

Review of Wafer Surface Defect Detection Methods

Jianhong Ma, Tao Zhang, Cong Yang, Yangjie Cao, Lipeng Xie *, Hui Tian * and Xuexiang Li *

School of Cyber Science and Engineering, Zhengzhou University, Zhengzhou 450000, China

* Correspondence: lipengxie@zzu.edu.cn (L.X.); tianhui@zzu.edu.cn (H.T.); lxx@zzu.edu.cn (X.L.)

Abstract: Wafer surface defect detection plays an important role in controlling product quality in semiconductor manufacturing, which has become a research hotspot in computer vision. However, the induction and summary of wafer defect detection methods in the existing review literature are not thorough enough and lack an objective analysis and evaluation of the advantages and disadvantages of various techniques, which is not conducive to the development of this research field. This paper systematically analyzes the research progress of domestic and foreign scholars in the field of wafer surface defect detection in recent years. Firstly, we introduce the classification of wafer surface defect patterns and their causes. According to the different methods of feature extraction, the current mainstream methods are divided into three categories: the methods based on image signal processing, the methods based on machine learning, and the methods based on deep learning. Moreover, the core ideas of representative algorithms are briefly introduced. Then, the innovations of each method are compared and analyzed, and their limitations are discussed. Finally, we summarize the problems and challenges in the current wafer surface defect detection task, the future research trends in this field, and the new research ideas.

Keywords: wafer defect pattern; image processing; defect identification method; feature extraction

1. Introduction

Silicon wafers are used in the manufacturing of semiconductor chips. The required patterns are formed on the wafers through lithography and other processes and are very important carriers in the semiconductor chip manufacturing process. In the manufacturing process, due to the influence of factors such as the environment and process parameters, defects will be generated on the surface of the wafer, which will affect the yield of wafer production. The accurate detection of wafer surface defects can accelerate the identification of abnormal faults in the manufacturing process as well as the adjustment of the manufacturing process, improve production efficiency, and reduce scrap rates.

Early detection of wafer surface defects is often performed manually by experienced inspectors, which has problems such as low efficiency, poor accuracy, high cost, and strong subjectivity, which are insufficient to meet the requirements of modern industrialized products. At present, defect detection methods based on machine vision [1] have replaced manual inspection in the field of wafer inspection. Traditional machine vision-based defect detection methods often use manual feature extraction, which is inefficient. The emergence of computer vision-based detection methods [2], especially the advent of neural networks such as convolutional neural networks, has addressed the limitations of data preprocessing, feature representation and extraction, and model learning strategies [3]. With their high efficiency, accuracy, low cost, and strong objectivity, neural networks have rapidly developed and have been widely applied in the field of surface defect detection in semiconductor wafers.

In recent years, with the development of electronic integrated circuits, such as smart terminals and wireless communication facilities, and the promotion of Moore's Law [4], while the global demand for chips has increased, the accuracy of the lithography process has increased. With the advancement of technology, process precision has reached below



Citation: Ma, J.; Zhang, T.; Yang, C.; Cao, Y.; Xie, L.; Tian, H.; Li, X. Review of Wafer Surface Defect Detection Methods. *Electronics* **2023**, *12*, 1787. <https://doi.org/10.3390/electronics12081787>

Academic Editor: Donghyeon Cho

Received: 6 February 2023

Revised: 24 March 2023

Accepted: 6 April 2023

Published: 10 April 2023



Copyright: © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

10 nm [5]. As a result, higher requirements have been placed on the yield of each process step, posing greater challenges for defect detection techniques in wafer manufacturing.

This paper mainly summarizes the related research on wafer surface defect detection algorithms, including traditional image processing, machine learning, and deep learning. According to the characteristics of the algorithm, the relevant literature is summarized and organized, and the problems and challenges faced in the field of wafer defect detection and future development are prospected. This paper intends to help in the quick understand of relevant methods and skills in the field of wafer surface defect detection.

2. Wafer Surface Defect Patterns

In actual production, there are many kinds of defects on wafers, and the shapes are not uniform, which increases the difficulty of wafer defect detection. Among the types of wafer defects, unpatterned wafer defects and patterned wafer defects are the two main forms of wafer defects. These two defects are the main cause of chip failure. Unpatterned wafer defects mostly occur in the pre-lithography stage of wafer production [6], which are wafer defects caused by machine failures. The scratch defect is shown in Figure 1a, and the particle contamination defect is shown in Figure 1b. Patterned wafer defects are mostly found in the middle process of wafer production. Improper exposure time, development time, and post-baking time lead to defects in photolithographic lines. Defects on the wafer surface generated during micro/nano-fabrication of spiral excitation coils and fork-shaped electrodes are shown in Figure 2. The open circuit defect is shown in Figure 2a, the short circuit defect is shown in Figure 2b, the line contamination defect is shown in Figure 2c, and the bite defect is shown in Figure 2d.

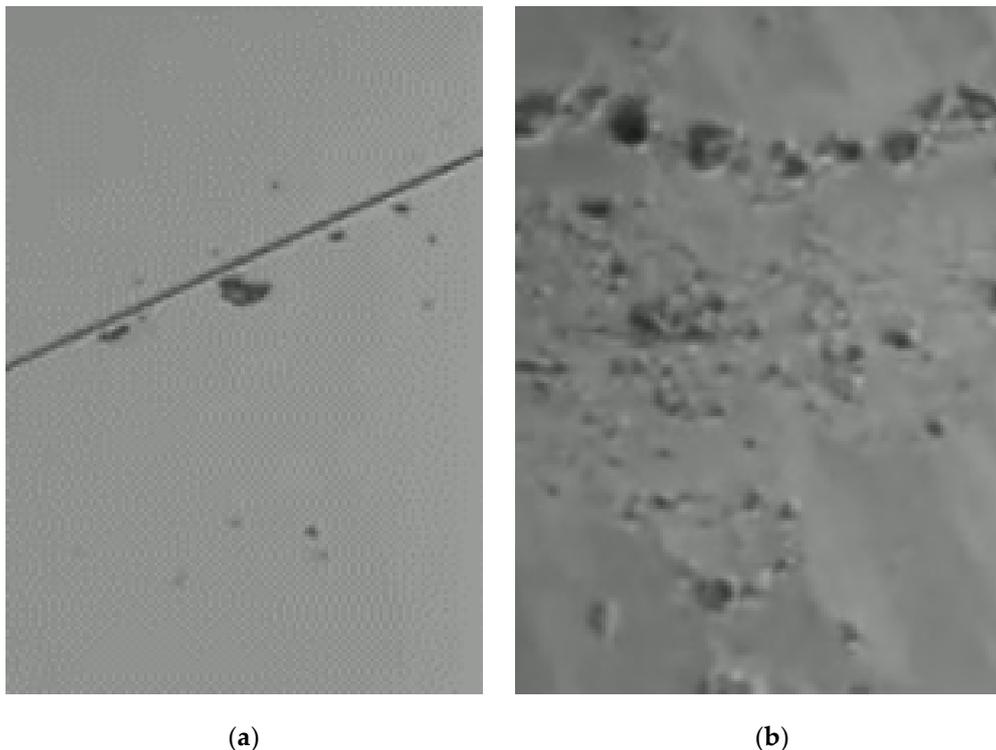


Figure 1. (a) Scratch defects in unpatterned wafers; (b) particulate contamination in unpatterned wafers.

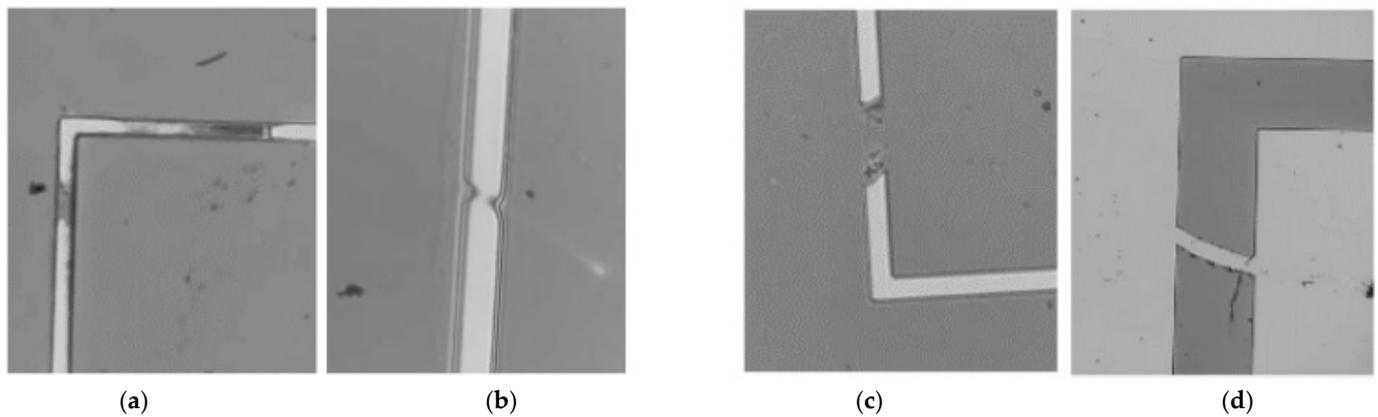


Figure 2. (a) Open defects, (b) short defects, (c) line contamination, and (d) bite defects in a patterned wafer defect map.

Due to the existence of the above-mentioned wafer defects, when the functional integrity test of all the chips on the wafer is performed, chip failures may occur. The chip engineer marks the test results with different colors to distinguish the position of the chip. Under the influence of different operating processes, corresponding specific spatial patterns will be generated on the wafer [7]. The wafer image data, i.e., the wafer map, is thus generated. As stated by Hansen et al. [8] in 1997, defective chips usually have a clustering phenomenon or show some systematic patterns, and this defect pattern usually contains the necessary information on process conditions [9]. The wafer map can not only reflect the integrity of the chip but also accurately describe the spatial position information corresponding to the defect data. Wafer maps may exhibit spatial dependencies across the wafer, and chip engineers can often trace the cause of defects and resolve problems based on defect type. Mirza et al. [10] divided wafer map defect patterns into general and local types, namely global random defects and local defects. The wafer map defect pattern diagram is shown in Figure 3, the local defects are shown in Figure 3a, and the global random defects are shown in Figure 3b. Global random defects are generated by uncertain factors, which are uncontrollable factors without specific clustering phenomena, such as dust particles in the environment. Global random defects can only be reduced through long-term, incremental improvements or expensive equipment overhaul programs. Localized defects are inherent to a system and are influenced by controllable factors during the wafer production process, such as process parameters, equipment issues, and improper operation. They appear on the wafer repeatedly and exhibit a certain level of clustering. Identifying and classifying local defects to locate equipment anomalies and inappropriate process parameters plays a vital role in improving the yield rate of wafer production.

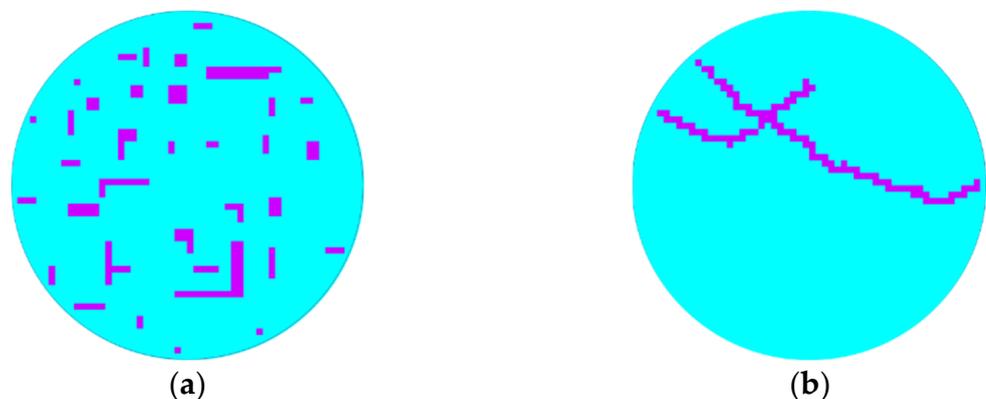


Figure 3. (a) Local defect pattern; (b) global defect pattern.

For wafer patterns with a large area, low feature size, low density, and low integration, the lithography path can be observed with an electron microscope, and trace detection can be performed directly. With the significant increase in the integration level of chip circuits, it has become increasingly difficult to perform chip-level inspection. This is because as integration level increases, the components on the chip become smaller, more complex, and more densely packed, resulting in a greater number of potential defects [11]. These defects are difficult to detect and repair through conventional inspection methods, requiring more sophisticated and advanced detection technologies and tools.

The research on wafer maps is a hot topic in wafer defect detection [12–15]. Liu Fengzhen of Tianjin University [16] researched wafer map defects caused by abnormal lithography equipment. For the defects in the actual production process of wafers, we have conducted in-depth research on photoresist, wafer dust particles, and wafer ring, scratch, spherical, linear, and other defects through equipment experiments, aiming to find the cause of defects to improve productivity. To determine the cause of wafer mode failure, Ming-Ju Wu [11] et al. collected 811,457 real wafer maps from actual manufacturing and created the WM-811K wafer map dataset, which is currently the most widely used wafer map. Semiconductor domain experts annotated eight defect pattern types for approximately 20 percent of the wafer maps in this dataset. The eight types of wafer map defect patterns are shown in Figure 4. Most of the articles cited in this review conducted their testing based on this dataset.

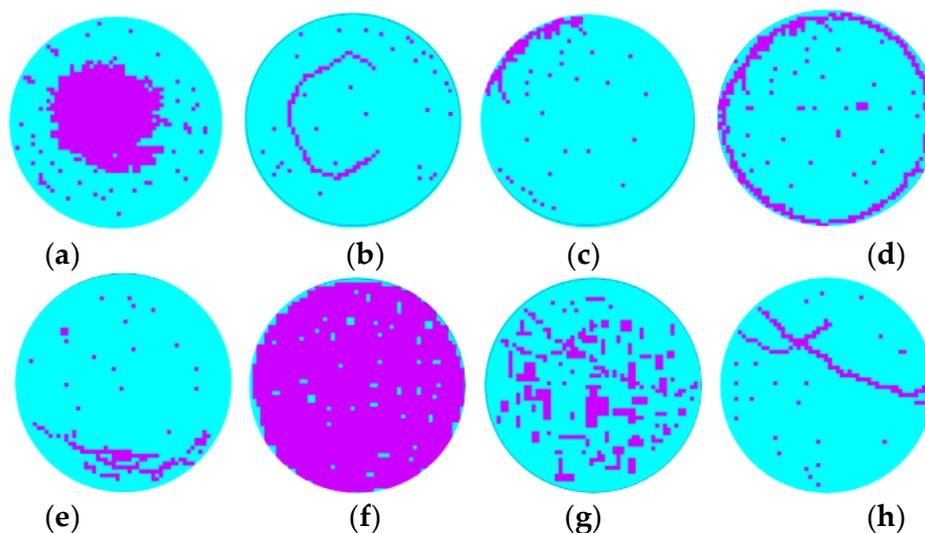


Figure 4. Eight types of wafer map defect pattern types: (a) center, (b) donut, (c) edge-loc, (d) edge-ring, (e) local, (f) near-full, (g) random, and (h) scratch.

3. Wafer Surface Defect Detection Based on Image Signal Processing

Image signal processing is used to convert image signals into digital signals, which are then processed via computer technology to achieve image transformation, enhancement, and detection [17–19]. Commonly used in the field of wafer inspection are wavelet transform (WT), spatial filtering (spatial filtering), and template matching (template matching). This section mainly introduces the application of these three algorithms on wafer surface defect detection. The comparison of image processing algorithms is shown in Table 1.

Table 1. Comparison of image processing algorithms.

| Model Algorithm | Innovation | Limitation |
|------------------------------|---|--|
| Wavelet transform [12,20–24] | The image can be decomposed into multiple resolutions and presented as local sub-images with different spatial frequencies. Anti-grain. | The selection of the threshold is very dependent and the adaptability is poor. |

Table 1. Cont.

| Model Algorithm | Innovation | Limitation |
|---------------------------------|---|---|
| Spatial filtering [25–33] | Based on spatial convolution, remove high-frequency noise, and perform edge enhancement. | Performance depends on the threshold parameter. |
| Template matching [11,17,34–36] | The template matching algorithm has strong anti-noise ability and fast calculation speed. | Sensitive to feature object size. |

3.1. Wavelet Transform

Wavelet transform (WT) [20] is a signal time–frequency analysis and processing technology. Firstly, the image signal is decomposed into different frequency subbands through a filter to perform wavelet decomposition. Then, by calculating the mean, standard deviation, or other statistical measures of the wavelet coefficients, each coefficient is analyzed to detect any anomalies or defects. Anomalies or defects may manifest as sudden changes or outliers in the wavelet coefficients. Based on the analysis results, pre-defined thresholds are used to determine the defects and anomalies in the signal, and the location of the defect is determined by identifying which time and frequency subband it is in. The diagram of the wavelet decomposition principle is shown in Figure 5, where L represents low-frequency information and H represents high-frequency information. Each time the image is decomposed, the image is decomposed into four frequency bands: LL, LH, HL, and HH. The lower layer decomposition repeats the decomposition on the upper layer LL band. Wavelet transform has good performance in boundary processing of wafer defect features [21] and multi-scale edge detection [22].

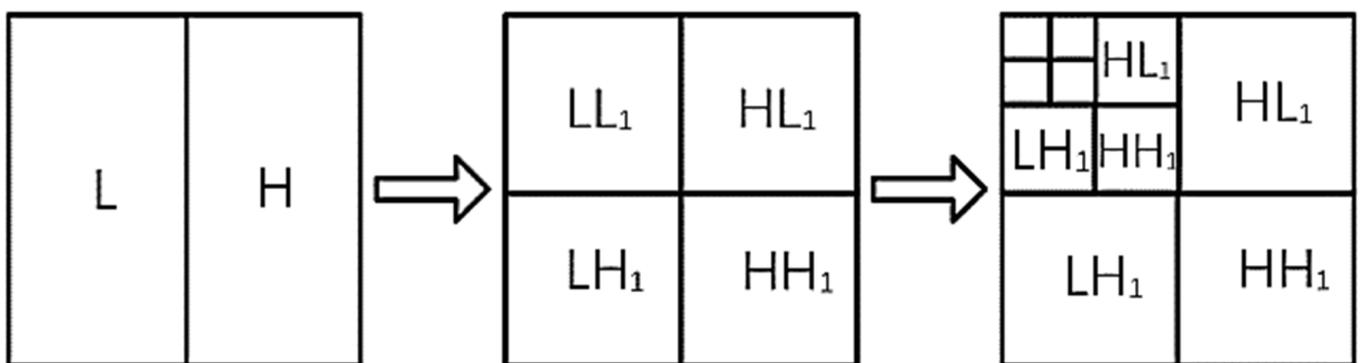


Figure 5. Schematic diagram of wavelet decomposition.

Yeh et al. [23] proposed a method based on two-dimensional wavelet transform (2DWT), where the ratio between scale coefficients is calculated by a modified wavelet to transform modulus (WTMS) for the localization of wafer defect pixels. By selecting the appropriate wavelet base and support length, accurate detection of wafer defects using a small amount of test data can be achieved. The image preprocessing stage takes a lot of time, which seriously affects the detection speed. Wen-Ren Yang et al. [24] proposed an online wafer microcrack detection system based on short-time discrete wavelet transform. There is no need to preprocess the wafer image. By emitting a continuous pulsed laser beam to the wafer surface, the reflected signal is collected by a space probe array, and analyzed by discrete wavelet transform to determine the reflection characteristics of microcracks. In the case of processing, it can also have a better detection effect on micro-cracks. A large number of random wafer particles exist on the surface of polycrystalline solar wafers, resulting in uneven texture in wafer sensing images. In response to this problem, Kim Y et al. [12] proposed a surface inspection method based on wavelet transform to detect solar wafer defects. To better distinguish defect edges from grain edges, the energy differences of the wavelet detail subgraphs of two consecutive decomposition levels are used as

weights to enhance the discriminative features proposed in each decomposition level. The experimental results show that this method has a good detection effect on fingerprints and stains, but the method is not effective on serious micro-crack defects with sharp edges, and cannot be applied to all defects.

3.2. Spatial Filtering

Spatial filtering is a mature image enhancement technique [25], which is realized by directly applying spatial convolution [26] on the gray value. The main role in image processing is image denoising, which is divided into smoothing filters [27] and sharpening filters [28], which are widely used in the field of defect detection. Figure 6 shows the denoising effect of the median filter and mean filter in the image with added noise.

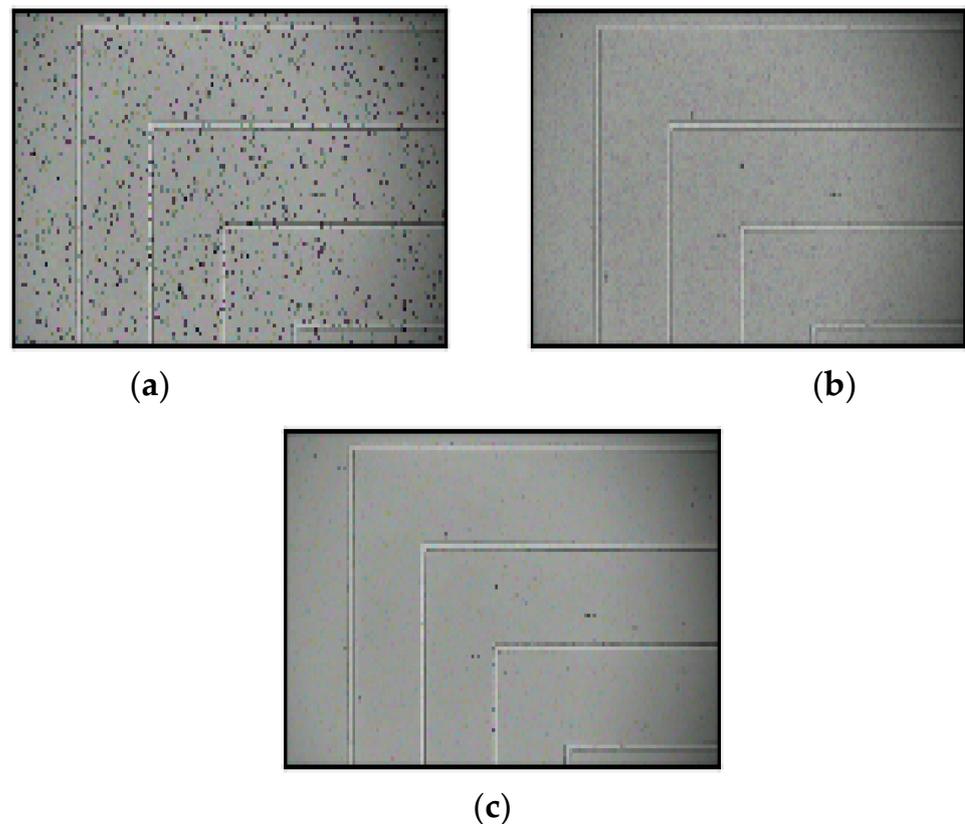


Figure 6. Effect picture of filtering and denoising: (a) original image, (b) median filter denoising, and (c) mean filter denoising.

Ohshige et al. [29] proposed a surface defect detection system based on spatial frequency filtering techniques. The method can effectively detect submicron defects or foreign particles on wafers. Effects of random defects in wafer manufacturing. C.H. Wang [30] proposed a wafer defect detection method based on spatial filtering, entropy-fuzzy c-means, and spectral clustering [31], which uses spatial filtering to denoise and extract defect regions, which are obtained by entropy-fuzzy c-means [29] and spectral clustering [32]. A hybrid algorithm combining mean and spectral clustering is used for defect classification. It solves the problem that traditional statistical methods cannot extract defect patterns with meaningful classification. For clustered defects in wafers, Chen SH et al. [33] developed a software tool based on median filtering and clustering methods, and the proposed algorithm effectively detected defect clusters. Often the performance of spatial filters is highly parameter-dependent, and it is often difficult to choose their values.

3.3. Template Matching

Template matching [36,37] detection is achieved by calculating the similarity between the template image and the image under test, to detect the difference area between the image under test and the template image.

Han H [36] et al. obtained templates from the wafer image itself mixed into the design layout scheme of the wafer fabrication process and used the mapping between the physical space and the pixel space to design a wafer image detection technology that combines existing a new method for circle template matching detection. Xifeng Liu [16] combined the SURF image registration algorithm to achieve spatial positioning matching between the test wafer and the standard wafer patterns. The feature point matching result between the test image and the standard image is shown in Figure 7. The contour extraction technique of pattern recognition is applied to wafer defect detection. Khalaj et al. [36] proposed a new technique that uses a high-resolution spectral estimation algorithm to extract wafer defect features and compare them with actual images to detect periodic 2D signals or the location of irregularities and defects in the image.

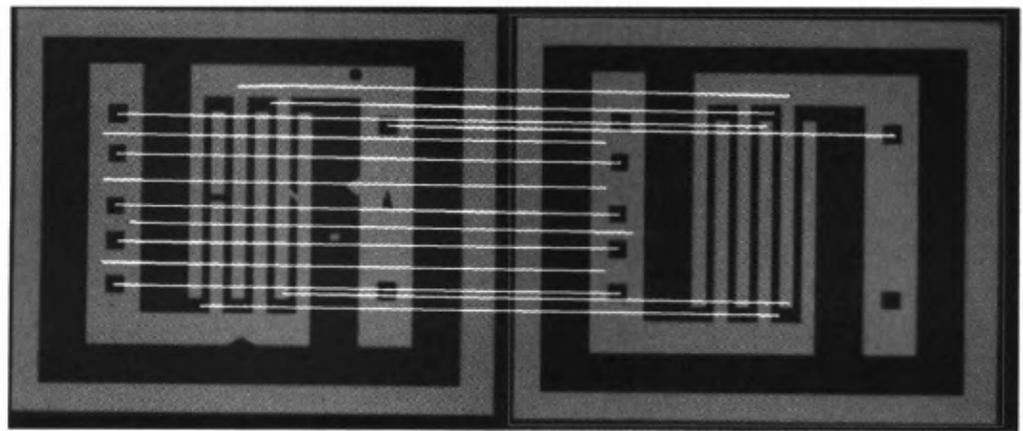


Figure 7. The feature point matching result between the test image and the standard image.

4. Wafer Surface Defect Detection Based on Machine Learning

Machine learning is mainly to abstract a specific problem into a mathematical model, solve the model through mathematical methods, solve the problem, and then evaluate the effect of the model on this problem. According to the characteristics of training data, it is divided into supervised learning, unsupervised learning, and semi-supervised learning. This paper mainly discusses the application of these three machine-learning methods in wafer surface defect detection. The machine learning model comparison is shown in Table 2.

Table 2. Comparison on machine learning algorithms.

| Classification | Algorithm | Innovation | Limitation |
|-----------------------------|---------------------|--|---|
| Supervised learning [38–46] | KNN | Insensitive to abnormal data and highly accurate. | High complexity and computation intensity. |
| | Decision Tree-Radon | Apply Radon to form new defect features. | Overfitting is highly proficient. |
| | SVM | SVM efficiently classifies multivariate, multi-modal, and indivisible data points. | It is not friendly to multiple samples, and the kernel function is difficult to locate. |

Table 2. Cont.

| Classification | Algorithm | Innovation | Limitation |
|----------------------------------|--|--|--|
| Unsupervised learning [47–50] | Multilayer Perceptron-Clustering Algorithm | The multilayer perceptron is used to enhance the feature extraction capability. | Depends on the choice of activation function. |
| | DBSCAN | Outliers can be selectively removed based on defect pattern characteristics. | The sample density is not uniform or the sample is too large, the convergence time is long, and the clustering effect is poor. |
| | SOM | High-dimensional data can be mapped to a low dimensional space and the structure of the high-dimensional space can be maintained. | The objective function is not easy to determine. |
| Semi-supervised learning [51–53] | A semi-supervised framework for augmented labeling Semi-Supervised Incremental Modeling Framework | A semi-supervised model is built by combining supervised ensemble learning and unsupervised SOM. Improve model performance by actively learning and labeling samples to enhance them. | Training is time-consuming and time-consuming. Performance depends on the amount of data tagged. |

4.1. Supervised Learning

Supervised learning is a learning model [38] created from labeled training data, and based on this model, it makes predictions about the new data samples needed. Supervised learning is currently a widely used machine learning algorithm in wafer surface defect detection and has high robustness in the field of object detection.

Yuan, T et al. [40] proposed a noise removal technique based on the k-Nearest Neighbor (KNN), which uses the k-Nearest Neighbor algorithm to separate global and local defects, providing information about all aggregated local defects of the wafer information, classify defects into clusters through similarity clustering techniques, and identify spatial patterns of defect clusters using a parametric model of clustered defects. Piao M et al. [41] proposed a decision tree-based wafer defect pattern recognition method. The radon transformation is used to extract the defect pattern features, the correlation analysis method is used to measure the correlation between the features, and the defect features are divided into feature subsets, and each feature subset builds a decision tree according to the C4.5 mechanism. The decision tree confidences are summed and the category with the highest overall confidence is selected. The decision tree shows better performance in specific categories of wafer defect detection, but the maximum, minimum, average, and standard deviation of projections are not enough to represent all the spatial information of wafer defects, so the edge defect detection performance is poor.

Support Vector Machine (SVM) in supervised learning is also a mature application of defect detection. When the samples are unbalanced, the k-nearest neighbor algorithm has a poor classification effect and a large amount of computation. Decision trees have similar problems and are prone to overfitting. The support vector machine still has good performance in the classification of small samples and high-dimensional features, and the computational complexity of the support vector machine does not depend on the dimension of the input space, and the multi-class support vector machines are robust to overfitting problems, so it is often used as a classifier [42–44]. R. Baly et al. [45] used a Support Vector Machine (SVM) classifier to classify 1150 wafer images into two categories, high-yield and low-yield, and then demonstrated through comparative experiments that relative to decision trees, k-Nearest Neighbor (KNN), Partial Least Squares Regression (PLS regression) and Generalized Regression Neural Network (GRNN), the nonlinear support vector machine model is better than the above four wafer classification methods. The multi-class SVM has better classification accuracy in the classification of wafer defect patterns. L. Xie et al. [46] proposed a wafer defect pattern detection scheme based on a support vector machine algorithm. The linear kernel, Gaussian kernel, and polynomial kernel are

used for selective testing, and the kernel with the smallest test error is selected through cross-validation for the next step of support vector machine training. The support vector machine method can deal with the problem of false positives caused by image translation or rotation. Compared with neural networks, SVM does not require a large number of training samples, so it does not need to spend a lot of time training data samples for classification. More robust performance for composite or diverse datasets.

4.2. Unsupervised Learning

In supervised learning, researchers need to classify defect sample types in advance as prior knowledge for training. In actual industrial production, there are a large number of unknown defects and the defect characteristics are ambiguous, and it is difficult for researchers to judge and classify through experience. In the early stages of process development, sample annotation is also limited. Aiming at these problems, unsupervised learning [47] has opened up new solutions, which do not need a lot of manpower for labeling data samples and performing clustering according to the feature relationship between samples. Unsupervised learning also has advantages when new defect patterns are added. In recent years, unsupervised learning has become one of the important research directions of industrial defect detection.

The defect patterns on the wafer pattern are unevenly classified and the features are irregular, and the unsupervised clustering algorithm has strong robustness against this situation and is widely used to detect complex wafer defect patterns. Due to the difficulty in detection as a result of cluster defects, such as scratches, stains, or local failure modes, C. J. Huang [48] proposed a new method to solve this problem. An automatic wafer defect clustering algorithm (k-means clustering) using self-supervised multilayer perceptrons to detect defects and label all defective chips was proposed. Jin C H et al. [49] proposed a wafer pattern detection and classification framework based on Density-Based Spatial Clustering of Applications with Noise (DBSCAN), which selectively removed outliers according to defect pattern features, and then the extracted defect features can complete the detection of abnormal points and defect patterns at the same time. Yuan, T et al. [40] proposed a multi-step wafer analysis method, which provides clustering results with different precisions based on a similarity clustering technique to identify six mixed-type defect patterns according to the spatial location of local defect patterns. Using location information to distinguish defect clusters has certain limitations, and when multiple clusters are close to or overlap with each other, the classification effect will be affected.

Di Palma, F et al. [50] employed unsupervised Self-Organizing Map (SOM) and Adaptive Resonance Theory (ART1) as wafer classifiers on nine different classes of wafers tested on a simulated dataset. Both SOM and ART1 rely on the competition between neurons to gradually optimize the network for unsupervised classification. Since ART1 is pushed to the reference vector by “AND” logic, when processing a large number of data sets, the number of calculations increases, and the real number of defect categories cannot be obtained. Adjusting the network identification threshold does not make any improvement. The SOM algorithm can map high-dimensional input data to a low-dimensional space while maintaining the topological structure of the input data in the high-dimensional space. First, the category and number of neurons are determined, and other parameters are determined through several comparison experiments. After determining the parameters, after several learning cycles, the data reaches an asymptotic value and performs well on both simulated and real datasets.

4.3. Semi-Supervised Learning

Semi-supervised learning is a machine learning method that combines supervised learning and unsupervised learning [51]. Semi-supervised learning can use a small amount of labeled data and a large amount of unlabeled data to solve problems. The ensemble-based semi-supervised learning process is shown in Figure 8. The cost consumption and

mislabeling of fully labeled samples are avoided. Semi-supervised learning has become a hot topic of research in recent years.

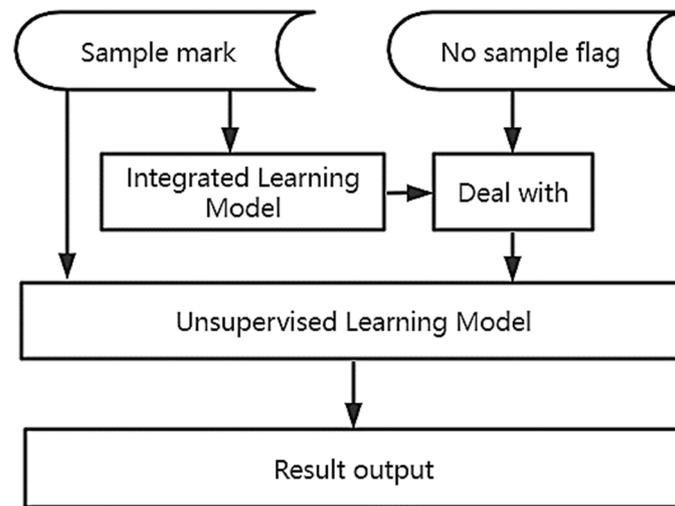


Figure 8. Ensemble-based semi-supervised learning.

Supervised learning usually achieves good recognition results, but relies on the accuracy of sample labeling. The wafer data samples may have the following problems. The first is that the wafer sample data needs to be manually marked by professionals. The manual marking process is subjective, and some mixed defect modes may be incorrectly marked. The second is that some defect modes have insufficient samples. The third is that some defect patterns are not marked in the first place. As a result, unsupervised learning methods cannot exert their performance. Aiming at this problem, Katherine Shu-Min Li [52] et al. proposed an ensemble-based semi-supervised framework to realize the automatic classification of defect patterns. First, a supervised ensemble learning model was trained on labeled data, and then the unlabeled data were trained by this model. Finally, the unsupervised learning algorithm is used to process the samples that cannot be correctly classified to achieve an enhanced labeling effect and improve the accuracy of wafer defect pattern classification. Yuting Kong and Dong Ni [53] proposed a semi-supervised incremental modeling framework for wafer map analysis. A semi-supervised incremental model improved by a ladder network and SVAE model is used to classify wafer maps, and then the model performance is improved by active learning and pseudo-labeling. Experiments show that it has a better performance than the CNN model.

5. Wafer Surface Defect Detection Based on Deep Learning

In recent years, with the development of deep learning algorithms, the improvement of GPU computing power, and the emergence of the convolutional neural network [54], the field of computer vision has been qualitatively developed, and it has also been widely used in the field of surface defect detection. Before deep learning, relevant personnel were required to have extensive knowledge of feature mapping and feature description to be able to draw features manually. Deep learning enables multi-layer neural networks to automatically extract and learn target features through abstraction layers, and detect target objects from images.

Cheng KCC et al. [55] used machine learning algorithms and deep learning algorithms for wafer defect detection, respectively. They used Logistic Regression [56], Support Vector Machine (SVM) [57], Adaptive Boosting Decision Tree (ADBT) [58], and deep neural networks to detect wafer defects. Experiments have proved that the average accuracy of the deep neural network is better than the above machine learning algorithms, and the wafer detection algorithm based on deep learning has a better performance. According to different application scenarios and task requirements, the deep learning model is di-

vided into a classification network, detection network, and segmentation network. This section discusses the innovations and compares the performance of each deep learning network model.

5.1. Classification Network

Classification networks are one of the older deep learning algorithms. The classification network extracts the characteristic information of the target object in the input image through a series of operations, such as convolution [54] and pooling [58], and then passes through the fully connected layer [59] to classify according to the preset label category. The network model is shown in Figure 9. In recent years, many classification networks for specific problems have emerged. In the field of wafer defect detection, focusing on defect features and enhancing feature extraction capabilities has promoted the development of wafer detection.

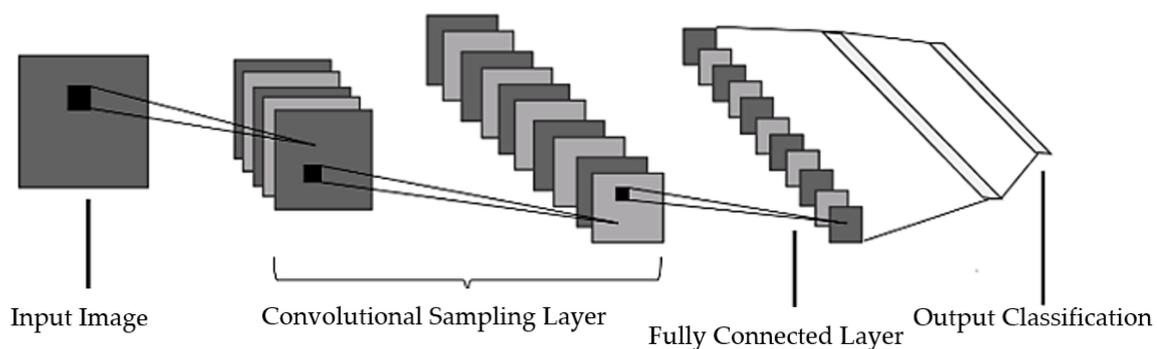


Figure 9. Classification network model structure diagram.

In the wafer manufacturing process, several different types of defects are coupled in a wafer, called mixed defects. These types of defects are complex, changeable, and highly random, and have become a key challenge for semiconductor companies. In response to this problem, Wang J et al. [60] proposed a hybrid DPR (MDPR) deformable convolutional network (DC-Net) for wafer defect classification. They designed a multi-label output of deformable convolution and a one-hot encoding mechanism layer, focusing the sampling area on the defect feature area, effectively extracting defect features, classifying mixed defects, outputting single defects, and improving the classification accuracy of mixed defects. Kyeong and Kim [61] designed a separate classification model for each type of defect in the wafer image of the mixed defect mode and detected the defect mode of the wafer through the combined classifier network. The authors tested the wafer map database of six different modes with MPL, SVM, and CNN combined classifiers, and only the algorithm proposed by the authors was classified correctly. Takeshi Nakazawa and Deepak V. Kulkarni [62] used CNN for wafer defect pattern classification. They trained and validated their CNN model using synthetically generated wafer images. Additionally, a method of using simulated generated data was proposed to solve the problem of unbalanced data of real defect categories in manufacturing and achieve reasonable classification accuracy. This effectively solves the problem of difficult wafer data collection and few available samples. The classification network model comparison is shown in Table 3.

Table 3. Classification network model comparison.

| Algorithm | Innovation | Acc |
|------------------------------------|---|-------|
| DC-Net [60] | The sampling area is focused on the defect feature area, which is very robust to mixed defects. | 93.2% |
| CNN-Based Combined Classifier [61] | Separately design classifiers for each defect, strong adaptability to new defect modes. | 97.4% |

Table 3. Cont.

| Algorithm | Innovation | Acc |
|---|---|-------|
| Classification Retrieval Method Based on CNN [62] | Simulated datasets can be generated to account for data imbalances. | 98.2% |

5.2. Object Detection Network

The target detection network can not only classify the target object but also identify its location. The target detection network is mainly divided into two types. The first type is the two-stage network [63,64], as shown in Figure 10. The candidate boxes are generated based on the region proposal network, and then the candidate boxes are classified and regressed. The second category is single-stage networks [65–67], as shown in Figure 11, that is, end-to-end object detection, which directly generates classification and regression information of target objects without generating candidate boxes. Relatively speaking, the two-stage network has a higher detection accuracy, and the single-stage network has a faster detection speed. The comparison of the detection network models is shown in Table 4.

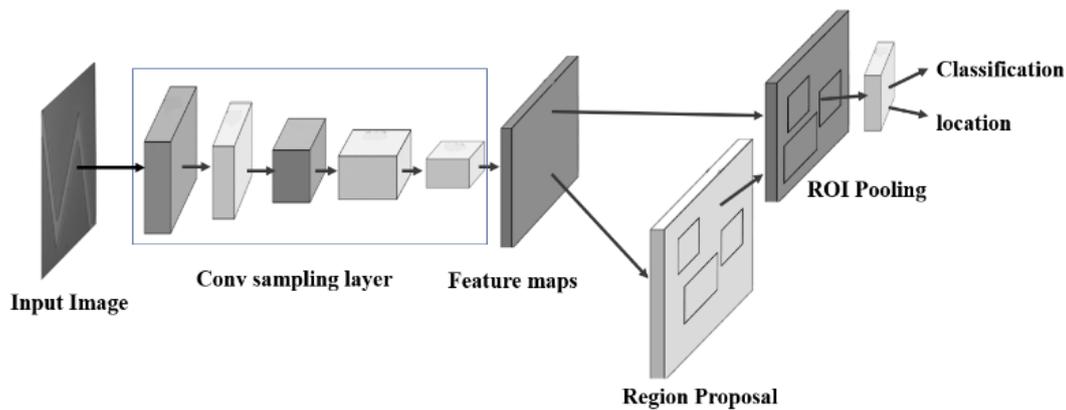


Figure 10. Schematic diagram of the structure of the two-stage detection network model.

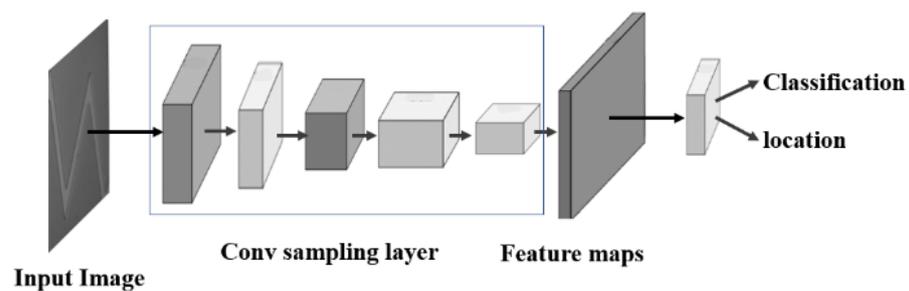


Figure 11. Schematic diagram of the structure of the one-stage detection network model.

Table 4. Detection network model comparison.

| Algorithm | Innovation | Acc | Ap |
|-----------------|--|--------|--------|
| PCACAE [68] | Automatic coding of concatenated roll types based on two-dimensional principal component analysis. | 97.27% | \ |
| YOLOv3-GAN [33] | GAN enhances the diversity of defect patterns and improves the versatility of YOLOv3. | \ | 88.72% |
| YOLOv4 [69] | Updated backbone network, enhanced with CutMix and Mosaic data. | 94.0% | 75.8% |

Yu J et al. [68] proposed a deep neural network PCACAE based on a convolutional autoencoder based on two-dimensional principal component analysis and designed a new convolution kernel for extracting wafer defect features. The product autoencoder is cascaded to further improve the performance of feature extraction. Aiming at the problems of difficult collection of wafer data and few public datasets, Ssu-Han Chen et al. [33] used a combination of the generative adversarial network and target detection algorithm YOLOv3 for the first time to detect wafer defects in small samples. The diversity of defects is enhanced by GAN, which improves the generalization ability of YOLOv3. Prashant P. SHINDE et al. [69] proposed the use of advanced YOLOv4 to detect and locate wafer defects. Compared with YOLOv3, the backbone extraction network was improved from Darknet-19 to Darknet-53, and the mish activation function was used to make the network robust. The stickiness is enhanced, the detection ability is greatly improved, and the detection and positioning performance of complex wafer defect modes is more efficient.

5.3. Segment Network

Segmentation networks perform pixel-level segmentation of regions of interest in input images [70]. Most of the segmentation network is based on the structure of the encoder and decoder, as shown in Figure 12 is a schematic diagram of the segmentation network model structure. Through the encoder and decoder, the ability to extract the features of the target object is improved, and the analysis and understanding of the image by the subsequent classification network are strengthened. It has a good prospect to be applied to wafer surface defect detection.

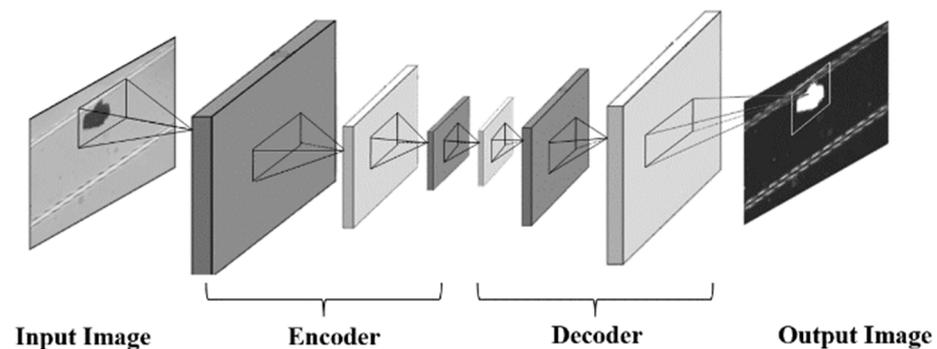


Figure 12. Schematic diagram of segmentation network model structure.

Takeshi Nakazawa et al. [70] proposed a deep convolutional encoder–decoder neural network structure for anomaly detection and segmentation of wafer defect patterns. The author designed three encoder–decoder wafer defect pattern segmentation networks based on FCN, U-Net, and SegNet to segment the wafer, local defect model. Global random defects in the wafer often lead to noise in the extracted features. After segmentation, the influence of global defects on local defects is ignored, and more information about defect clusters is helpful for further analysis of the reasons. Aiming at the problem of unbalanced wafer defect pixel categories and insufficient samples, Han Hui et al. [36] designed an improved segmentation system based on a U-net network. Based on the original UNet network, the RPN network is added to obtain defect area suggestions, and then input to the Unit network for segmentation. The designed two-stage network has an accurate segmentation effect on wafer defects. Subhrajit Nag [71] et al. proposed a new network structure WaferSegClassNet with a decoder–encoder architecture. The encoder extracts better multi-scale local details through a series of convolutional blocks, and the decoder is used to perform classification and generation. Segmentation mask, which is the first wafer defect detection model that can classify and segment at the same time, has a good segmentation and classification effect for mixed wafer defects. The segmentation network model comparison is shown in Table 5.

Table 5. Segmentation network model comparison.

| Algorithm | Innovation | Acc |
|-----------------------|--|-------|
| FCN [62] | Replacing fully connected layers with convolutional layers to output 2D heatmaps. | 97.8% |
| SegNet [62] | Combining encoder–decoder and pixel-level classification layers. | 99.0% |
| U-net [36] | Copy and crop the feature maps in each encoder layer to the corresponding decoder layer. | 98.9% |
| WaferSegClassNet [70] | Simultaneous classification and Segmentation using shared encoders. | 98.2% |

6. Conclusions and Outlook

With the continuous development of electronic information technology and the continuous improvement of lithography technology, wafer surface defect detection occupies an important position in the semiconductor industry, and has attracted increasing attention from scholars in this field. This paper analyzes and summarizes the research on image signal processing, machine learning, and deep learning related to wafer surface defect detection. In the early days, image signal processing methods were mainly used, among which wavelet transform methods and spatial filtering methods were widely used. Machine learning is very robust in wafer defect detection. Algorithms such as k-Nearest Neighbor (KNN), Decision Tree (Decision Tree), and Support Vector Machine (SVM) are widely used in this field and have achieved good results. Deep learning has brought vitality to the field of wafer inspection with its powerful feature extraction capabilities. The latest manufacturing technology for integrated circuits has advanced to 4 nm, and predictions suggest it will continue to develop towards even smaller scales. However, as these trends emerge, the complexity of surface defects on wafers will also increase, posing more stringent challenges to the reliability and robustness of models. Therefore, the analysis and processing of these defects become increasingly important to ensure high-quality manufacturing of integrated circuits. Although some achievements have been made in the field of wafer surface defect analysis, there are still many problems and challenges.

1. There are few public datasets of wafer defects. Due to the high cost of wafer production and labeling, there are very few high-quality public datasets, and the few datasets are not enough to support training. It is possible to consider creating a synthetic wafer defect database and performing data augmentation on the existing dataset to provide more accurate and comprehensive data samples for neural networks. Due to the versatility of defect types in gradient features, such problems can be addressed using transfer learning, mainly to solve problems such as negative transfer and model inappropriateness in transfer learning [72]. A flexible and efficient migration model does not currently exist. Using transfer learning to solve the problem of a few samples in wafer surface defect detection is a difficult topic for future research.
2. During the wafer fabrication process, new defects are continuously generated, and the number and types of defect samples are continuously accumulated. Using incremental learning [73] can improve the recognition accuracy of the network model for new defects and the ability to maintain the classification of old defects. It can also be used as a research direction for the expanded sample method.
3. With the rapid development of technological progress, the chip feature size is becoming smaller and more complex, resulting in multiple defect types in a wafer, and the defects are folded with each other, resulting in non-uniform and inconspicuous defect features. increase the difficulty of detection. Multi-step, multi-method hybrid models have become the mainstream method for detecting hybrid defects. How to optimize the performance of the deep network model and maintain a high detection efficiency is a problem that needs to be further solved.
4. During the wafer fabrication process, wafer patterns for different purposes will produce different defects. Currently, a network model trained on a single data set is

not sufficient to identify defects in all wafers for different purposes. How to design a universal network model to detect all defects, thereby avoiding the waste of resources caused by designing a training model separately for all wafer defect data sets, is a direction worth thinking about in the future.

5. The majority of defect detection models are offline models, which are unable to meet the real-time requirements of industrial production. To address this issue, an autonomous learning model system needs to be established, which enables the model to rapidly learn and adapt to new production environments, thereby achieving more efficient and accurate defect detection.

Author Contributions: Conceptualization, J.M.; writing—review and editing, T.Z.; investigation, C.Y.; resources, Y.C.; writing—review and editing, H.T.; supervision, L.X.; project administration, X.L. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by National Key Research and Development Program Key Special Project (2020YFB1712401-1), ZhengZhou Collaborative Innovation Major Project (20XTZX06013), China Postdoctoral Science Foundation (2021M692926), HeNan Science and Technology Research (222102310647) and Strategic Research and Consulting Project of Chinese Academy of Engineering (2022HENYB03).

Data Availability Statement: Not applicable.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Gao, Y.; Li, X.; Wang, X.V.; Wang, L.; Gao, L. A review on recent advances in vision-based defect recognition towards industrial intelligence. *J. Manuf. Syst.* **2021**, *62*, 753–766. [\[CrossRef\]](#)
2. Zhu, Y.; Ling, Z.; Zhang, Y. Research progress and Prospect of machine vision technology. *J. Graph.* **2020**, *41*, 871–890.
3. Kim, T.; Behdinin, K. Advances in Machine Learning and Deep Learning Applications towards Wafer Map Defect Recognition and Classification: A Review. *J. Intell. Manuf.* **2022**, *33*, 1805–1826. [\[CrossRef\]](#)
4. Thomas, N.; Wong, H.S.P. The End of Moore’s Law: A New Beginning for Information Technology. *Comput. Sci. Eng.* **2017**, *19*, 41–50.
5. Zhu, J.; Liu, J.; Xu, T.; Yuan, S.; Zhang, Z.; Jiang, H.; Gu, H.; Zhou, R.; Liu, S. Optical Wafer Defect Inspection at the 10 nm Technology Node and Beyond. *Int. J. Extrem. Manuf.* **2022**, *12*, 2631–8644. [\[CrossRef\]](#)
6. McLaughlin, M.P.; Mennell, P.; Stamper, A.; Barber, G.; Paduano, J.; Benn, E.; Linnane, M.; Zwick, J.; Khatumria, C.; Isaacson, R.L.; et al. Improved Color Defect Detection with Machine Learning for After Develop Inspections in Lithography. *IEEE Trans. Semicond. Manuf.* **2022**, *35*, 418–424. [\[CrossRef\]](#)
7. Chen, S.H. Research on Machine Learning and Intelligent Classification and Identification of Wafer Defect Patterns in Microelectronics Manufacturing. Ph.D. Thesis, Jiangsu University, Zhenjiang, China, 2021.
8. Hansen, M.H.; Nair, V.N.; Friedman, D.J. Monitoring wafer map data from integrated circuit fabrication processes for spatially clustered defects. *Technometrics* **1997**, *39*, 241–253. [\[CrossRef\]](#)
9. Jeong, Y.-S.; Kim, S.J.M. Automatic Identification of Defect Patterns in Semiconductor Wafer Maps Using Spatial Correlogram and Dynamic Time Warping. *IEEE Trans. Semicond. Manuf.* **2008**, *21*, 625–637. [\[CrossRef\]](#)
10. Mirza, A.I.; O’Donoghue, G.; Drake, A.W.; Graves, S.C. Spatial yield modeling for semiconductor wafers. In Proceedings of the SEMI Advanced Semiconductor Manufacturing Conference and Workshop, Cambridge, MA, USA, 13–15 November 1995; pp. 276–281.
11. Wu, M.J.; Jang, J.S.R.; Chen, J.L. Wafer Map Failure Pattern Recognition and Similarity Ranking for Large-Scale Data Sets. *IEEE Trans. Semicond. Manuf.* **2014**, *28*, 1–12.
12. Kim, Y.; Cho, D.; Lee, J.H. Wafer map classifier using deep learning for detecting out-of-distribution failure patterns. In Proceedings of the 2020 IEEE International Symposium on the Physical and Failure Analysis of Integrated Circuits (IPFA), Singapore, 20–23 July 2020; pp. 1–5.
13. Kang, S.; Cho, S.; An, D.; Rim, J. Using wafer map features to better predict die-level failures in final test. *IEEE Trans. Semicond. Manuf.* **2015**, *28*, 431–437. [\[CrossRef\]](#)
14. Saqlain, M.; Jargalsaikhan, B.; Lee, J.Y. A voting ensemble classifier for wafer map defect patterns identification in semiconductor manufacturing. *IEEE Trans. Semicond. Manuf.* **2019**, *32*, 171–182. [\[CrossRef\]](#)
15. Shim, J.; Kang, S.; Cho, S. Active learning of convolutional neural network for cost-effective wafer map pattern classification. *IEEE Trans. Semicond. Manuf.* **2020**, *33*, 258–266. [\[CrossRef\]](#)
16. Feng, L. Research on Wafer Defects in Semiconductor Lithography. Master’s Thesis, Tianjin University, Tianjin, China, 2021.

17. Caiazzo, F.; Gao, R.; Teti, R. Machine learning-based image processing for on-line defect recognition in additive manufacturing. *CIRP Ann.* **2019**, *68*, 451–454.
18. Zhang, H. Research on the Optimizing Process of the Basic Image Processing Algorithms. In Proceedings of the International Conference on Cyber Security Intelligence and Analytics, Shenyang, China, 21–22 February 2019; pp. 212–217.
19. Rajeswari, J.; Jagannath, M. Advances in biomedical signal and image processing—A systematic review. *Inform. Med. Unlocked* **2017**, *8*, 13–19.
20. Guo, T.; Zhang, T.; Lim, E.; Lopez-Benitez, M.; Ma, F.; Yu, L. A Review of Wavelet Analysis and Its Applications: Challenges and Opportunities. *IEEE Access* **2022**, *10*, 58869–58903. [[CrossRef](#)]
21. Wang, F.; Liu, Q.; Hu, D.; Wang, Y.; Zheng, T. Research of the image restoration algorithm based on boundary pre-processing in the space domain and fast computing in the frequency domain. In Proceedings of the Fifth Symposium on Novel Optoelectronic Detection Technology and Application, Xi'an, China, 12 March 2019; Volume 11023, pp. 143–154.
22. Prasath, V.B.S.; Thanh, D.N.H.; Hung, N.Q.; Hieu, L.M. Multiscale gradient maps augmented fisher information-based image edge detection. *IEEE Access* **2020**, *8*, 141104–141110. [[CrossRef](#)]
23. Yeh, C.H.; Wu, F.C.; Ji, W.L.; Huang, C.Y. A wavelet-based approach in detecting visual defects on semiconductor wafer dies. *IEEE Trans. Semicond. Manuf.* **2010**, *23*, 284–292. [[CrossRef](#)]
24. Yang, W.R. Short-time discrete wavelet transform for wafer microcrack detection. In Proceedings of the 2009 IEEE International Symposium on Industrial Electronics, Seoul, Republic of Korea, 5–8 July 2009.
25. Villa, J.; González, E.; Moreno, G.; de la Rosa, I.; Flores, J.L.; Alaniz, D. Fringe pattern denoising using spatial oriented gaussian filters. *Opt. Commun.* **2020**, *457*, 124704. [[CrossRef](#)]
26. Allen, J.; Brewster, C.; Slone, D. Spatially explicit ecological models: A spatial convolution approach. *Chaos Solitons Fractals* **2001**, *12*, 333–347. [[CrossRef](#)]
27. Kuan, D.T.; Sawchuk, A.A.; Strand, T.C.; Chavel, P. Adaptive noise smoothing filter for images with signal-dependent noise. *IEEE Trans. Pattern Anal. Mach. Intell.* **1985**, *221*, 165–177. [[CrossRef](#)]
28. Zhang, B.; Allebach, J.P. Adaptive bilateral filter for sharpness enhancement and noise removal. *IEEE Trans. Image Process.* **2008**, *17*, 664–678. [[CrossRef](#)] [[PubMed](#)]
29. Ohshige, T.; Tanaka, H.; Miyazaki, Y.; Kanda, T.; Ichimura, H.; Kosaka, N.; Tomoda, T. Defect inspection system for patterned wafers based on the spatial-frequency filtering. In Proceedings of the 1991 Proceedings Eleventh IEEE/CHMT International Electronics Manufacturing Technology Symposium, San Francisco, CA, USA, 16–18 September 1991.
30. Wang, C.H. Recognition of semiconductor defect patterns using spatial filtering and spectral clustering. *Expert Syst. Appl.* **2008**, *34*, 1914–1923. [[CrossRef](#)]
31. Zarinbal, M.; Zarandi, M.F.; Turksen, I. Relative entropy fuzzy c-means clustering. *Inf. Sci.* **2014**, *260*, 74–79. [[CrossRef](#)]
32. Jia, H.; Ding, S.; Xu, X.; Nie, R. The latest research progress on spectral clustering. *Neural. Comput. Appl.* **2014**, *24*, 1477–1486. [[CrossRef](#)]
33. Chen, S.H.; Kang, C.H.; Perng, D.B. Detecting and measuring defects in wafer die using gan and yolov3. *Appl. Sci.* **2020**, *10*, 8725. [[CrossRef](#)]
34. Briechle, K.; Hanebeck, U.D. Template matching using fast normalized cross correlation. In *Optical Pattern Recognition XII*; SPIE: Bellingham, WA, USA, 2001; Volume 4387, pp. 95–102.
35. Brunelli, R. *Template Matching Techniques in Computer Vision: Theory and Practice*; John Wiley & Sons: Hoboken, NJ, USA, 2009.
36. Han, H.; Gao, C.; Zhao, Y.; Liao, S.; Tang, L.; Li, X. Polycrystalline silicon wafer defect segmentation based on deep convolutional neural networks. *Pattern Recognit. Lett.* **2020**, *130*, 234–241. [[CrossRef](#)]
37. Khalaj, B.H.; Aghajan, H.K.; Kailath, T. Digital image processing techniques for patterned-wafer inspection. In Proceedings of the SPIE'S 1993 Symposium on Microlithography, San Jose, CA, USA, 4 August 1993; Volume 1926, pp. 508–516.
38. Kotsiantis, S.B.; Zaharakis, I.; Pintelas, P. Supervised machine learning: A review of classification techniques. *Emerg. Artif. Intell. Appl. Comput. Eng.* **2007**, *160*, 3–24.
39. Kotsiantis, S.B.; Zaharakis, I.D.; Pintelas, P.E. Machine learning: A review of classification and combining techniques. *Artif. Intell. Rev.* **2006**, *26*, 159–190. [[CrossRef](#)]
40. Yuan, T.; Kuo, W.; Bae, S.J. Detection of spatial defect patterns generated in semiconductor fabrication processes. *IEEE Trans. Semicond. Manuf.* **2011**, *24*, 392–403. [[CrossRef](#)]
41. Piao, M.; Jin, C.H.; Lee, J.Y.; Byun, J.Y. Decision tree ensemble-based wafer map failure pattern recognition based on radon transform-based features. *IEEE Trans. Semicond. Manuf.* **2018**, *31*, 250–257. [[CrossRef](#)]
42. Burges, C.J. A tutorial on support vector machines for pattern recognition. *Data Min. Knowl. Discov.* **1998**, *2*, 121–167. [[CrossRef](#)]
43. Xie, L.; Li, D.; Simske, S.J. Feature dimensionality reduction for example-based image super-resolution. *J. Pattern Recognit. Res.* **2011**, *6*, 130–139. [[CrossRef](#)] [[PubMed](#)]
44. Song, H.; Choi, K.K.; Lee, I.; Zhao, L.; Lamb, D. Adaptive virtual support vector machine for reliability analysis of high-dimensional problems. *Struct. Multidiscip. Optim.* **2013**, *47*, 479–491. [[CrossRef](#)]
45. Baly, R.; Hajj, H. Wafer classification using support vector machines. *IEEE Trans. Semicond. Manuf.* **2012**, *25*, 373–383. [[CrossRef](#)]
46. Xie, L.; Huang, R.; Gu, N.; Cao, Z. A novel defect detection and identification method in optical inspection. *Neural Comput. Appl.* **2014**, *24*, 1953–1962. [[CrossRef](#)]
47. Barlow, H.B. Unsupervised learning. *Neural Comput.* **1989**, *1*, 295–311. [[CrossRef](#)]

48. Huang, C.J. Clustered defect detection of high quality chips using self-supervised multilayer perceptron. *Expert Syst. Appl.* **2007**, *33*, 996–1003. [CrossRef]
49. Jin, C.H.; Na, H.J.; Piao, M.; Pok, G.; Ryu, K.H. A novel DBSCAN-based defect pattern detection and classification framework for wafer bin map. *IEEE Trans. Semicond. Manuf.* **2019**, *32*, 286–292. [CrossRef]
50. Di Palma, F.; De Nicolao, G.; Miraglia, G.; Pasquinetti, E.; Piccinini, F. Unsupervised spatial pattern classification of electrical-wafer-sorting maps in semiconductor manufacturing. *Pattern Recognit. Lett.* **2005**, *26*, 1857–1865. [CrossRef]
51. Xiaoxi, W.; Wenjun, W.; Feng, Y.; Pengbo, S.; Xuanyi, Z.; Yanhua, Z. Pseudo-label based semi-supervised learning in the distributed machine learning framework. *High Technol. Lett.* **2022**, *28*, 172–180.
52. Li, K.S.M.; Jiang, X.H.; Chen, L.L.Y.; Wang, S.J.; Huang, A.Y.A.; Chen, J.E.; Liang, H.C.; Hsu, C.L. Wafer Defect Pattern Labeling and Recognition Using Semi-Supervised Learning. *IEEE Trans. Semicond. Manuf.* **2022**, *35*, 291–299. [CrossRef]
53. Kong, Y.; Ni, D. A semi-supervised and incremental modeling framework for wafer map classification. *IEEE Trans. Semicond. Manuf.* **2020**, *33*, 62–71. [CrossRef]
54. Albawi, S.; Mohammed, T.A.; Al-Zawi, S. Understanding of a convolutional neural network. In Proceedings of the 2017 International Conference on Engineering and Technology (ICET), Antalya, Turkey, 21–23 August 2017; pp. 1–6.
55. Cheng, K.C.C.; Chen, L.L.Y.; Li, J.W.; Li, K.S.M.; Tsai, N.C.Y.; Wang, S.J.; Huang, A.Y.A.; Chou, L.; Lee, C.S.; Chen, J.E.; et al. Machine learning-based detection method for wafer test induced defects. *IEEE Trans. Semicond. Manuf.* **2021**, *34*, 161–167. [CrossRef]
56. Wright, R.E. Logistic Regression. 1995. Available online: <https://psycnet.apa.org/record/1995-97110-007> (accessed on 5 February 2023).
57. Chen, Y.; Zhou, X.S.; Huang, T.S. One-class SVM for learning in image retrieval. In Proceedings of the 2001 International Conference on Image Processing (Cat. No. 01CH37205), Thessaloniki, Greece, 7–10 October 2001; Volume 1, pp. 34–37.
58. Song, Y.Y.; Ying, L. Decision tree methods: Applications for classification and prediction. *Shanghai Arch. Psychiatry* **2015**, *27*, 130. [PubMed]
59. Ma, W.; Lu, J. An equivalence of fully connected layer and convolutional layer. *arXiv* **2017**, arXiv:1712.01252 2017.
60. Wang, J.; Xu, C.; Yang, Z.; Zhang, J.; Li, X. Deformable convolutional networks for efficient mixed-type wafer defect pattern recognition. *IEEE Trans. Semicond. Manuf.* **2020**, *33*, 587–596. [CrossRef]
61. Kyeong, K.; Kim, H. Classification of mixed-type defect patterns in wafer bin maps using convolutional neural networks. *IEEE Trans. Semicond. Manuf.* **2018**, *31*, 395–402. [CrossRef]
62. Nakazawa, T.; Kulkarni, D.V. Wafer map defect pattern classification and image retrieval using convolutional neural network. *IEEE Trans. Semicond. Manuf.* **2018**, *31*, 309–314. [CrossRef]
63. Girshick, R.; Donahue, J.; Darrell, T.; Malik, J. Rich feature hierarchies for accurate object detection and semantic segmentation. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Washington, DC, USA, 23–28 June 2014; pp. 580–587.
64. Ren, S.; He, K.; Girshick, R.; Sun, J. Faster r-cnn: Towards real-time object detection with region proposal networks. *Adv. Neural Inf. Process. Syst.* **2015**, *39*, 1137–1149. [CrossRef]
65. Redmon, J.; Divvala, S.; Girshick, R.; Farhadi, A. You only look once: Unified, real-time object detection. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Las Vegas, NV, USA, 27–30 June 2016; pp. 779–788.
66. Redmon, J.; Farhadi, A. YOLO9000: Better, faster, stronger. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Honolulu, HI, USA, 21–26 June 2017; pp. 7263–7271.
67. Redmon, J.; Farhadi, A. Yolov3: An incremental improvement. *arXiv* **2018**, arXiv:1804.02767 2018.
68. Yu, J.; Liu, J. Two-dimensional principal component analysis-based convolutional autoencoder for wafer map defect detection. *IEEE Trans. Ind. Electron.* **2020**, *68*, 8789–8797. [CrossRef]
69. Shinde, P.P.; Pai, P.P.; Adiga, S.P. Wafer Defect Localization and Classification Using Deep Learning Techniques. *IEEE Access* **2022**, *10*, 39969–39974. [CrossRef]
70. Nakazawa, T.; Kulkarni, D.V. Anomaly detection and segmentation for wafer defect patterns using deep convolutional encoder-decoder neural network architectures in semiconductor manufacturing. *IEEE Trans. Semicond. Manuf.* **2019**, *32*, 250–256. [CrossRef]
71. Nag, S.; Makwana, D.; Mittal, S.; Mohan, C.K. A light-weight network for classification and segmentation of semiconductor wafer defects. *Comput. Industry* **2022**, *142*, 103720. [CrossRef]
72. Torrey, L.; Shavlik, J. *Handbook of Research on Machine Learning Applications and Trends: Algorithms, Methods, and Techniques*; Information Science Reference: Hershey, PA, USA, 2010.
73. Castro, F.M.; Marin-Jiménez, M.J.; Guil, N.; Schmid, C.; Alahari, K. End-to-end incremental learning. In Proceedings of the European Conference on Computer Vision (ECCV), Munich, Germany, 8–14 October 2018; pp. 233–248.

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.