Similarity-based Local Feature Extraction for Wafer Bin Map Pattern Recognition

Jieun Kim
Department of Industrial and Management Engineering
Korea University
Seoul, South Korea
techzt@korea.ac.kr

Jun-Geol Baek*
Department of Industrial and Management Engineering
Korea University
Seoul, South Korea
jungeol@korea.ac.kr

***Abstract*— A wafer bin map consists of a local chip containing key information and a global chip present in all patterns. The defect pattern shows a specific pattern shape on the wafer bin map and is defined based on the existing area information. Global information is not differentiated from local information in classification problems and is recognized as a major characteristic, so it affects the identification of the characteristics of defective patterns. In preparation for this, a method of extracting key local information has been proposed. In this paper, we propose a Skip Connections Denoising Autoencoder-based methodology to extract regional information of defect patterns. Randomly distributed chips are recognized as noise by defining anomaly scores based on the probability of each chip appearing in the wafer bin map. We propose a data transformation and reconstruction methodology for extracting local information based on the anomaly score, which is an uncertainty score index. Through the proposed methodology, it was confirmed that the main information that could not be extracted from the convolutional neural network (CNN) was extracted, and it was confirmed that the method proposed in this paper for WM-811K data is superior to the existing method.**

Keywords*— Semiconductor manufacturing process, Defect pattern recognition, Data augmentation, Anomaly localization, Skip connections denoising autoencoder

I. INTRODUCTION

A wafer produced through a thin substrate for making an integrated circuit. Thousands of integrated circuit (IC) chips are obtained from wafer, and actual wafers are produced through several steps such as etching and surface polishing [1].

The processed wafer goes through the Electrical Die Sorting (EDS) process, which verifies that the chip reaches desired quality level through electrical property inspection. In semiconductor manufacturing process, chips consist of binary value that indicate each die are classified as defective or normal. When defect dies are concentrated in a specific area and show a certain pattern occur, the label of the defect pattern is defined [2].

Fig. 1 shows eight WBM defect patterns and one normal pattern defined in the last stage of the manufacturing process. Among the defective patterns, the center is caused by surface polishing, and the edge-loc is caused by misalignment between layers [3]. If the defect pattern is accurately classified, the frequency of occurrence of defects can be reduced by identifying the causal factors that contributed to the occurrence of defects during the process.

To identify the cause of WBM patterns, deep learning based method for extracting features are being attempted. It learns common information and classifies labels based on learned image characteristics. Since the existing deep learning-based defect pattern classification method uses all chip information of WBM, it has a limitation in learning common chip information appearing in multiple defect patterns as main information of each defect patterns.

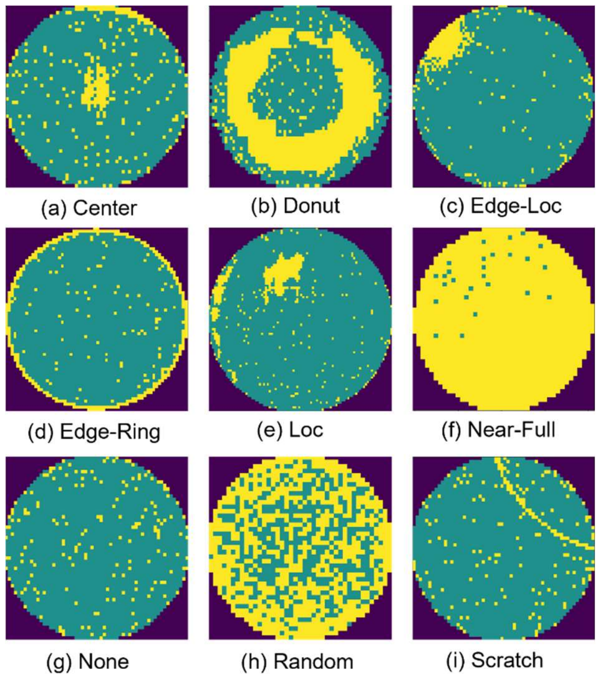


Fig. 1. Wafer bin maps by defect types

In this paper, we propose a regional information extraction methodology for each label of the defect pattern of WBM. As

* Corresponding author-Tel: +82-2-3290-3396; Fax: +82-2-3290-4550

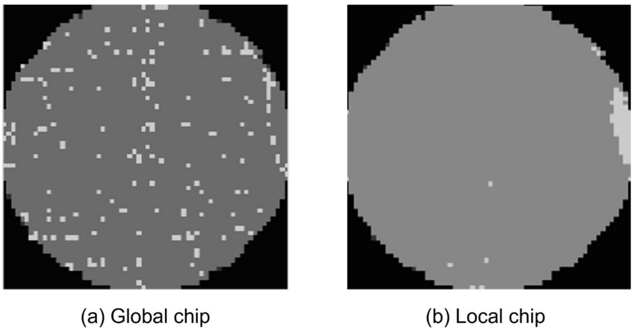


Fig. 2. Significant chip information in Edge-Loc pattern

shown in Fig. 2(a), common global information is removed regardless of the characteristics of the defect pattern.

Our paper is organized as follows. Section II review related studies, and in Section III, describe the structure of method for classifying defect patterns based on local information extraction. In Section IV, the performance of the proposed method is verified by conducting an experiment and describes the conclusion in Section V.

II. RELATED WORK

The purpose is to classify defect patterns by extracting only regional information representing defective patterns from the WBM. To learn the important characteristics of each pattern, consider main information without high uncertainty information.

This paper assumes that global information is an anomaly element. We introduce existing research on image anomaly detection method, feature extraction-based method and convolutional neural networks (CNN).

A. Image Anomaly Detection

In order to detect anomaly information in the image, only normal images are trained. Autoencoder (AE) learns by reducing the dimension of the input image through a bottleneck structure and then restoring it. An abnormal image entered for test, restored image only reflect the normal image feature. The reconstructed image is different from the abnormal image used as an input value

Problem of classifying images containing anomalous information, it is determined whether the image is anomaly as shown in Fig. 3 based on the difference from the generated image [4].

Fig. 3 show the framework for anomaly detection, finding difference between the input value x and the output value x’, threshold classify normal and abnormal. When the AE is trained using only normal data, a high error value is emitted for abnormal. Since the error value is larger than the set threshold, it can be classified as abnormal.

B. Feature Extraction-based Method

Denoising Autoencoder (DAE) learns noisy image by removing noise and extracting more features. Recently, a study

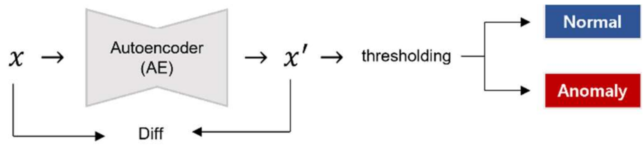


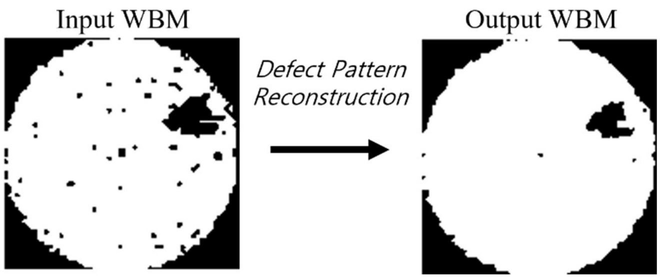
Fig. 3. Image anomaly detection using autoencoder

of classification by applying DAE method to extract features of wafers has been conducted [5]. This paper conducts Stacked Convolutional Sparse Denoising Autoencoder (SCSDAE) to filter noise on the wafer surface. However, the method compresses and classifies the features of the training data through a layered structure, an overfitting problem occurs in the training data. This has a limitation in misclassifying data with noise or deformed data.

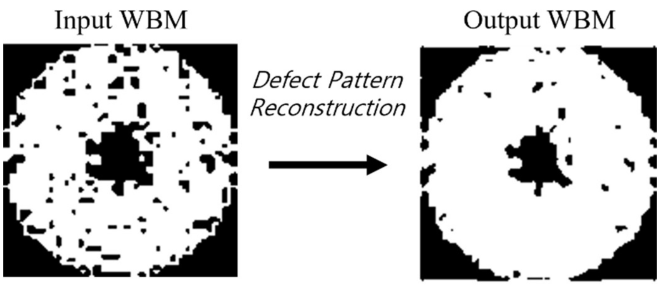
To solve this limitation, using skip connection technique was proposed. It does not learn details such as global information and noise in image. It learns without a bottleneck structure that stores dimensionally reduced features and delivers uncompressed information to decoder [6]. AE framework with skip connection structure shows good performance in removing noise from the image and obtaining key information.

III. METHOD

In this paper, we propose multi step learning model for regional information extraction of WBM. Step 1 is the process of extracting main features from images including noise and global information using Skip Connections Denoising Autoencoder (SCDAE). In step 2, anomaly score is calculated using the distance similarity between pixels of the existing WBM. In the last step, an image is generated based on anomaly score and CNN is trained to classify labels.



(a) Local pattern reconstruction result



(b) Center pattern reconstruction result

Fig. 4. Reconstruction result using SCDAE

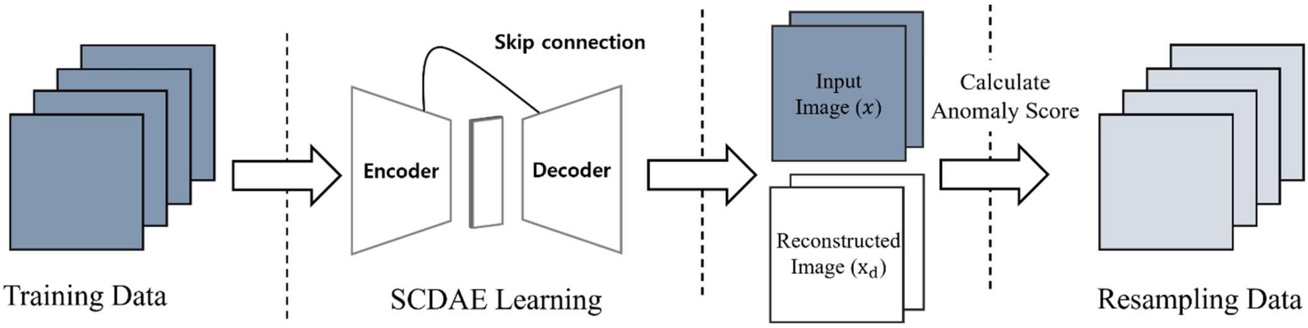


Fig. 5. Pipeline of proposed resampling method using SCDAE

A. SCDAE based Feature Extraction

WBM include regional and global chip information. Each chip has a binary value by setting a threshold value based on the pixel value of WBM that has completed the EDS process

The existing WBM pattern classification model was trained by reflecting all chip information in the data. This means that global information is also judged as important information and learned in the learning process.

Proposed method extracts an image $\tilde{x}$ is created by adding Gaussian noise to the input data. $\tilde{x}$ provides additional random information to remove global information. After that, it learns by comparing $\tilde{x}$ with the input x through the SCDAE model.

To extract local feature, add random noise to all data. Each pixel value with consecutive values in the range [0, 1]. Information from the front part of the encoder is transmitted to the decoder through the skip connection, and only major regional information that can be intuitively identified is learned. Through the learning process, image x_d including main information can be obtained as shown in Fig. 4. Fig. 4 shows the results of extracting the features of Local and Center patterns by applying SCDAE. Through this, it can be confirmed that the global chip information to the two defective patterns is removed, and local region information that can define the pattern is extracted.

B. Similarity based Image Generation

The generated image x_d is the result value of p(x|$\tilde{x}$) that follows a stochastic distribution for the image x to which noise is added during the learning process of SCDAE. In the learning process, values other than local information are removed to have a small value. Therefore, an anomaly score for the importance of each chip can be defined as in Equation 1.

$$Anomaly\ score = 1 - |\,x - x_d\,| \tag{1}$$

Based on the anomaly score defined in Equation 1, it is possible to determine the critical information of each chip in the WBM. Since the anomaly score is based on the difference between x and x_d, the chip information removed that has a high value. This mean that it is insignificant information of the defect pattern.

A. Stepwise Method for Classification

The proposed method is a stepwise process as shown in Fig. 3. x_r generated based on the threshold value includes area information of the WBM defect pattern. x_r not include information on chips with low importance.

The structure for each stage learned from data containing only local information contains only the main information of each defect pattern. Therefore, when new data is entered, the chip corresponding to global information is judged to be insignificant information and has a low probability value.

Last stage CNN consider WBM local location information together. Since the generated image reflects only local information, it learns location information and features for each bad pattern through CNN.

IV. EXPERIMENTS

A. Data Description

WM-811K used to classify defect patterns in WBM [7]. WM-811K consists of Center (2.5%), donut (0.3%), Edge-Loc (3.0%), Edge-Ring (5.6%), Loc (2.1%), Random (0.5%), Scratch (0.7%), Near-Full (0.1%) defect pattern.

8 types of defect pattern are used except the None pattern in Fig. 1 to perform verification to extract the main features of the defect pattern. Total of 17,625 data is used, 13,128 are used as training and 4,407 are used as test.

B. Wafer Map Classification Accuracy

For the performance evaluation of a model with image reconstructed by SCDAE as an input, accuracy is used for image classification model performance evaluation. Each index of the evaluation matric indicates the following. TP (True Positive) and TN (True Negative) in Equation 2 are cases in which labels are correctly classified. FP (False Positive) and FN (False Negative) are indicators indicating the result of classification differently from the actual label. The combined accuracy represents the percentage of correctly classified labels for the entire data.

$$\text{Accuracy} = \frac{TP+TN}{TP+TN+FP+FN} \qquad (2)$$

Among the 8 types of defective patterns presented in Fig. 1, 8 types of defective patterns excluding normal patterns are classified by applying the proposed model.

TABLE I. CLASSIFICATION ACCURACY

Model	CNN	AE CNN	DAE CNN	Proposed Method
Center	99	100	99	100
Donut	93	77	92	**94**
Edge-Loc	85	91	88	**95**
Edge-Ring	97	94	98	**98**
Loc	83	90	83	83
Near-Full	92	85	98	**100**
Random	87	4	93	93
Scratch	63	63	73	**73**

Table I shows the results of classifying 8 types of defect patterns by applying 4 types of method including the proposed model. All CNNs included in the four methodologies have the same structure.

Based on the accuracy of the proposed method, Donut, Edge-Loc, Edge-Ring, Near-Full, and Scratch patterns showed higher performance than the existing methodologies and showed similar performance for three defective patterns. This is a result showing that the proposed method extracts the local information of the wafer bin map defect pattern well and the model shows robust characteristics against new data.

V. CONCLUSION

The method proposed in this paper is to classify defect patterns by extracting regional information from the WBM. Related research has limitations in applying the deep learning method that reflects all chip information in the WBM. Inconsequential information for each defect pattern was also considered, resulting in confusion among some patterns.

SCDAE is an object localization methodology that used to extract local information of WBM defect patterns. Through this, an anomaly score was defined to generate an image including major features for each defect pattern and to define global information that could be generated from a new input value. Based on the uncertainty, global information was removed and data reflecting only the main information of the defect pattern was generated. Afterwards, it was confirmed that unique regional information for each defect pattern was extracted for new data through the learning process. Through this, it was found through performance that it is possible to robustly classify the deformed data or the data with added noise.

The main information extraction-based learning method proposed in this study confirms the main characteristics of each defect pattern and contributes to identifying the cause of the defect. Through this, it is possible to define the main characteristics of each defect pattern, and it is possible to easily classify the defect patterns collected in the future process.

ACKNOWLEDGMENT

This work was supported by the National Research Foundation of Korea (NRF) grant funded by the Korean government (MSIT) (NRF-2019R1A2C2005949). Also, this work was supported by Brain Korea 21 FOUR and Samsung Electronics Co., Ltd(IO201210- 07929-01).

REFERENCES

[1] Uzsoy, R., Lee, C. Y., Martin-Vega, L. A. "A review of production planning and scheduling models in the semiconductor industry part I: system characteristics, performance evaluation and production planning." IIE transactions, vol. 24, no. 4, pp. 47-60, 1992.

[2] Do, H., Lee, C., and Kim, S. B. "A Hierarchical Spatial-Test Attention Network for Explainable Multiple Wafer Bin Maps Classification." IEEE Trans. Semicond. Manuf., 2021.

[3] Saqlain, M., Jargalsaikhan, B., and Lee, J. Y. "A voting ensemble classifier for wafer map defect patterns identification in semiconductor manufacturing." IEEE Trans. Semicond. Manuf., vol. 32, no. 2, pp 171-182, 2019.

[4] Bergmann, P., Löwe, S., Fauser, M., Sattlegger, D., amd Steger, C. "Improving unsupervised defect segmentation by applying structural similarity to autoencoders." *arXiv,* 2018.

[5] Yu, J., Zheng, X., and Liu, J. "Stacked convolutional sparse denoising auto-encoder for identification of defect patterns in semiconductor wafer map." Computers in Industry 109, pp. 121-133, 2019.

[6] Mao, X., Shen, C., and Yang, Y. B. "Image restoration using very deep convolutional encoder-decoder networks with symmetric skip connections." Advances in neural information processing systems 29, pp. 2802-2810, 2016.

[7] Wu, Ming-Ju, Jyh-Shing R. Jang, and Jui-Long Chen. "Wafer map failure pattern recognition and similarity ranking for large-scale data sets." IEEE Trans. Semicond. Manuf., vol. 28, no.1. pp. 1-12, 2014.