Hindawi
International Journal of Intelligent Systems
Volume 2023, Article ID 4096164, 13 pages
https://doi.org/10.1155/2023/4096164

Research Article

LED Chip Defect Detection Method Based on a Hybrid Algorithm

Pengfei Zheng,[1,2,3] **Jingjing Lou**,[1] **Xiyuan Wan**,[1] **Qingdong Luo**,[1] **Yunhan Li**,[4] **Linsheng Xie**,[2] **and Zegang Zhu**[5]

[1]*School of Mechanical Information, Yiwu Industrial & Commercial College, Yiwu 322000, China*
[2]*School of Mechanical and Power Engineering, East China University of Science and Technology, Shanghai 200237, China*
[3]*Zhejiang Provincial Key Laboratory of Third Generation Semiconductor Materials and Devices, Yiwu 322000, China*
[4]*School of Intelligent Manufacturing, Huzhou College, Huzhou 313000, China*
[5]*Jinhua Academy of Agricultural Sciences, Jinhua 321000, China*

Correspondence should be addressed to Pengfei Zheng; pfzheng@126.com

Received 2 October 2022; Revised 5 February 2023; Accepted 9 February 2023; Published 21 February 2023

Academic Editor: B. B. Gupta

Copyright © 2023 Pengfei Zheng et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

LED is an extremely important energy-saving lighting products, which has greatly facilitated human life. Meanwhile, it also makes a positive contribution to global carbon neutrality and carbon peaking. Defect detection is a vital part of the production process to control the quality of LED chips. The traditional methods use a microscope for manual visual inspection, which is time-consuming and has inconsistent testing standards, low efficiency, and other deficiencies. To solve these problems, a hybrid algorithm based on geometric computation and a convolutional neural network is proposed for LED chip defect detection. The method takes advantage of the dimensionality reduction of geometric computation to perform coarse detection of defects on preprocessed chip lithography graphs in the form of grid segmentation, which realizes fast coarse screening of large-scale chip samples and reduces postcomputational costs. The convolutional neural network model is used for the secondary fine detection of "suspected defective" chips after geometric coarse screening, and the SPP (spatial pyramid pooling) network model is improved by directly introducing the original feature map into the SPP pooling layer for summation to enhance the global and local feature information of the output feature map. Furthermore, we construct an LED chip image acquisition platform using a high-frequency multi-magnification zoom lens, collect training samples of defective chips, and increase the number of samples through image processing techniques. The research introduces the R-CNN, SDD, and YOLO methods to evaluate the superiority of our method in a number of experiments. The experimental results show that our algorithm proposed in this paper has an average precision (*AP*) of 96.7% for large-scale chip detection with a low defect rate. Compared with other methods, the testing mean average precision (*mAP*) is 10.39% higher than traditional YOLO v2. The testing *mIoU* is also 3.63% higher than traditional YOLO v5, the detection speed is also significantly improved, and it has good robustness.

1. Introduction

The light-emitting diode (LED) is a commonly used light-emitting device that releases energy to emit light through the compound of electrons and holes and is extremely widely used in the lighting field due to its small structure, low energy consumption, and long life. According to statistics, the scale of the global LED lighting industry in 2021 will exceed 100 billion U.S. dollars. Analogously, China's LED chip market scale has grown from 8.25 billion RMB in 2013 to nearly 30 billion RMB in 2021, with huge development space. LED chip manufacturing is divided into wafer processing, wafer pin testing, packaging, testing, and other processes: wafer processing and pin testing for the front process, packaging and testing for the back process. Among them, wafer processing refers to the production of circuits and electronic components (such as transistors, capacitors, logic switches, etc.) on wafers, the basic steps of which are cleaning, chemical vapor deposition, coating, exposure, development, lithography, etc. As LED chips are generally lithographically stacked through multiple layers of circuits and then protected by the packaging process and connected

to external circuits, the integrity of each layer of lithography will be critical to the luminous performance of the chip. However, during the lithography process, some defects will inevitably be generated on the chip surface or outside, and these defects will lead to the degradation of semiconductor components and integrated circuits or even damage. Therefore, detecting defects in chips not only greatly improves performance, but the feedback data can also be used to analyze the failures of the production line and improve the yield rate. At present, in enterprises with monthly production of more than 80 k chips, which commonly use the manual visual inspection method for quality control, repetitious examination under strong lights and microscopes may cause persistence of vision or dizziness, and examiners are susceptible to environmental, emotional, and other factors, so that the accuracy rate of human inspection may be no more than 80%–85%, and the standard may not be uniform. Consequently, there is an urgent need to study the LED chip defect detection method based on machine vision theory.

In recent years, scholars have conducted extensive research on product defect detection using machine vision technology, especially for PCBs, chips, and other products. Thomas and Gehrig [1] designed a multifunctional tool for object detection and localization in microscopes based on the multitemplate matching method. Hu and Wang [2] proposed a deep learning-based image detection method for PCB defect detection based on faster R-CNN. Huang et al. [3] adopted a reference-based method to inspect and train an end-to-end convolutional neural network to classify the defects, which is referred to as the RBCNN approach. Liu and Qu [4] proposed a key technology for PCB defect online detection based on a visual detection algorithm, image smoothing, image contrast enhancement, image sharpening, and other preprocessing operations. Liao et al. [5] improved the backbone network and the activation function in the neck/prediction network and proposed a cost-effective deep learning-based detector based on the cutting-edge YOLO v4 to detect PCB surface defects quickly and efficiently. Li et al. [6] proposed an effective self-adaption method that collected "exception data," retrained with exceptions on the server, and deployed back to the edge. An et al. [7] proposed a transformer-based model, LPViT, for defect detection and classification of PCBs. Ling et al. [8] developed a novel deep Siamese semantic segmentation network which integrated the similarity measurement of the Siamese network with the encoder-decoder semantic segmentation network for PCB welding defect detection. Jeon et al. [9] proposed a contactless inspection method that can detect PCBA defects without the use of a fixture or ICT by using the comparison of thermal images and deep learning analysis. Jiang et al. [10] proposed a data-driven spike detection method using multitemplate matching, feature extraction, and a threshold method to improve the detection performance under different circumstances. Chen et al. [11] presented a machine vision method for auto-detecting defects embedded on LED substrates. Lin et al. [12] proposed a method for LED chip defect inspection based on a deep convolutional neural network named LEDNet. Shu et al. [13] proposed a parallel deep convolution model and a parallel spatial pyramid pooling network (PSPP-net) for LED chip surface quality detection. Stern and Schellenberger [14] proposed a full convolutional network for chip defect detection and applied it to a photoluminescence image dataset for training. Lin et al. [15] proposed a defect contour detection method for complex structural chips based on region segmentation search. Otherwise, YOLO (You Only Look Once) has been gradually introduced as an efficient target detection algorithm, into the field of machine vision inspection. Chen and Tsai [16] proposed a dense YOLO v3 variant algorithm for detecting defects such as incorrect placement, reverse polarity, and missing wires in LED placement assemblies by using Darknet-53 as the prime grid and improving the YOLO v3 network model. Chen et al. [17] developed a new end-to-end framework that can detect multiple hotspots in a large region at a time and promises better hotspot detection performance. Xu et al. [18] improved the normalized correlation coefficient matching algorithm and proposed an LED chip defect detection method based on position preestimation and feature enhancement. Analogously, some scholars have conducted in-depth research on segmentation and detection of crack damage, which are advanced levels of damage identification. Kang et al. [19, 20] improved the tubularity flow field (TuFF) algorithm and distance transform method (DTM) to propose an automatic crack detection, localization, and quantification method. A novel semantic transformer representation network (STRNet) is developed for crack segmentation at the pixel level in complex scenes in a real-time manner. For the segmentation problem of internal damages of concrete members, Ali and Cha [21] developed an attention-based generative adversarial network (AGAN) to generate synthetic images for training the proposed internal damage segmentation network (IDSNet). Cha et al. [22, 23] proposed a vision-based method using a deep architecture of convolutional neural networks (CNNs) for detecting concrete cracks without calculating the defect features and a structural vision inspection method based on a fast regional convolutional neural network (Faster R-CNN) for the simultaneous detection of multiple types of damage in quasi-real time. Wang and Cha [24] proposed an unsupervised deep learning method based on a deep auto-encoder with a one-class support vector machine to detect damage in civil infrastructures. Additionally, other scholars [25–36] have applied vision inspection to many fields, such as machining, loosened bolts, automatic volumetric damage quantification, and food processing, to promote the wider application of machine vision technology.

It is not difficult to find that neural networks, machine learning, deep learning, and other artificial intelligence algorithms have been commonly used in the field of chip detection, and better detection results have been achieved. However, its application to LED chip defect detection still has some shortcomings: Compared with traditional detection algorithms, the R-CNN algorithm has made great progress in terms of accuracy, but its drawback is also very obvious: the redundant computation of overlapping frame features makes the detection of the whole network slow.

Although the Fast R-CNN algorithm has further improved in detection accuracy, it still uses a selective search algorithm to find the region of interest, a process that is usually slow. The faster R-CNN still suffers from computational redundancy, and the improper selection of IOU threshold is prone to noise detection or overfitting problems. The SDD algorithm has better detection accuracy and speed than the YOLO v1 algorithm. The YOLO v2 algorithm has only one detection branch, and the network lacks the ability to capture multiscale information, which is still less effective for detecting targets of different sizes. YOLO v3, YOLO v4, and YOLO v5 are not accurate in target localization, although they have a significant improvement in detection speed. In general, the deeper convolutional layers lead to the low detection rates of SSD and YOLO algorithms for small targets, and they suffer from deficiencies such as insufficient localization accuracy. The "black box" characteristics of neural networks easily lead to unpredictable detection results; machine learning in the process of constructing a decision tree requires a large number of data scans and sorting, which is computationally intensive and leads to a reduction in the effectiveness of the algorithm. The efficient ability of deep learning relies on the training of large data samples, but the large number of defective LED chip samples is difficult to collect, while the network model is more complex. Accordingly, this paper proposes a hybrid algorithm combining graphical geometric computation and a convolutional neural network model, which takes into account algorithm efficiency and detection accuracy and is applied to multi-type LED chip defect detection.

The remainder of this study is organized as follows: Section 2 builds a high-frequency multimagnification LED chip image acquisition platform. In Section 3, we discuss the LED chip image preprocessing methods in terms of morphological processing and chip feature extraction. In Section 4, the architecture of the chip defect detection hybrid algorithm is described, and then the localization detection algorithm based on geometric calculation and the CNN-based defect detection model are presented. Section 5 shows the experimental results and a comparison analysis. Finally, Section 6 involves the conclusion of the studies. The flow of the work is shown in Figure 1.

2. Image Acquisition Platform Construction

The circuit fabrication of LED chips is mainly done in the wafer processing process, which creates circuits and electronic components on wafers in the form of photolithographic images. Wafer size is small, generally, there are different specifications, such as 4, 6, 8 inches, etc. This work mainly focuses on 4-inch wafers (diameter of 101.6 mm, thickness of 1.0 mm wafer) as the object of study. Due to the tiny chip size, a 4-inch wafer contains 100,000–400,000 LED chips, so it is necessary to design and build a high magnification function for the image acquisition platform. According to the detection work scene, we choose the appropriate industrial camera and lens after comprehensive consideration of resolution, depth of field, aberration, environmental interference, and other factors to determine the camera calibration and focal length, relative aperture (aperture coefficient), field of view angle, and other parameters.

As shown in Figure 2, the image acquisition platform uses a high-performance CCD camera with a resolution of 1920×1080, 2 million pixels, and a frame rate of 30 f/s. It is equipped with a high-frequency multimagnification zoom lens with a focal length of 10–15 mm and a zoom magnification of 250–2000 times, as well as an adjustable high-brightness LED coaxial light source.

3. Image Preprocessing

3.1. Morphological Processing. Before chip defect detection, morphological processing methods such as image filtering or open and closed operations are commonly used to reduce noise on the source image in order to reduce the impact of external and internal noise on the image quality. The common image filtering functions are median filtering, mean filtering, and Gaussian filtering, as shown in Figure 3. In order to better highlight the chip circuit features and retain the chip edge information, Gaussian filtering is selected as the preprocessing method for chip images based on the comparison of various filtering results.

3.2. Chip Feature Extraction. The reason why the chip manufacturing technology is called "neck" technology is largely due to the complexity of the chip lithography process. A chip is usually made up of multiple layers of lithographic overlays, each with different circuits and different functions and using different film materials. An LED chip generally has several or a dozen layers of lithographic graphs, and the thickness of each layer is on the nanometer scale. As a result, it can be detected as a two-dimensional target.

Although the chip lithography process is exceptionally complex, the difficulty is greatly reduced when each layer of the lithography graph is used as the object of study. As shown in Figure 4, the lithographic features of each layer are not complicated from the perspective of geometry, and its P and N poles are common curve types, and the electrodes are generally circular in character.

In order to extract the chip lithography features more accurately, the chip source image needs to be segmented between foreground and background to fully retain the chip lithography and separate the background objects to reduce the interference of later recognition. Since chip lithography graphics have the good property of being single-color, this feature can be used to extract the image foreground. On this basis, the chip lithography image is binarized to facilitate the edge detection of subsequent lithography graphic features. The commonly used edge detection operators are: Sobel, Roberts, Prewitt, Canny, Laplacian operator, etc. According to the chip features and the actual testing results, the Canny operator is selected in this study to extract its edge features, as shown in Figure 5.

4. Hybrid Algorithm Architecture

Currently, the defect rate of the large-scale LED chip lithography process is low, typically only 100–500 ppm.

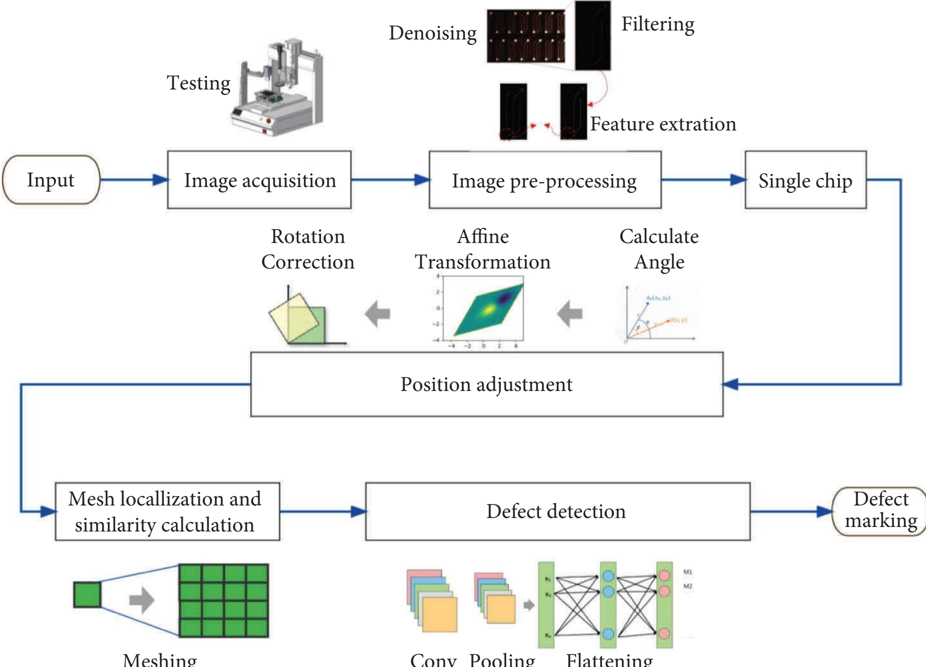

Figure 1: The flow of this work.

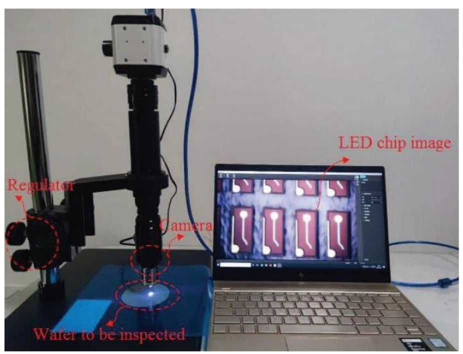

Figure 2: Chip image acquisition experimental platform.

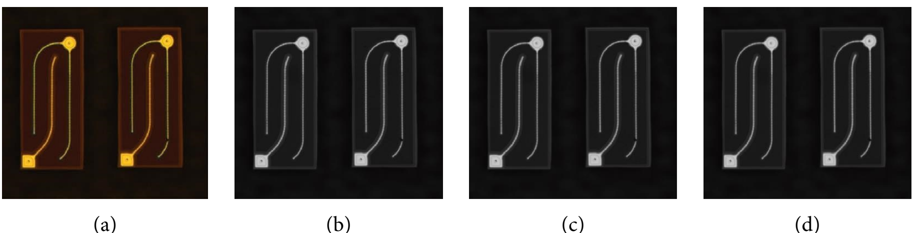

Figure 3: Comparison of chip image filtering results. (a) Original image, (b) median filtering, (c) mean filtering, and (d) Gaussian filtering.

Therefore, the adoption of an undifferentiated inspection strategy would result in excessive computational costs, which is clearly inappropriate. To locate and detect multiple types of LED chip defects, we propose a hybrid algorithm that combines geometric computation with a convolutional neural network model. The method framework is composed of a geometrically computed chip substrate localization model, a mesh segmentation model, a geometrically computed coarse inspection model, and a detection network model based on a hybrid of improved SPP and CNN. This section summarizes the entire process of our framework, and an overall schematic view of the proposed hybrid method is presented in Figure 6. In order to make full use of the respective advantages of geometric computing and deep

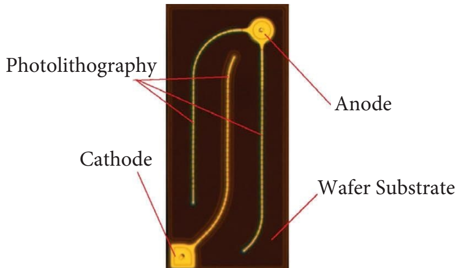

Figure 4: Chip lithography image.

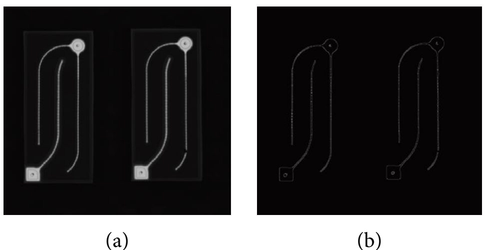

Figure 5: Edge detection results of chip lithography graphs. (a) After image preprocessing. (b) Edge detection result.

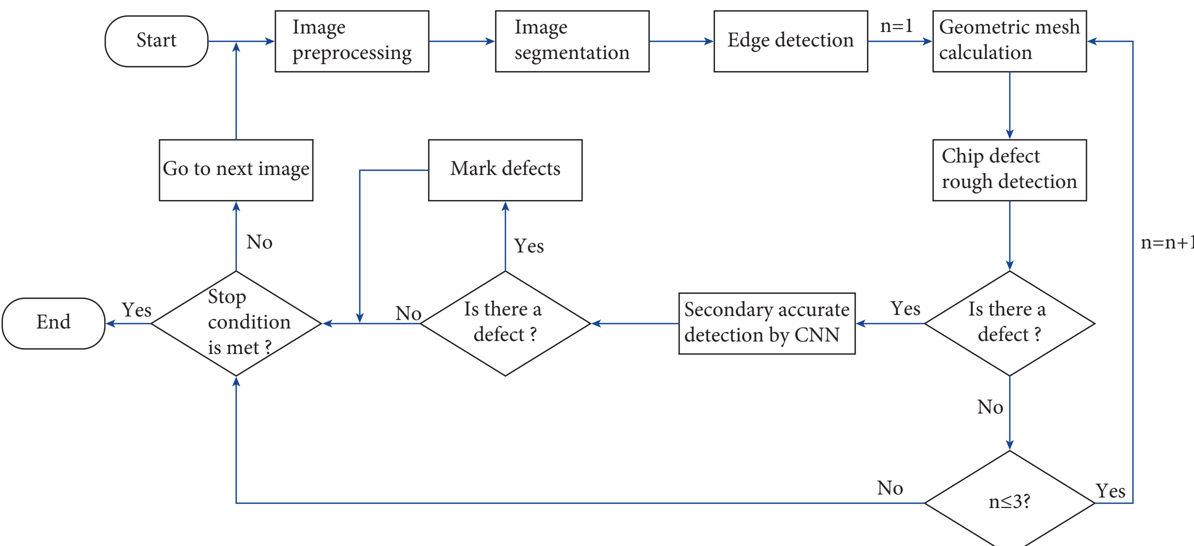

Figure 6: The architecture of chip defect detection hybrid algorithm.

learning in multi-type LED chip defect detection, initially, we designed a chip substrate localization model and a mesh segmentation model for fast localization and coarse detection of LED chip defects by geometric computing. Additionally, we improved the traditional SPP structure and CNN network structure to extract feature values of various typical defects for secondary accurate detection of LED chips suspected of having defects. Eventually, the sampled defect samples were expanded by image processing techniques, and the generated training samples were used for training this method for LED chip defect detection. The details of the developed modules and their roles are described in the following subsections.

4.1. Localization Detection Algorithm Based on Geometric Calculation. Prior to defect detection, the chip to be inspected must be positioned. Usually, chip substrates are rectangular and evenly and regularly distributed on the wafer, and the spacing between each chip is also equal, so only one chip needs to be positioned in the chip positioning calculation, and the adjacent chips can be moved and positioned by its uniform layout. Since the color difference between the chip and the wafer background is obvious, the tricolor value of the pixel point can be used to extract the chip area, as shown in Figure 7, and A is the area range of a chip. If the set of pixel points in region A is: $\{p_{i,j}\}$, it is easy to find the maximum and smallest value of x, y direction coordinates in this point set is:

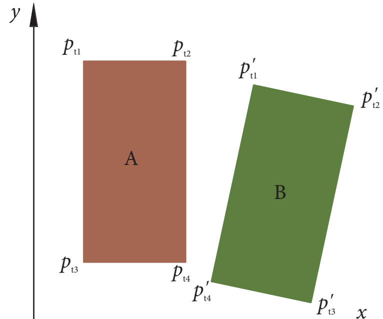

Figure 7: Position and calculation model of chip substrate.

$(p_{A(x\max)}, p_{A(x\min)}, p_{A(y\max)}, p_{A(y\min)})$. Then the coordinates of the four corner points of chip A can be obtained: p_{t1} $(p_{A(x\min)}, p_{A(y\max)})$, $p_{t2}(p_{A(x\max)}, p_{A(y\max)})$, $p_{t3}(p_{A(x\min)}, p_{A(y\min)})$, $p_{t4}(p_{A(x\max)}, p_{A(y\min)})$, and the chip corner points are the important basis for its positioning. If the chip position is not flush, as with chip B shown in Figure 7, we can still use the pixel point coordinates to find its four corner points and then make a radial transformation for position correction.

Since the lithography graph of each chip only accounts for a small part of the single chip area, the adoption of nondifferentiated search algorithm and constant weight calculation method will inevitably lead to a large amount of invalid calculation costs. Therefore, an optimized geometric calculation method is used to implement a strategy that combines coarse inspection of large size screening with fine inspection of the target area. As shown in Figure 8, the area occupied by a single chip is *abcd*, which is regarded as an image, and a grid segmentation is adopted to coarsely screen the whole image in a large size. It is easy to see that when using 2×2 grid partitioning, each subgrid has a target object, and when using 3×8 grid partitioning, some empty set grids can be delineated. Thus, the iterative model with grid partitioning can be used to achieve fast target selection and localization while reducing the computational effort to a large extent. However, grid segmentation only serves to quickly distinguish the target and blank areas and does not have the capability of target detection. As seen in Figure 8, a complete lithographic is divided into 24 subgraphs by a 3×8 grid, which causes a loss of image continuity, but the process actually transforms a linear problem into a discrete one. It is easy to imagine that if the grid is iteratively subdivided, a series of lattices containing only a few pixels can be obtained, which essentially transforms the curve calculation into a point calculation, achieving the purpose of geometric dimensionality reduction. As shown in Figure 9, a unit is used as an example to briefly introduce the target coarse inspection process. If $C_{i,j}$ is a subdivision unit of a chip containing the corresponding pixel point, and $T_{i,j}$ is the unit of its corresponding standard template, then the probability value of defect for unit $C_{i,j}$ is $\eta_{i,j}$. After testing different chip lithography graphics, the range of $\eta_{i,j}$ is set to 0.95~1.05 in this paper. It means, the image is considered no defects when it satisfied $0.95\le\eta_{i,j}\le1.05$; otherwise, it is transferred to the

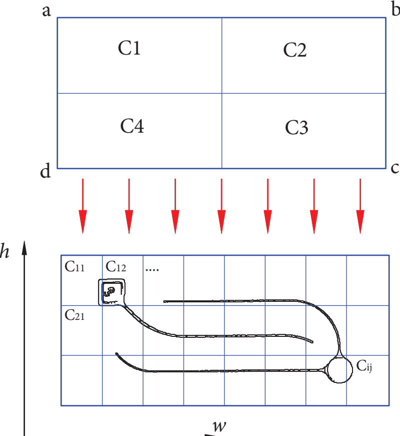

Figure 8: Mesh segmentation model.

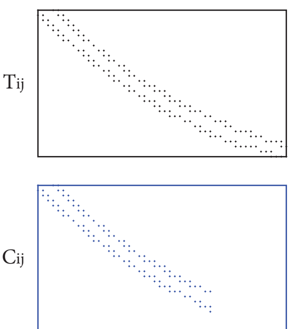

Figure 9: Rough detection model of geometric calculation.

convolutional neural network model for secondary fine inspection. Since the main types of chip defects are vestigial or missed engraving of photolithographic, the coarse screening effect of this method is as follows:

$$\eta_{i,j}=\frac{\sum C_{i,j}}{\sum T_{i,j}}=\frac{\text{num}\left(p_{i,j}\right)_C}{\text{num}\left(p_{i,j}\right)_T}\times100\%, \tag{1}$$

where $\text{num}(p_{i,j})_C$ is the number of pixel points contained in grid $C_{i,j}$ and $\text{num}(p_{i,j})_T$ is the number of pixel points contained in grid $T_{i,j}$.

4.2. Convolutional Neural Network-Based Defect Detection Model. Convolutional Neural Network (CNN) is a class of feedforward neural networks with deep structure containing convolutional computation, and is one of the hot algorithms of deep learning, from the initial LeNet, AlexNet, InceptionNet, and VGG to the later ResNet, MobileNet, and EfficientNet, the algorithms are continuously optimized and developed. Due to the various types of chip defects, including cathode, anode along the defective, broken electrodes, substrate contamination, etc., and the incomplete lithography at different locations that belong to the defect

category, there is a diversity of detection object characteristics, so the convolutional neural network model is first trained to learn and then detected. In this paper, a hybrid detection network model based on SPP-Net and CNN is used for accurate detection of "suspected defect" images after geometrically calculated coarse screening. A typical convolutional neural network generally consists of five layers, the first layer is the input layer, which input an original two-dimensional image; the second layer is the convolutional layer, using different convolutional kernels such as 3×3, 5×5, and 7×7, and the input layer for slip convolution calculation, and each convolutional kernel and the original image convolution can be obtained after a corresponding feature map; the third layer is the pooling layer, which uses the pooling process to further compress the feature map, extract the most significant features, reduce the amount of computation, and improve the robustness of the algorithm. The fourth layer is the fully connected layer, which is consistent with artificial neural networks and fully connects each neuron with a series of feature values to ensure that all features of the whole image are retained to the maximum extent; the last layer is the output layer, which integrates the features and outputs a probability value. As shown in Figure 10, for a chip image as an example, the collected defective images are trained.

The structure of the hybrid network model constructed in this paper is shown in Figure 11. Taking a 128×128 input image as an example, a $16 \times 16 \times 64$ feature model is obtained after convolution and activation operations with different convolution kernels in multiple layers and channels. Due to the tiny size of LED chip images, the loss of certain features is inevitable after the multi-layer convolution operation with different convolution kernels. In order to increase the perceptual field of the image and reduce the feature loss caused by the convolution operation, the SPP-Net feature pyramid model is used for feature enhancement.

As shown in Figure 12, the SPP-Net network uses a variety of different pooling kernels (1×1, $m \times m$, $n \times n$, etc.) to maximize pooling with the input layer to extract more significant global and local features, and finally performs a fusion operation to obtain a more comprehensive set of original image features. SPP-Net was proposed in 2014 as a network structure that disregards image size, outputs images of fixed length, and can perform stably despite image distortion. SPP-Net collects surrounding information by increasing the filter size, but this will inevitably lead to an exponential increase in computational cost. To overcome this drawback, DeepLabV3+ presented a modified Atlas Spatial Pyramid Pooling (ASPP) module. Choi and Cha [37] adopted this ASPP module in the SDDNet, proposed a new deep learning architecture, and applied it to the real-time segmentation of superficial cracks in structures with very good results. In order to retain the global features of the original image to the greatest extent, this paper improves the structure of the traditional SPP-Net model by adding another superposition operation on the pooling kernel 1×1 in the original SPP-Net model, and the superposition features are taken directly from the original input map. In order to ensure the same size superposition law for both, the original input map is subjected to a corresponding maximum pooling process to match the size of the input feature map of the SPP-Net model. The improved SPP-Net model integrates the combined information of the four scales and the original image, which is richer than the previous feature information, and this form of hybrid grouping scheme facilitates a significant improvement in the computational efficiency of feature extraction.

The signal processing nature of neurons is a linear combination, and even if the depth of the convolutional layer is increased, the whole network is still a linear combination, which can easily lead to a gradient disappearance problem during the convolutional neural net training. The activation function can effectively fit the output value of the neural network, effectively improve the learning efficiency, accelerate the convergence speed of the algorithm, and suppress the gradient disappearance phenomenon. The commonly used activation functions are SIGMOD, Tanh, and ReLU. Since it has the advantages of unilateral inhibition and sparse activation, ReLU is used for activation, and its function is defined as follows:

$$f(x) = \begin{cases} 0, x < 0, \\ x, x \geq 0. \end{cases} \tag{2}$$

In order to improve the learning efficiency, the weights and biases of the output parameters need to be optimized, the essence of which is to minimize the loss of the training model, which is continuously optimized through sample training to reach the ideal value, and the iterative formula for the weights and biases is as follows:

$$w_{\text{new}} = w_{\text{old}} - \eta \frac{d\,\text{loss}}{dw_{\text{old}}}, \tag{3}$$

where w_{old} is the current weight value, loss is the loss function, η is the learning rate, and w_{new} is the updated weight value.

The loss function of the hybrid algorithm proposed in this paper consists of two parts: the bounding box error and the classification error, where the mean square error is used for the bounding box error and the multi-categorycross-entropy error is used for the classification error, and an $m \times m$ prediction bounding box is used, and this loss function is as follows:

$$\text{loss} = \text{loss}_{\text{box}} + \text{loss}_{\text{class}} = \lambda_{\text{box}} \sum_{i=0}^{m \times m} \left[\left(x_{it} - x_{it}^{*}\right)^2 + \left(\text{y}_{it} - y_{it}^{*}\right)^2 + \left(x_{ib} - x_{ib}^{*}\right)^2 + \left(\text{y}_{ib} - y_{ib}^{*}\right)^2 \right] - \sum_{c=1}^{M} \log\left(\text{p}_{\text{ic}}\right), \tag{4}$$

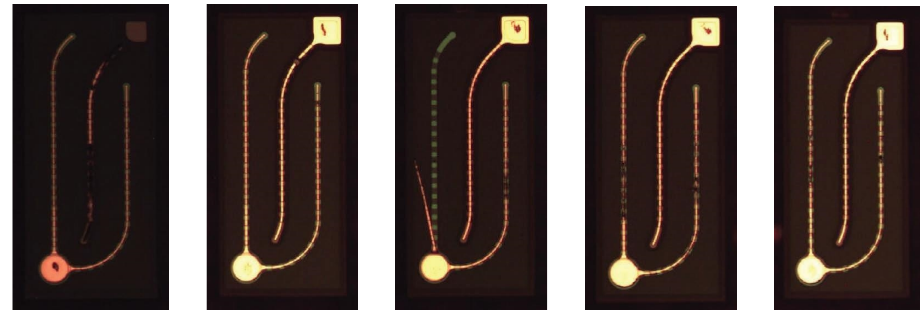

Figure 10: Various chip defect samples.

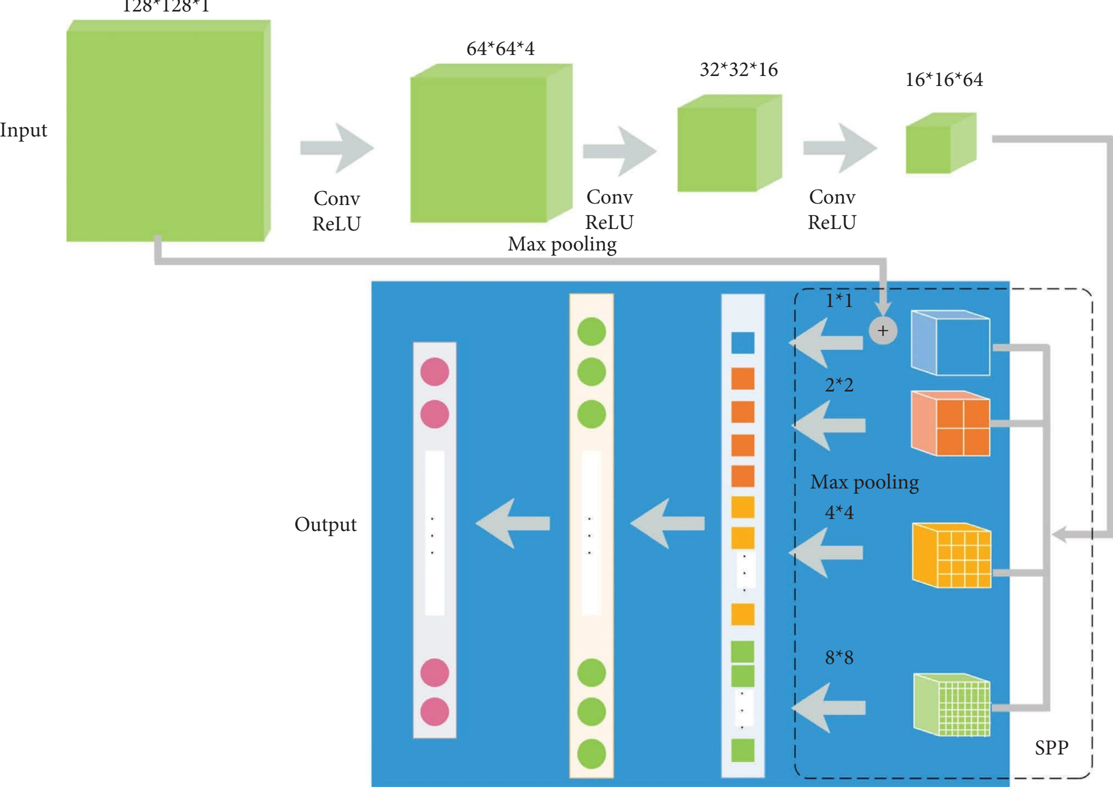

Figure 11: Detection network model based on improved SPP and CNN hybrid.

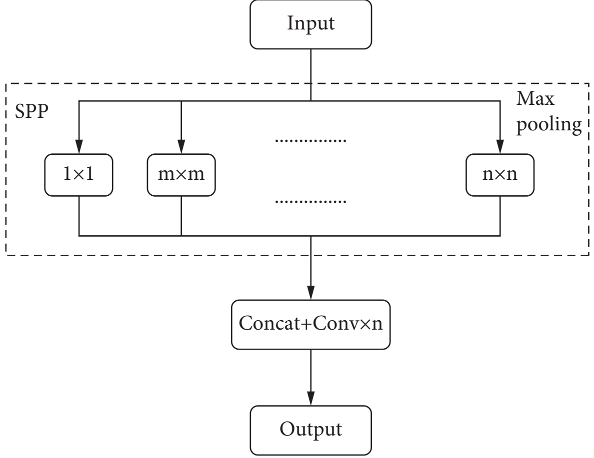

Figure 12: SPP-net network structure.

where λ_{box} is the penalty coefficient of the bounding box prediction. $(x_{it}, y_{it}, x_{ib}, y_{ib})$ and $(x_{it}^*, y_{it}^*, x_{ib}^*, y_{ib}^*)$ are the horizontal and vertical coordinate values of the upper left corner point and lower right corner point of the predicted and real bounding boxes, respectively. M is the number of target categories, and p_{ic} is the predicted probability that observation sample i belongs to category c.

5. Experiments and Discussion

To verify the effectiveness of this method, the hybrid algorithm is applied to LED chip defect detection for the corresponding tests. The experimental environment is 8.00 GB of memory, an Intel Core i5-5200U with 4 cores and 8 threads at 2.20 GHz, Intel HD Graphics 5500 as the graphics card, and Windows 7 as the operating system.

5.1. Sample Testing. The test samples are obtained from the LED chip images acquired by the acquisition platform in Figure 2 after 2000 times magnification. 300 defective chip images are collected, of which 200 are arbitrarily selected as training objects and another 100 are used as test objects. In order to expand the number of training sets, image processing techniques are used to flip the position, add different noises, and sharpen 200 images, expanding the 200 original images into 2000 training samples. The details of the datasets are presented in Table 1.

According to the characteristics of the training set, the following evaluation metrics are introduced to test and analyze the proposed hybrid algorithm: average precision (AP), mean average precision (mAP), mean intersection over union ($mIoU$) and precision (P) and recall (R), and each metric is calculated as follows:

$$
\begin{aligned}
P &= \frac{\mathrm{TP}}{\mathrm{TP}+\mathrm{FP}},\\
R &= \frac{\mathrm{TP}}{\mathrm{TP}+\mathrm{FN}},\\
AP &= \int_0^1 P(R)\mathrm{d}R,\\
mAP &= \frac{\sum_{i=1}^{M} AP_i}{M},\\
mIoU &= \operatorname{mean}\left(\frac{\mathrm{TP}}{\mathrm{TP}+\mathrm{FP}+\mathrm{FN}}\right),
\end{aligned}
\tag{5}
$$

where TP is the number of qualified samples correctly detected; FP is the number of defective samples mistakenly detected as qualified samples; and FN is the number of qualified samples mistakenly detected as defective samples.

As shown in Figure 13, some results of defect detection of a certain type of LED chip are shown. It can be seen that the method can effectively detect defects in different areas of the LED chip lithography graph and complete the marking of the defect location, and the algorithm has a high defect recognition rate and good robustness.

5.2. Discussion and Analysis. To further validate the adaptability of this method for LED chip defect detection and to more comprehensively evaluate the detection performance and efficiency, it is compared and analyzed with various classical convolutional neural network detection algorithms such as R-CNN, Fast R-CNN, Faster R-CNN, SDD, YOLO v2, YOLO v3, YOLO v4, and YOLO v5. R-CNN, Fast R-CNN and Faster R-CNN are classical two-stage detection algorithms, which are computationally intensive and the detection process is time-consuming; SDD and YOLO v2 are typical single-stage target detection methods, and although they improve the detection speed, they do not have high detection accuracy when locating tiny targets. For the abovementioned different detection methods for LED chip detection comparison experiments, the respective performances are shown in Table 2.

Accordingly, the above comparison experiments show that: the hybrid algorithm-based LED chip defect detection method proposed in this paper is better than other methods in terms of correct detection rate and speed, the method is better adapted to the characteristics of the tiny size of LED chips, the accuracy of LED chip defect detection reached 96.7%, and the average detection time of each LED chip defect image reached 43 ms to meet the LED chip defect online detection needs.

To summarize, the proposed approach has the following merits concerning the existing ones:

(1) As discussed above, the hybrid algorithm proposed in this paper combines the respective advantages of geometric dimensionality reduction calculation and a convolutional neural network model and uses a step-by-step detection mechanism to achieve fast screening of large sample data with a low defect rate by using the geometric dimensionality reduction calculation method, which largely overcomes the limitation of a single detection method for undifferentiated detection of large sample data and avoids a large amount of invalid detection data. This is also reflected in the detection comparison test result, which shows that the geometric dimensionality reduction method effectively filters a large amount of useless image information and contributes positively to the improvement of detection speed, which is the intrinsic mechanism of the method's improvement in detection speed relative to other methods such as R-CNN, SDD, and YOLO.

(2) LED chips are tiny and diverse, and their detection belongs to the category of small target detection. In this paper, we improve the SPP-Net model by superimposing a layer of the original input information on the SPP output to avoid using a deeper convolutional layer model in order to increase the image perceptual field. The advantage of this is that it not only overcomes the loss of small target features due to deep convolutional layer calculation but also makes up for the original feature information to a greater extent, reduces the number of layers of convolutional layers, and effectively reduces the computational effort while ensuring detection

Table 1: The datasets for training and testing.

Defect class	Training			Testing		Total
	Size	Number of images	Number of augmented images	Size	Number of images	
P-electrode	1536 × 1024	10	100	1536 × 1024	5	115
P-goldwire	1280 × 720	90	900	1280 × 720 1024 ×512	50	1040
N-electrode	1536 × 1024	15	150	1536 × 1024	5	170
N-goldwire	1024 × 512	85	850	1280 × 720 1024 × 512	40	975

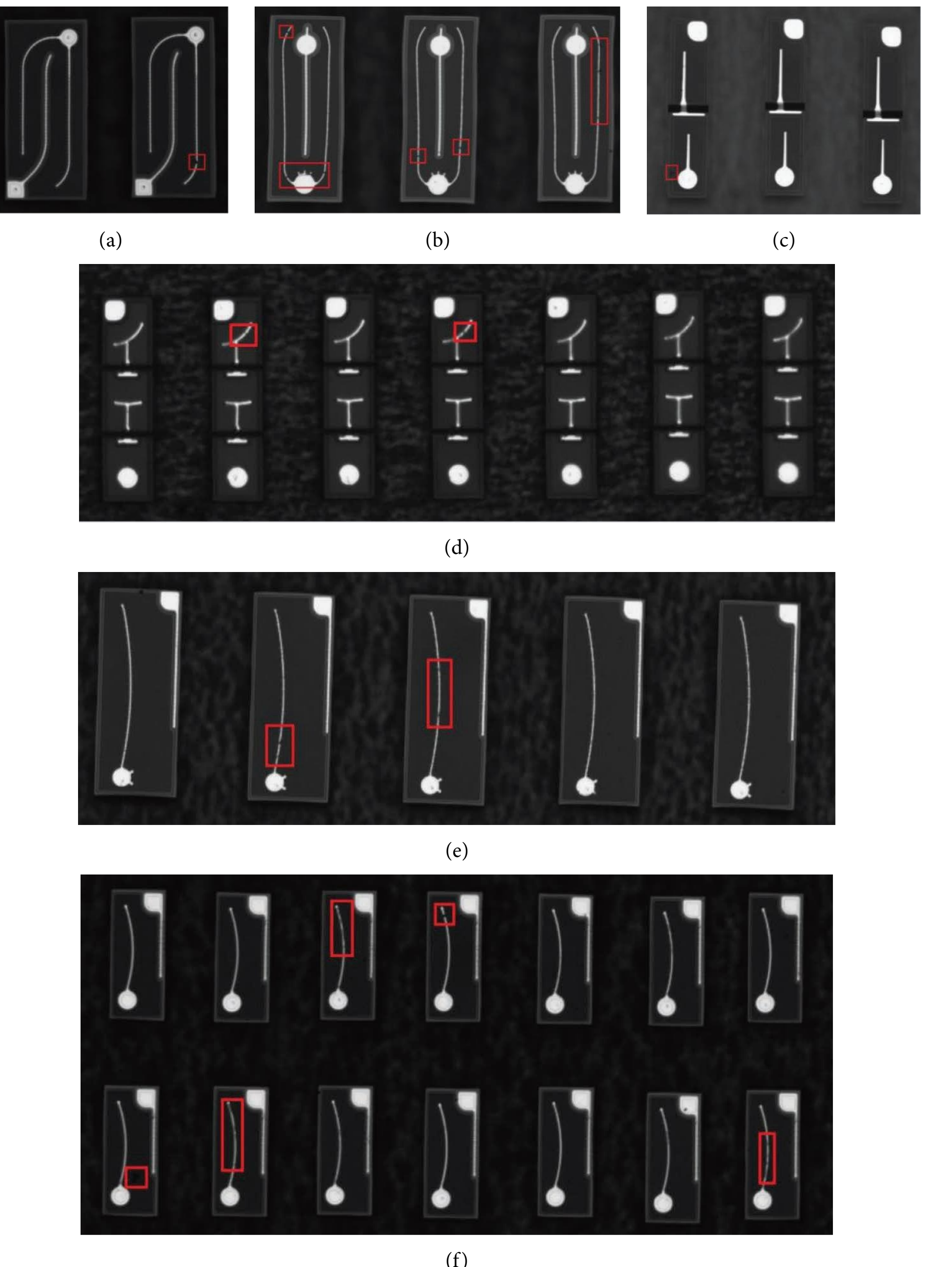

Figure 13: LED chip defect detection results based on a hybrid algorithm. (a) Defect detection result 1. (b) Defect detection result 2. (c) Defect detection result 3. (d) Defect detection result 4. (e) Defect detection result 5. (f) Defect detection result 6.

Table 2: Comparison of the results of the eight methods in the LED chip detection experiment.

Methods	mIoU (%)	mAP (%)	AP (%)	Speed (ms)
R-CNN	49.37	61.59	63.45	1546
Fast R-CNN	57.76	71.05	71.39	862
Faster R-CNN	69.41	82.13	82.53	469
SDD	71.29	83.18	85.64	215
YOLO v2	73.35	84.16	86.42	156
YOLO v3	77.63	90.32	91.52	96
YOLO v4	80.51	92.74	94.16	73
YOLO v5	82.16	93.21	95.45	65
Ours	85.79	94.55	96.70	59

accuracy. From the comparison experiments with other methods, it is found that the proposed detection network model based on the improved SPP and CNN hybrid is positively helpful in improving the *AP*, *mAP,* and *mIoU* of detection.

(3) The existing defect detection algorithms focus more on the neural network model or deep neural network model constructs, and with the increase of convolutional layer depth, the computational volume also increases, and its computational efficiency is difficult to achieve linear improvement. The geometric calculation method has the characteristics of high accuracy and accurate localization, but its self-learning ability is poor, and it is difficult to cope with diverse detection objects, so it has great limitations for the detection of a wide variety of LED chips. The efficiency of the proposed method lies in the organic combination of geometric computation and a neural network model. On the one hand, the advantage of high accuracy in geometric computation is used to overcome the difficulty of the low detection rate of the neural network model for small targets; on the other hand, the self-learning function of the neural network model is fully utilized to adapt to the application scenario of diversified LED chips. The proposed approach in this paper takes into account the simultaneous existence of the above two relations, so the detection results are more reliable and credible.

6. Conclusion

On the basis of analyzing and summarizing the characteristics of existing convolutional neural network algorithms and their adaptability, we propose a hybrid algorithm integrating geometric computation and a convolutional neural network model for the LED chip defect detection problem. Geometric grid segmentation is used to reduce the dimensionality of the preprocessed LED chip images to achieve coarse screening and fast detection of defects in large quantities of chips and save computational costs. On this basis, the convolutional neural network model is used for secondary fine inspection of "suspected defect" images, and the structure of the SPP model is improved to directly introduce the original image information into the SPP network for feature enhancement processing, which takes into account both global and local features of the image and reduces the risk of losing small information during the convolutional operation. In the experimental stage, the test samples are effectively expanded by using image processing techniques to enrich the training samples of the convolutional neural network model. According to the comparison experiments with other algorithms, it is concluded that the LED chip defect detection method proposed in this paper is effective, with significantly higher detection accuracy than other methods and faster detection speeds, good robustness, and is suitable for large-scale LED chip detection applications. However, LED chip preparation belongs to the field of microscopic inspection, and there are characteristics such as huge production volumes, high integration, and difficult collection of defective samples, the way of expanding samples by means of image processing in this paper still has limitations in dealing with diverse chip type scenarios, and its detection rate will be affected. In addition, there is variability in the background of different types of chip images, leading to certain limitations in the application scenarios of this image preprocessing method.

Despite the many advantages of the proposed algorithm, some limitations remain, and further modifications of the algorithm will be carried out in the following areas. We will extend the proposed approach to effectively deal with the LED chips defect detection problem involving lightweight design of the algorithm, weakly supervised target detection, and other optimization of the detection network model. Equally, we will optimize and improve the convolutional neural network model, increase the color channels, reduce the information loss caused by image preprocessing methods, and improve the detection efficiency. Additionally, we will categorize the typical LED chip package defect types and characteristics and establish the LED chip package defect detection method. Nevertheless, a CNN-based method requires a large amount of training data to train a robust classifier. The main limitation of the proposed method is that due to the absence of the samples of sufficient and diverse defect types in its required training datasets, large-scale and high-precision detection is still challenging tasks. In the future, more samples with different defects will be collected and added to the database to increase the accuracy and robustness of the proposed method, extensive comparative studies with YOLO v6, YOLO v7 and other advanced algorithms will be performed. Simultaneously, we will expand our approach to other fields such as welding quality inspection, pattern recognition, and so forth.

Consequently, we will integrate geometric computation methods into machine learning theory and apply them to more engineering fields to provide a novel solution for large-scale online inspection issues.

Data Availability

The data used to support the findings of this study are included within the article.

Conflicts of Interest

The authors declare that they have no conflicts of interest with respect to the research, authorship, and/or publication of this article.

Acknowledgments

This study was supported by the Special Project of Scientific Research and Development Center of Higher Education Institutions, Ministry of Education, China (Grant no. ZJXF2022126), the New Product Pilot Program Projects of Zhejiang Province in 2022 (the second batch), China (Grant nos. 2022D60SA7020371 and 2022D60SA7017931), the Public Welfare Science and Technology Research Project of Jinhua, Zhejiang Province, China (Grant no. 2022-4-302), and the Ywicc Special Interest Group on Machine Vision (YSIGMV).

References

[1] L. S. V. Thomas and J. Gehrig, "Multi-template matching: a versatile tool for object-localization in microscopy images," *BMC Bioinformatics*, vol. 21, no. 1, p. 44, 2020.

[2] B. Hu and J. H. Wang, "Detection of PCB surface defects with improved faster-RCNN and feature pyramid network," *IEEE Access*, vol. 8, pp. 108335–108345, 2020.

[3] W. B. Huang, P. Wei, M. H. Zhang, and H. Liu, "HRIPCB: a challenging dataset for PCB defects detection and classification," *Journal of Engineering*, vol. 2020, no. 13, pp. 303–309, 2020.

[4] Z. C. Liu and B. D. Qu, "Machine vision based online detection of PCB defect," *Microprocessors and Microsystems*, vol. 82, Article ID 103807, 2021.

[5] X. T. Liao, S. P. Lv, D. H. Li, Y. Luo, Z. C. Zhu, and C. Jiang, "YOLOv4-MN3 for PCB surface defect detection," *Applied Sciences*, vol. 11, no. 24, Article ID 11701, 2021.

[6] Y. T. Li, P. Kuo, and J. I. Guo, "Automatic industry PCB board DIP process defect detection system based on deep ensemble self-adaption method," *IEEE Transactions on Components, Packaging, and Manufacturing Technology*, vol. 11, no. 2, pp. 312–323, 2021.

[7] K. An, Y. P. Zhang, and LpviT, "LPViT: a transformer based model for PCB image classification and defect detection," *IEEE Access*, vol. 10, pp. 42542–42553, 2022.

[8] Z. G. Ling, A. R. Zhang, D. E. Ma, Y. X. Shi, and H. Wen, "Deep siamese semantic segmentation network for PCB welding defect detection," *IEEE Transactions on Instrumentation and Measurement*, vol. 71, pp. 5006511–11, 2022.

[9] M. G. Jeon, S. Y. Yoo, and S. W. Kim, "A contactless PCBA defect detection method: convolutional neural networks with thermographic images," *IEEE Transactions on Components, Packaging, and Manufacturing Technology*, vol. 12, no. 3, pp. 489–501, 2022.

[10] T. J. Jiang, D. P. Wu, F. Gao et al., "Improved spike detection algorithm based on multi-template matching and feature extraction," *IEEE Transactions on Circuits and Systems II: Express Briefs*, vol. 69, no. 1, pp. 249–253, 2022.

[11] S. H. Chen, A. H. Chiou, and C. C. Wang, "A NMF-based image restoration scheme with applications to LED integrated substrate defect detection," *IEEE Transactions on Semiconductor Manufacturing*, vol. 31, no. 4, pp. 486–494, 2018.

[12] H. Lin, B. Li, X. G. Wang, Y. F. Shu, and S. L. Niu, "Automated defect inspection of LED chip using deep convolutional neural network," *Journal of Intelligent Manufacturing*, vol. 30, no. 6, pp. 2525–2534, 2019.

[13] Y. F. Shu, B. Li, and H. Lin, "Quality safety monitoring of LED chips using deep learning-based vision inspection methods," *Measurement*, vol. 168, Article ID 108123, 2021.

[14] M. L. Stern and M. Schellenberger, "Fully convolutional networks for chip-wise defect detection employing photoluminescence images," *Journal of Intelligent Manufacturing*, vol. 32, no. 1, pp. 113–126, 2021.

[15] B. Lin, J. Wang, X. Yang et al., "Defect contour detection of complex structural chips," *Mathematical Problems in Engineering*, vol. 2021, Article ID 5518675, 11 pages, 2021.

[16] S. H. Chen and C. C. Tsai, "SMD LED chips defect detection using a YOLOV3-dense model," *Advanced Engineering Informatics*, vol. 47, Article ID 101255, 2021.

[17] R. Chen, W. Zhong, H. Y. Yang et al., "Faster region-based hotspot detection," *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, vol. 41, no. 3, pp. 669–680, 2022.

[18] L. Xu, X. J. Hu, T. He, K. Hu, and J. M. Zhang, "Defect detection on LED chips based on position pre-estimation and feature enhancement," *Applied Sciences*, vol. 12, no. 3, p. 1265, 2022.

[19] D. H. Kang, S. S. Benipal, D. L. Gopal, and Y. J. Cha, "Hybrid pixel-level concrete crack segmentation and quantification across complex backgrounds using deep learning," *Automation in Construction*, vol. 118, Article ID 103291, 2020.

[20] D. H. Kang and Y. J. Cha, "Efficient attention-based deep encoder and decoder for automatic crack segmentation," *Structural Health Monitoring*, vol. 21, no. 5, pp. 2190–2205, 2022.

[21] R. Ali and Y. J. Cha, "Attention-based generative adversarial network with internal damage segmentation using thermography," *Automation in Construction*, vol. 141, Article ID 104412, 2022.

[22] Y. J. Cha, W. Choi, and O. Buyukozturk, "Deep learning-based crack damage detection using convolutional neural networks," *Computer-Aided Civil and Infrastructure Engineering*, vol. 32, no. 5, pp. 361–378, 2017.

[23] Y. J. Cha, W. Choi, G. Suh, S. Mahmoudkhani, and O. Buyukozturk, "Autonomous structural visual inspection using region-based deep learning for detecting multiple damage types," *Computer-Aided Civil and Infrastructure Engineering*, vol. 33, no. 9, pp. 731–747, 2018.

[24] Z. L. Wang and Y. J. Cha, "Unsupervised deep learning approach using a deep auto-encoder with a one-class support vector machine to detect damage," *Structural Health Monitoring*, vol. 20, no. 1, pp. 406–425, 2021.

[25] T. J. Zhang, C. R. Zhang, Y. J. Wang, X. F. Zou, and T. L. Hu, "A vision-based fusion method for defect detection of milling

cutter spiral cutting edge," *Measurement*, vol. 177, Article ID 109248, 2021.

[26] M. C. Zhang, H. Shi, Y. Zhang, Y. Yu, and M. S. Zhou, "Deep learning-based damage detection of mining conveyor belt," *Measurement*, vol. 175, Article ID 109130, 2021.

[27] Z. P. Wang, L. Y. Jin, S. Wang, and H. R. Xu, "Apple stem/calyx real-time recognition using YOLO-v5 algorithm for fruit automatic loading system," *Postharvest Biology and Technology*, vol. 185, Article ID 111808, 2022.

[28] T. H. Xie, X. X. Li, X. S. Zhang, J. Y. Hu, and Y. Fang, "Detection of Atlantic salmon bone residues using machine vision technology," *Food Control*, vol. 123, Article ID 107787, 2021.

[29] M. Farhan Khan, A. Alam, M. Ateeb Siddiqui et al., "Real-time defect detection in 3D printing using machine learning," *Materials Today Proceedings*, vol. 42, pp. 521–528, 2021.

[30] Y. J. Cha, K. S. You, and W. Choi, "Vision-based detection of loosened bolts using the Hough transform and support vector machines," *Automation in Construction*, vol. 71, pp. 181–188, 2016.

[31] A. John Rajan, K. Jayakrishna, T. Vignesh, J. Chandradass, and T. T. M. Kannan, "Development of computer vision for inspection of bolt using convolutional neural network," *Materials Today Proceedings*, vol. 45, pp. 6931–6935, 2021.

[32] M. K. Li, J. Y. Jia, X. Lu, and Y. Zhang, "A method of surface defect detection of irregular industrial products based on machine vision," *Wireless Communications and Mobile Computing*, vol. 2021, Article ID 6630802, 10 pages, 2021.

[33] G. H. Beckman, D. Polyzois, and Y. J. Cha, "Deep learning-based automatic volumetric damage quantification using depth camera," *Automation in Construction*, vol. 99, pp. 114–124, 2019.

[34] A. Rasheed, B. Zafar, A. Rasheed et al., "Fabric defect detection using computer vision techniques: a comprehensive review," *Mathematical Problems in Engineering*, vol. 2020, Article ID 8189403, 24 pages, 2020.

[35] C. F. Zhao, H. C. Ding, G. H. Cao, and Y. Zhang, "A new method for detecting compensation hole parameters of automobile brake master cylinder based on machine vision," *Journal of Advanced Transportation*, vol. 2021, Article ID 8864679, 14 pages, 2021.

[36] L. Y. Liu, "Moving object detection technology of line dancing based on machine vision," *Mobile Information Systems*, vol. 2021, Article ID 9995980, 9 pages, 2021.

[37] W. Choi and Y. J. Cha, "SDDNet: real-time crack segmentation," *IEEE Transactions on Industrial Electronics*, vol. 67, no. 9, pp. 8016–8025, 2020.