WBMs that are ultimately identified as defects of unknown type are assigned a new label based on the engineer's judgment. Through this multi-step process, the anomaly detector and classifier were updated to detect unknown defect classes, allowing the corresponding defect patterns to be accurately identified and predicted in Steps 1 and 2.

Table 8 shows the performance of the updated classifier after the model is updated in Step 4. Each row represents the average classification performance of the model on test data where all types of defects coexist, assuming the corresponding defect in the WM-811K data as an "unknown pattern" and removing it from the classifier's training process. As shown in Table 8 the Performance Gain Ratio, represents the performance improvement in percent of the updated classifier after Step 4 compared to the original classifier (Performance in Table 6) prior to being updated. Additionally, Fig. 9 compares the average classification performance on the entire dataset between the original classifier before applying the proposed methodology and the updated classifier after its application. These results reveal that without Step 4, the model performance is significant drop because "unknown patterns" cannot be verified by the detector and classifier (Classification performance index "0" for each unknown defect pattern). Without Step 4, the classifier

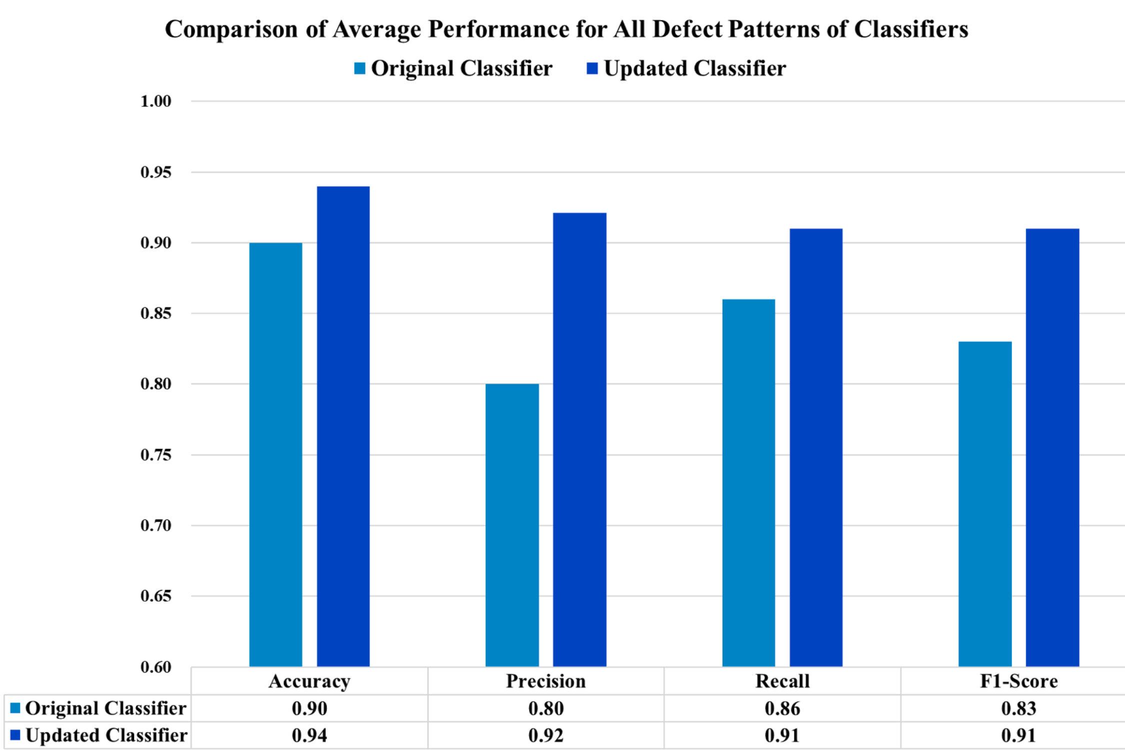

Fig. 9. Comparison of classifier's average performance for all defect patterns before and after using the proposed methodology.

Table 10
Performance Comparison with Related Studies: (a) Macro Average Accuracy, (b) Micro Average Accuracy.

(a) Author	Macro Accuracy	(b) Author	Micro Accuracy
Nakazawa and Kulkarni (2018)	0.94	Chen et al. (2022)	0.98
Yu et al. (2019)	0.93	Wu et al. (2015)	0.95
Xu et al. (2022)	0.95	Yu and Liu (2021)	0.97
Shinde et al. (2022)	0.96	Updated Classifier	0.98
Yu et al. (2021)	0.96		
Wu et al. (2015)	0.83		
Yu and Liu (2021)	0.94		
Kong and Ni (2020)	0.90		
Updated Classifier	0.94		

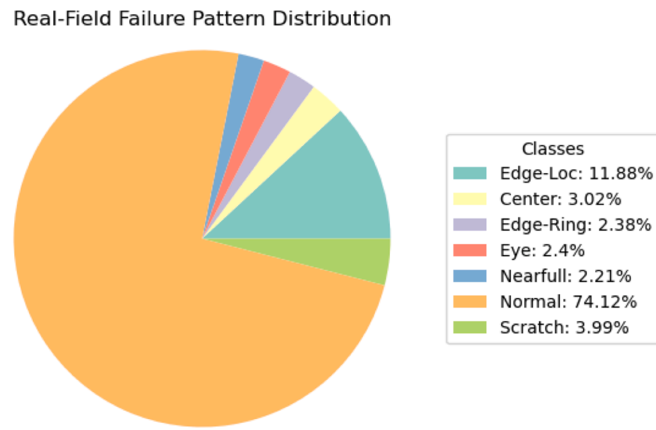

Fig. 11. Distribution of defect patterns from real field data.

showed a decrease in precision, recall, F1-score, and accuracy by 14 %, 6 %, 10 %, and 4 %, respectively, compared to the performance of the model with Step 4 implemented. Specifically, for the "Edge-Ring" defect, all performance metrics decreased by more than 20 %.

To further validate the superior performance of the proposed method, the performance of the "Benchmark Classifier", trained using ResNet50 without removing all defect types from the WM-811K dataset (Table 9) was compared with the performance of the proposed method after Step 4. Increasing the amount of training data generally improves classifier performance. However, label noise can particularly impair performance when class boundaries are unclear (Krawczyk, 2016). WBM patterns often lack clear boundaries, causing confusion. For instance, defect patterns like "Edge-Loc" and "Loc" may have similar shapes and sizes but differ in location. Similarly, a "Scratch" defect could be mistaken for "Loc" or "Edge-Loc". Research by Kim et al. (2021), suggests that selecting consistent and clear images for training enhances classifier performance more effectively than using all labeled WBM data. The proposed methodology demonstrates improved performance when learning new defect classes that resemble known patterns such as "Loc", "Edge-Loc", or "Scratch". This improvement stems from selecting training data that distinctly differs from existing defect patterns. Conversely, for patterns like "Center", "Edge-Ring", and "Random" which exhibit clear distinctions from known defects, benchmark classifiers trained on larger datasets generally perform better.

Table 10 compares the classification performance of studies using the WM-811K dataset mentioned in Tables 2 and 3 of Section 2 with the performance of the "updated classifier" applying the proposed methodology. The classification accuracy of existing studies is presented in terms of Micro-Average and Macro-Average, and the performance of the proposed updated classifier is compared using Macro average accuracy in Table 10(a) and Micro average accuracy in Table 10(b). As shown in Fig. 2, 85 % of the WM-811K data consists of "None" pattern data, resulting in higher Micro average accuracy compared to Macro average accuracy. This is because the classification performance of the "None" pattern class significantly influences the Micro average accuracy. A similar phenomenon was observed in Wu et al. (2015), demonstrating the impact of data imbalance on performance metrics. The "updated classifier" showed performance similar to Chen et al. (2022) in Micro average accuracy in Table 10(b), achieving the best performance. On the other hand, Macro average accuracy shown in Table 10(a) surpassed the average performance of other models, although it did not record the highest performance. The updated classifier differs from existing

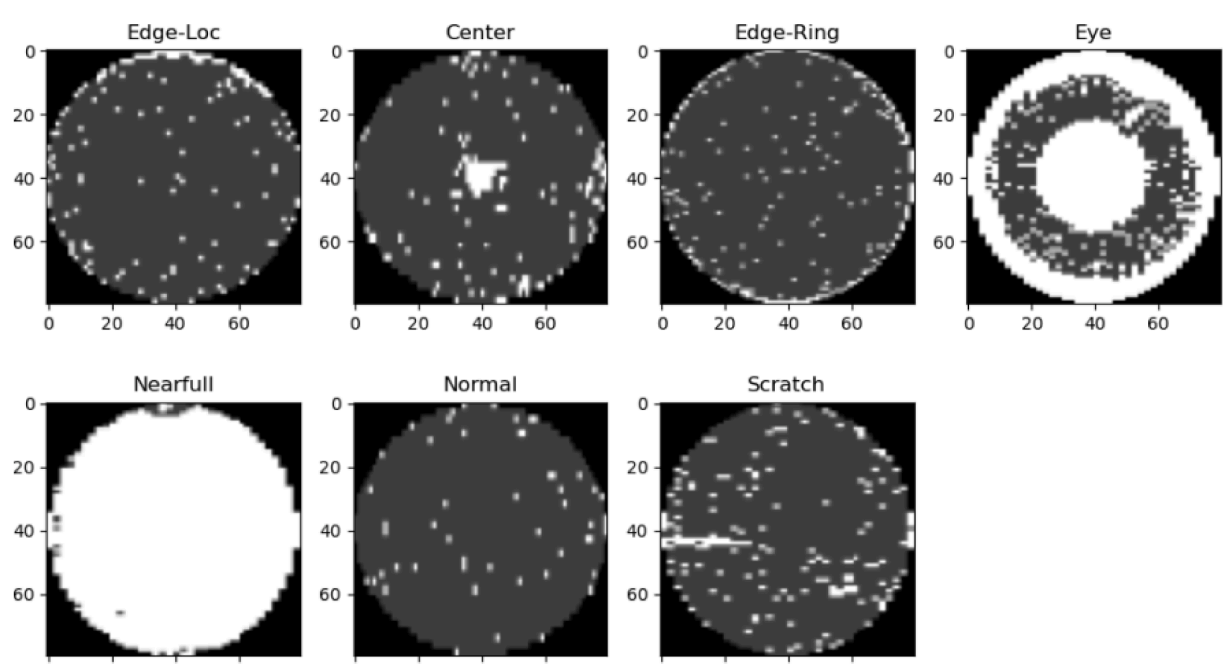

Fig. 10. Defect classification after visualization using real field data.

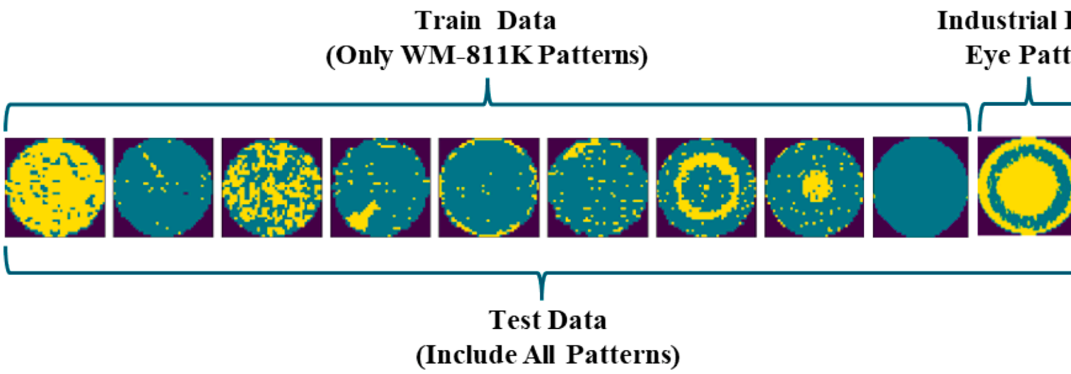

Fig. 12. Composition of real field industrial dataset.

Table 11
Classification performance of the existing classifier.

Defect Name	Precision	Recall	F1-Score	Number of Samples
Center	**0.91**	**0.98**	**0.91**	891
Donut	0.85	0.95	0.90	122
Edge-Loc	0.76	0.90	0.82	1081
Edge-Ring	**0.82**	**0.98**	**0.90**	1860
Loc	0.79	0.81	0.80	688
Random	**0.69**	**0.86**	**0.76**	174
Scratch	0.91	0.72	0.80	225
Near-Full	**0.19**	**0.97**	**0.32**	34
Eye (Real Industrial Data)	**0.00**	**0.00**	**0.00**	715
None	0.96	0.93	0.95	2029
Average (Macro)	**0.69**	**0.81**	**0.72**	**Accuracy: 84 %**

Table 12
Classification performance of the proposed classifier.

Defect Name	Precision	Recall	F1-Score	Number of Samples
Center	0.95	0.96	0.96	891
Donut	0.95	0.85	0.90	122
Edge-Loc	0.87	0.87	0.87	1081
Edge-Ring	0.95	0.98	0.96	1860
Loc	0.78	0.85	0.81	688
Random	0.90	0.83	0.86	174
Scratch	0.77	0.84	0.80	225
Near-Full	0.62	0.88	0.73	34
Eye (Real Industrial Data)	1.00	0.88	0.94	715
None	0.96	0.94	0.95	2029
Average (Macro)	0.88	0.89	0.88	Accuracy: 92 %
Performance Gain Ratio	+28 %	+10 %	+22 %	+10 %

studies' classifiers that train all defect pattern classes through supervised learning, as it detects previously untrained unknown defect patterns and then updates the classifier using active learning techniques. Nevertheless, the performance results in Table 10 demonstrate that the proposed methodology effectively detects unknown defect pattern data, and the updated classifier exhibits excellent performance. It shows that there is no significant performance degradation compared to classifiers trained through supervised learning with all defect pattern classes predefined. Moreover, while the updated classifier presented here uses the ResNet50 model, if various State of The Art deep learning networks are used as the backbone model when configuring the known defect pattern classifier in step 2 of the proposed methodology, the classification accuracy could be further improved.

5. Case study 2: using real field data

In this section, field data obtained from the semiconductor manufacturing industry was used to evaluate the applicability of the proposed method on the manufacturing site. The dataset involved selecting wafer bin numbers in the region of interest from electrical die alignment results obtained from an actual semiconductor wafer fab and preprocessing these into images representing WBM defect patterns. After that, the entire WBM is labeled by process engineers using domain knowledge. To validate the utility of the proposed framework on real industrial data, a classifier ("Existing Classifier") was pre-trained using only the WM-811K dataset, and then the performance of the existing classifier was evaluated using a test dataset containing the 'Eye' defect pattern, an unknown defect pattern in real field data. The existing classifier was trained using the entire pattern set of WM-811K as training data, following the same approach as the benchmark classifier introduced in Section 4.5. Additionally, the existing classifier utilized a ResNet50 model that was fine-tuned through transfer learning with the ImageNet 1 K dataset, applying the same parameters as the benchmark classifier. Then, performances of the existing classifier and the classifier of the proposed method ("Proposed Classifier"), the effectiveness of the proposed method in detecting unknown defect patterns in actual semiconductor manufacturing data were evaluated and compared.

5.1. Real field industrial data

The defect patterns from actual semiconductor manufacturing data are visualized in Fig. 10. There are six defect patterns: 'Edge-Loc', 'Center', 'Edge-Ring', 'Eye', 'Near-full' and 'Scratch'. The 'Eye' defect pattern is an unknown single defect pattern that does not exist in the WM-811K data set and can also be seen as a mixed defect pattern that occurs by combining 'Center' defects and 'Edge-Ring' defects. Fig. 11 indicates the data distribution for each defect pattern, showing similarities in the distribution ratios between 'Normal' and 'Defect' patterns compared to WM-811K. Similar to Case Study 1, Case Study 2 utilized real field industrial data to evaluate the methodology by adding a new defect pattern called "Eye". This defect pattern does not exist in the WM-811K dataset. Fig. 12 illustrates the composition of the dataset in detail. The real field industrial dataset used the 9 patterns from WM-811K as known patterns in the training data. Subsequently, a new defect type, "Eye", which occurs in real industrial environments, was added only to the test dataset, resulting in test data containing a total of 10 patterns. This approach enabled the evaluation of the proposed methodology's ability to detect unknown defect patterns and the performance of the updated classifier.

5.2. Performance for real field industrial data

The performance metrics of the existing classifier are summarized in Table 11. Since the "Eye" defect cannot be identified by the classifier trained only on the WM-811K dataset, all classification performance

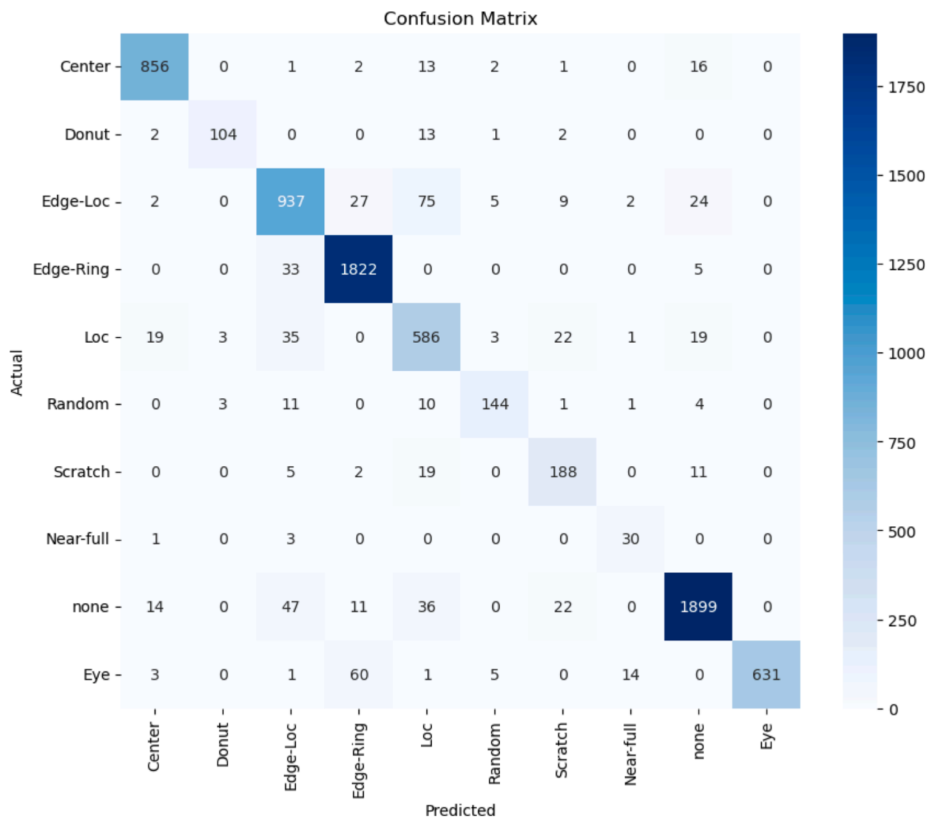

Fig. 13. Confusion matrix of real field industrial test data.

Table 13
Detection performance of the proposed model.

Model Performance With real industrial dataset		
Unsupervised learning Performance	Unknown Pattern Detect Ratio (Recall)	0.91
	Anomaly Detection Accuracy	0.89
	Unknown Pattern Cluster Proportion	1.00
Updated Model Performance (With active learning)	Updated Model Classification Accuracy	0.92
	Unknown Pattern (Eye) Classification Precision	1.00
	Unknown Pattern (Eye) Classification Recall	0.88
	Unknown Pattern (Eye) Classification F1-Score	0.94
	Updated Model Precision	0.88
	Updated Model Recall	0.89
	Updated Model F1 Score	0.88

metrics for this defect type are 0 %. Additionally, it was found that the detection performance for the "Center" and "Edge-Ring" defect patterns, which are similar to the "Eye" defect pattern, is significantly reduced due to misclassification. Similarly, the detection performance for the "Near-full" and "Random" defect patterns is also greatly deteriorated by misclassifications caused by the "Eye" defect. As a result, the existing classifier was performed significantly worse than the "Benchmark Classifier" (a classifier trained on all defect patterns in the WM-811K dataset using supervised learning as presented in Table 9). Defect patterns that cannot be classified using existing classifiers ("Eye") or those in which the classification performance significantly deteriorated are marked in bold in Table 11. Compared to the benchmark classifier, the performance of the existing classifier decreased by 23 %, 10 %, 20 %, and 10 % in average precision, recall, F1-score, and accuracy, respectively.

Table 12 shows the classification performance of the proposed classifier. The proposed classifier successfully detected the "Eye" defect with excellent performance metrics of 1.00, 0.88, and 0.94 in precision, recall, and F1-score, respectively. Additionally, it was confirmed that the classification performance for "Center" and "Edge-Ring" defects, which are misclassified by the existing classifier, is significantly improved due to the existence of the "Eye" defect pattern. The detection performance of defect patterns such as "Random" and "Near-full" in Table 12, which were greatly reduced by the existing classifier, was also greatly improved by the proposed classifier. Fig. 13 shows the confusion matrix of the test data.

As a result, the proposed classifier matched up to 99 % of the performance of the benchmark classifier, much better than the existing classifier. The overall performance of the proposed model measured using test data from actual field data can be seen in Table 13. Using One-Class SVM, the identification accuracy for unknown defects is 89 %, the recall rate for unknown defects is 91 %. The F1-Score, a measure of detection performance for the "Eye" defect pattern using the proposed classifier, is 0.94. These results indicate that the proposed method has excellent classification performance for unknown defect patterns. The results also imply that the performance of the pre-trained classifier is not deteriorated by the proposed method, even when unknown defect patterns occur in actual semiconductor manufacturing environments. On the other hand, if a defect with an unknown pattern occurs, the classification performance of the pre-trained classifier can significantly deteriorate, as shown in Table 9. A more critical issue is that this degradation in performance is not immediately obvious. Therefore, in order to maintain the acceptable performance of the pre-trained classifier, the classifier should be periodically retrained, data for unknown defect types should be continuously obtained, and labels must be accurately assigned. These measures cost a lot and are thus difficult to implement and expose the limitations of applying existing research to actual field data. Therefore, in this study, we proposed a method that can recognize newly occurring unknown defect types, including a comprehensive framework for updating the classifier with unknown defect patterns. The results prove that the proposed method can accurately detect unknown defects in an actual industrial setting and minimize the decline in performance of the classifier when new unknown defects are present.

6. Conclusions

This study proposes a framework for accurately and efficiently detecting unknown defect patterns that have been overlooked in previous WBM defect pattern classification studies. Previous studies have not considered the effect of classifier performance degradation or the emergence of unknown defect patterns, thus making continuous classifier retraining and data labeling for unknown defect patterns cost a lot and difficult in actual semiconductor manufacturing sites. In this study, unknown defect patterns were effectively detected while good classification performance for known defect patterns was maintained. Additionally, by utilizing active learning techniques to continuously update the classifier and detector with detected unknown defects, an innovative strategy for wafer bin map defect analysis was introduced. This allows semiconductor process engineers to react quickly when unknown fault patterns arise while reducing the cost and complexity associated with constant classifier retraining.

To evaluate the performance of our proposed defect detection method, firstly, it was assumed the defect patterns in the WM-811K dataset as unknown defect patterns and trained the pre-trained classifier. The performances of the classifier before and after applying active learning to update the model (Step 4) were compared. When Step 4 was not implemented, the classifier failed to detect defects with unknown defect patterns, showing significant performance degradation of at least 4 % and up to 14 % in all performance indicators. However, after implementing Step 4, the classifier exhibited a significant performance improvement compared to the existing classifier. This was further validated by comparing the results to that using a benchmark classifier trained using supervised learning techniques on all defect pattern types using the WM-811K dataset. In addition, the performance of the classifier was evaluated using real semiconductor manufacturing data containing 'Eye' defect patterns. When compared to the existing classifier, the performance evaluation indicators of the proposed classifier were significantly improved by 10 % to 28 %. These results confirm that the proposed framework is very efficient. Applying the proposed framework can immediately produce significant cost savings and enhanced quality control in actual semiconductor manufacturing environments.

Building upon the results of this study, we aim to enhance the detection rate and labeling efficiency of unknown defect patterns. We intend to explore various methodologies such as Open-Set Recognition and Zero-shot learning to achieve more efficient and high-performance detection of unknown defect patterns using only known patterns. Additionally, leveraging the layers of deep networks trained with known defect patterns, techniques like uncertainty estimation and clustering will be introduced to improve the efficiency of active learning through advanced clustering for data selection. These approaches are expected to enhance the overall performance of clustering through data selection, utilizing diverse model-based approaches during the process of selecting data, based on known defect patterns.

CRediT authorship contribution statement

Jin-Su Shin: Conceptualization, Methodology, Software, Writing – original draft. **Min-Joo Kim:** Validation. **Dong-Hee Lee:** Writing – review & editing, Supervision, Project administration.

Declaration of competing interest

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

Data availability

Data will be made available on request.

Acknowledgements

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

References

Assent, I. (2012). Clustering high dimensional data. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery, 2*(4), 340–350. https://doi.org/10.1002/widm.1062

Bae, Y., & Kang, S. (2023). Supervised contrastive learning for wafer map pattern classification. *Engineering Applications of Artificial Intelligence, 126*. https://doi.org/10.1016/j.engappai.2023.107154

Baly, R., & Hajj, H. (2012). Wafer classification using support vector machines. *IEEE Transactions on Semiconductor Manufacturing, 25*(3), 373–383. https://doi.org/10.1109/TSM.2012.2196058

Batool, U., Shapiai, M. I., Tahir, M., Ismail, Z. H., Zakaria, N. J., & Elfakharany, A. (2021). A systematic review of deep learning for silicon wafer defect recognition. *IEEE Access, 9*, 116572–116593. https://doi.org/10.1109/ACCESS.2021.3106171

Cacciarelli, D., & Kulahci, M. (2024). Active learning for data streams: A survey. *Machine Learning, 113*(1), 185–239. https://doi.org/10.1007/s10994-023-06454-2

Cha, J., & Jeong, J. (2022). Improved U-Net with residual attention block for mixed-defect wafer maps. *Applied Sciences (Switzerland), 12*(4). https://doi.org/10.3390/app12042209.

Chen, S., Zhang, Y., Hou, X., Shang, Y., & Yang, P. (2022). Wafer map failure pattern recognition based on deep convolutional neural network. *Expert Systems with Applications, 209*. https://doi.org/10.1016/j.eswa.2022.118254

Hansen, C. K., & Thyregodb, P. (1998). Use of wafer maps in integrated circuit manufacturing. In *Microelectronics Reliability Pergamon microelectronics Reliability* (Vol. 38).

Hansen, M. H., Nair, V. N., & Friedman, D. J. (1997). *Monitoring wafer map data from integrated circuit fabrication processes for spatially clustered defects* (Vol. 39, Issue 3).

He, K., Zhang, X., Ren, S., & Sun, J. (2015). *Deep residual learning for image recognition.* http://image-net.org/challenges/LSVRC/2015/.

Hsu, C. Y., Chen, W. J., & Chien, J. C. (2020). Similarity matching of wafer bin maps for manufacturing intelligence to empower Industry 3.5 for semiconductor manufacturing. *Computers and Industrial Engineering, 142*. https://doi.org/10.1016/j.cie.2020.106358.

Kang, H., & Kang, S. (2023). Semi-supervised rotation-invariant representation learning for wafer map pattern analysis. *Engineering Applications of Artificial Intelligence, 120*. https://doi.org/10.1016/j.engappai.2023.105864

Kim, E. S., Choi, S. H., Lee, D. H., Kim, K. J., Bae, Y. M., & Oh, Y. C. (2021). An oversampling method for wafer map defect pattern classification considering small and imbalanced data. *Computers and Industrial Engineering, 162*. https://doi.org/10.1016/j.cie.2021.107767

Kim, M., Tak, J., & Shin, J. (2023). A deep learning model for wafer defect map classification: Perspective on classification performance and computational volume. *Physica Status Solidi (B) Basic Research*. https://doi.org/10.1002/pssb.202300113.

Kim, T., & Behdinan, K. (2023). Advances in machine learning and deep learning applications towards wafer map defect recognition and classification: a review. In *Journal of Intelligent Manufacturing* (Vol. 34, Issue 8, pp. 3215–3247). Springer. https://doi.org/10.1007/s10845-022-01994-1.

Kim, Y., Cho, D., & Lee, J.-H. (2020). Wafer map classifier using deep learning for detecting out-of-distribution failure patterns. *2020 IEEE International Symposium on the Physical and Failure Analysis of Integrated Circuits (IPFA), Singapore, 2020, pp. 1-5*. https://doi.org/ 10.1109/IPFA49335.2020.9260877.

Kong, Y., & Ni, D. (2020). A semi-supervised and incremental modeling framework for wafer map classification. *IEEE Transactions on Semiconductor Manufacturing, 33*(1), 62–71. https://doi.org/10.1109/TSM.2020.2964581

Krawczyk, B. (2016). Learning from imbalanced data: open challenges and future directions. In *Progress in Artificial Intelligence* (Vol. 5, Issue 4, pp. 221–232). Springer Verlag. https://doi.org/10.1007/s13748-016-0094-0.

Lee, H., & Kim, H. (2020). Semi-supervised multi-label learning for classification of wafer bin maps with mixed-type defect patterns. *IEEE Transactions on Semiconductor Manufacturing, 33*(4), 653–662. https://doi.org/10.1109/TSM.2020.3027431

Lee, J. H., Moon, I. C., & Oh, R. (2021). Similarity search on wafer bin map through nonparametric and hierarchical clustering. *IEEE Transactions on Semiconductor Manufacturing*. https://doi.org/10.1109/TSM.2021.3102679

Manivannan, S. (2024). Pseudo-labeling and clustering-based active learning for imbalanced classification of wafer bin map defects. *Signal, Image and Video Processing, 18*(3), 2391–2401. https://doi.org/10.1007/s11760-023-02915-2

Nag, S., Makwana, D. R. S. C. T., Mittal, S., & Mohan, C. K. (2022). WaferSegClassNet – A light-weight network for classification and segmentation of semiconductor wafer defects. *Computers in Industry, 142*. https://doi.org/10.1016/j.compind.2022.103720

Nakazawa, T., & Kulkarni, D. V. (2018). Wafer map defect pattern classification and image retrieval using convolutional neural network. *IEEE Transactions on Semiconductor Manufacturing, 31*(2), 309–314. https://doi.org/10.1109/TSM.2018.2795466

Shim, J., Kang, S., & Cho, S. (2020). Active learning of convolutional neural network for cost-effective wafer map pattern classification. *IEEE Transactions on Semiconductor Manufacturing, 33*(2), 258–266. https://doi.org/10.1109/TSM.2020.2974867

Shinde, P. P., Pai, P. P., & Adiga, S. P. (2022). Wafer defect localization and classification using deep learning techniques. *IEEE Access, 10*, 39969–39974. https://doi.org/10.1109/ACCESS.2022.3166512

Tello, G., Al-Jarrah, O. Y., Yoo, P. D., Al-Hammadi, Y., Muhaidat, S., & Lee, U. (2018). Deep-structured machine learning model for the recognition of mixed-defect patterns in semiconductor fabrication Processes. *IEEE Transactions on Semiconductor Manufacturing, 31*(2), 315–322. https://doi.org/10.1109/TSM.2018.2825482

Wang, R., & Chen, N. (2019). Wafer map defect pattern recognition using rotation-invariant features. *IEEE Transactions on Semiconductor Manufacturing, 32*(4), 596–604. https://doi.org/10.1109/TSM.2019.2944181

Wang, R., & Wang, S. (2023). Similarity searching for fault diagnosis of defect patterns in wafer bin maps. *Computers and Industrial Engineering, 185*. https://doi.org/10.1016/j.cie.2023.109679

Wu, M. J., Jang, J. S. R., & Chen, J. L. (2015). Wafer map failure pattern recognition and similarity ranking for large-scale data sets. *IEEE Transactions on Semiconductor Manufacturing, 28*(1), 1–12. https://doi.org/10.1109/TSM.2014.2364237

Wu, X., Xiao, L., Sun, Y., Zhang, J., Ma, T., & He, L. (2022). A survey of human-in-the-loop for machine learning. In *Future Generation Computer Systems* (Vol. 135, pp. 364–381). Elsevier B.V. https://doi.org/10.1016/j.future.2022.05.014.

Xu, Q., Yu, N., & Essaf, F. (2022). Improved wafer map inspection using attention mechanism and cosine normalization. *Machines*, *10*(2). https://doi.org/10.3390/machines10020146.

Yu, J., & Liu, J. (2021). Two-dimensional principal component analysis-based convolutional autoencoder for wafer map defect detection. *IEEE Transactions on Industrial Electronics, 68*(9), 8789–8797. https://doi.org/10.1109/TIE.2020.3013492

Yu, J., Shen, Z., & Wang, S. (2021). Wafer map defect recognition based on deep transfer learning-based densely connected convolutional network and deep forest. *Engineering Applications of Artificial Intelligence, 105*. https://doi.org/10.1016/j.engappai.2021.104387

Yu, N., Xu, Q., & Wang, H. (2019). Wafer defect pattern recognition and analysis based on convolutional neural network. *IEEE Transactions on Semiconductor Manufacturing, 32*(4), 566–573. https://doi.org/10.1109/TSM.2019.2937793

Zhu, J., Liu, J., Xu, T., Yuan, S., Zhang, Z., Jiang, H., Gu, H., Zhou, R., & Liu, S. (2022). Optical wafer defect inspection at the 10 nm technology node and beyond. In *International Journal of Extreme Manufacturing* (Vol. 4, Issue 3). Institute of Physics. https://doi.org/10.1088/2631-7990/ac64d7.