Surface Defect Detection Methods Based on Deep Learning: a Brief Review

Guanlin Liu
School of information and engineering
Guangxi University of Foreign Languages
Guangxi China
liuguanlin.article@foxmail.com

Abstract—Surface defect detection techniques based on deep learning have been widely used in various industrial scenarios. This paper reviews the latest works on deep learning-based surface defect detection methods. They are classified into three categories: full-supervised learning model method, unsupervised learning model method, and other methods. The typical methods are further subdivided and compared. The advantages and disadvantages of these methods and their application scenarios are summarized. This paper analyzes three key issues in surface defect detection and introduces common data sets for industrial surface defects. Finally, the future development trend of surface defect detection is predicted.

Keywords— Deep learning, surface defect detection, machine vision, convolutional neural network

I. INTRODUCTION

Surface defect detection is a very important research content in the field of machine vision, also known as Automated Optical Inspection (AOI) or Automated surface inspection (ASI), which uses machine vision equipment to obtain images to determine whether there are defects in the collected images Technology. At present, surface defect equipment based on machine vision has widely replaced artificial naked eye detection in various industrial fields, including 3C, automobiles, home appliances, machinery manufacturing, semiconductors and electronics, chemicals, medicine, aerospace, light industry, and other industries. Traditional Surface defect detection methods based on machine vision often use conventional image processing algorithms or artificially designed features plus classifiers. Generally speaking, the different properties of the inspected surface or defects are usually used to design imaging schemes. Reasonable imaging scheme is that it helps to obtain a uniformly illuminated image and reflect the surface defects of the object. One way is to select the light source for the color of the inspected surface. For example, in the literature [1], the composite white light source is selected to image the surface defects of the colored cloth. Another common method is to select different imaging schemes according to the reflection properties of the detected surface, mainly including bright field imaging, dark-field imaging, and hybrid imaging. For example, Chen [2] et al. designed two concentric ring-shaped bright field light source placed concentrically is used to simultaneously illuminate the central and peripheral areas of the bottom of the can. Tao [3] et al. used dark-field imaging to detect weak scratches on the surface of large-aperture optical elements. Although carefully constructed imaging, the solution can greatly reduce the difficulty of designing classic detection algorithms, it also increases the application cost of the detection system. At the same time, in many open industrial environments, look forward to the designed imaging system eliminates the influence of changes in the scene or the inspected material on the inspection system, which is often not realistic. In the real and complex industrial environment, surface defect detection often faces many challenges, such as the existence of defect imaging and the small difference between the background and the contrast the defect size is low, the defect size varies greatly, and the types are diverse. There is a lot of noise in the defect image, and even there is a lot of interference in the image of the defect in the natural environment. effect.

Nowadays, deep learning models represented by convolutional neural networks (CNN) have been successfully applied in many computer vision (CV) fields, such as face recognition, pedestrian re-recognition, scene text detection, target tracking, and automatic Many deep learning-based defect detection methods are also widely used in various industrial scenarios, such as driving, and even some companies at home and abroad have developed a variety of deep learning-based commercial-industrial surface defect detection software, as shown in Table 1. The market size of traditional industrial vision and its components will reach 19.2 billion U.S. dollars in 2025[4], of which China accounts for about 30%, and maintains an average annual growth rate of 14%. This field is gradually being used by the new generation based on depth. Industrial vision technology to replace learning. At the same time, China proposed in the "Made in China 2025" white paper... Accelerate the improvement of product quality. Implement the action plan for the improvement of industrial product quality... Promote the use of intelligent testing equipment... Make key physical products performance stability, quality reliability, environmental adaptability, service life and other indicators have reached the international advanced level of similar products... ". Therefore, the surface defect detection method based on deep learning not only has important academic research value but also has very broad Market application prospects.

Since there is no comprehensive and detailed review literature on surface defect detection methods based on deep learning in China, this article summarizes the relevant literature from 2014-2019 to help researchers quickly and systematically understand the relevant methods and technologies in this field. This article is arranged as follows: The first part gives the
definition of defect detection problems. The second part focuses on the detailed introduction, subdivision and comparison of related methods in recent years. The third part analyzes three key problems in surface defect detection based on deep learning. Part 4 introduces the public defect detection data set in the industrial field. Finally, it looks forward to the possible future research focus and development direction.

II. DEFINITION OF DEFECT DETECTION PROBLEM

Definition of defects: In machine vision tasks, defects tend to be a concept based on human experience, rather than a purely mathematical definition. Different perceptions of the defect pattern will lead to two completely different detection methods. Detection of defects on the surface of cloth the first supervised method is embodied in the use of labeled defect images (including categories, rectangular boxes, or pixel-by-pixel, etc.) to input them into the network for training. At this time 'defects mean Marked area or image. Therefore, this method pays more attention to defect features. For example, in the training phase, areas containing large black areas or images are marked as different color defects for network training. In the testing phase, when a large amount of black features are detected in the cloth image, that is to say, there is a "different color" defect. The second is an unsupervised defect detection method, which usually only requires normal non-defective samples for network training, also known as one-class learning. This method focuses more on non-defect (ie normal Sample) feature, when a feature (abnormal feature) that has not been seen before is found during the defect detection process, it is considered that a defect has been detected. At this time 'defect" means an abnormality, so this method is also called anomaly detection (Anomaly Detection)

Definition of defect detection: Compared with the clear classification, detection and segmentation tasks in computer vision, the requirements for defect detection are very general. Its requirements can be divided into three different levels: 'What is the defect", "Where is the defect" and 'how much is the defect". The first stage 'what is the defect" corresponds to the classification task in computer vision, three types of defects are classified: heterochromatic, void and warp. The task at this stage can be called \ "Defect classification", only gives the category information of the image. The second stage 'Where is the defect" corresponds to the positioning task in computer vision. The defect location at this stage is the detection in the strict sense. It is not only to obtain what types of images exist Defects, and also give the specific location of the defect, the heterochromatic defect is marked with a rectangular box. The third stage 'Defect is "corresponding to the segmentation task in computer vision, as shown in the area of defect segmentation. It shows that the defect is segmented from the background pixel by pixel, and a series of information such as the length, area, and location of the defect can be further obtained. This information can assist the high-level quality evaluation of the product, such as the judgment of the quality of the product. Although the defect the functional requirements and objectives of the three phases of inspection are different, but the three phases are mutually included and can be converted into each other. For example, the second phase 'defect location" includes the first phase 'defect classification" process, the third phase 'defect Segmentation "can also complete the second stage 'defect location". The first stage 'defect classification" can also achieve the goals of the second and third stages through some methods. Therefore, in the following, it will be collectively referred to as defects according to traditional industrial customs. Detection is only distinguished when targeting different network structures and target functions.

III. SURFACE DEFECT DETECTION DEEP LEARNING METHOD

This section summarizes and summarizes the surface defect detection methods based on deep learning. According to the different labels of the data, the whole is divided into fully supervised learning models, unsupervised learning models, and other methods (semi-supervised learning models and weakly supervised learning models). Learning model). In the fully-supervised model, according to the difference between the input image method and the loss function, it is divided into the method based on representation learning and metric learning. In the representation learning, according to the different network structures, it can be further subdivided into classification network and detection networks, and segmentation networks. At present, a large amount of research work focuses on the direction of fully supervised learning, but unsupervised learning is also a direction worth studying. This paper is subdivided according to the processing characteristics of each type of method for several different sub-methods.

Nowadays, most surface defect detection based on deep learning is based on supervised representation learning methods. The essence of representation learning is to treat defect detection as a classification task in computer vision, including coarse-grained image label classification or region classification, and the finest pixel classification. Since the achieved goal is completely consistent with the computer vision task, the defect detection method based on representation learning can be regarded as an application of its related classic network in the industrial field.

In real industrial production, the huge differences in the shape, size, texture, color, background, layout, and imaging lighting of the detected object make the classification of defects in complex environments a difficult task. Due to the powerful feature extraction capabilities of CNN, the use of CNN's classification network has now become the most commonly used mode in surface defect classification. Usually, the feature extraction part of the CNN classification network consists of a cascaded convolutional layer + pooling layer, followed by a fully connected layer (or average pooling layer) + softmax structure Used for classification. Generally speaking, existing networks for surface defect classification often use existing network structures in computer vision, including AlexNet [5], VGG [6], GoogLeNet [7], ResNet [8], DenseNet [9], SENet[10], ShuffleNet[11], MobileNet [12], etc. Or build a simple network structure for practical problems, by inputting a test image to the classification network, the network outputs the category of the image and its category confidence. According to the differences in the tasks achieved by the classification network method, we subdivide it into three small categories: directly use the network for classification, use the network for defect location and use the network as a feature extractor.

Directly using the classification network to do the classification task of defects is the earliest common method
used by CNN in surface defect detection. According to the characteristics of the research work, it can be further subdivided into original image classification, ROI (Region of Interest) positioning, and multi-class classification. There are three categories.

A. Original image classification

That is, the collected complete defect images are directly put into the network for learning and training. In 2014, the Austrian Institute of Science and Technology [13] was the first to collect photometric stereo images to train the CNN network to realize the classification of the orbit surface void defects. The entire network contains a total of two convolutional layers. With the pooling layer and the last fully connected layer, the final error recognition rate reached on the rail surface data set is 1.108%. Park et al. [14] designed a simple CNN classification network to automatically detect dirt on surface parts,Scratches, burrs, and wear, and other defects. The average detection accuracy rate of this method on the experimental defect data set is 98%, and its detection speed is 5285 samples/min (image resolution is 32×32 pixels). Kyeong [15] A convolutional neural network framework is proposed to classify the mixed-type defect patterns in the wafer warehouse map WBM in the semiconductor industry. Literature [16] uses a modified VGG19 network to identify solar panel image defects with a resolution of 300×300. The accuracy of the network reaches 88.42%, which exceeds the effect of a variety of manual design features (including KAZE [17], SIFT [18], SURF[19]) and Support Vector Machine (SVM, Support Vector Machine) methods. Liang et al. [20] proposed a method based on ShuffLeNetV2 network classification of bottle inkjet code defects under the complex background. The proposed method obtained 99.88% classification accuracy on the online inkjet code detection equipment in the plastic container industry. Directly use the original image to classify the method is widely used and can be used for defect classification in many fields, such as welding defect classification [21], polymer lithium battery blistering defect classification [22] and PCB board defect classification [23].

B. Classification after positioning the ROI

It is more common in many industrial applications. Generally speaking, we usually only pay attention to whether there are defects in a certain fixed area in the entire image obtained, so we often obtain the region of interest (ROI) in advance, and then ROI input network to judge the defect category. Shang [24] and others proposed a two-stage rail defect-recognition algorithm, first using canny and straight-line fitting algorithm to crop the rail area on the entire original image. Then the cropped image is put into the Inception V3 network to extract features for orbital image analysis class. In the literature [25], the bolt area of the high-speed railway catenary is obtained through the cascaded target detection network, and then the cropped bolt image is input into the CNN network for defect classification. Li [26] and others first used the local binary The cascaded target detector with the characteristics of the pattern (LBP, Local Binary Pattern) realizes the positioning of the scanning tunnel microscope imaging material to be detected, and then uses the CNN model to obtain the specific types of surface defects.

C. Multi-category classification:

When the defect type to be classified exceeds two types, the conventional defect classification network is the same as the original image classification method, that is, the output node of the network is the number of defect types + 1 (including the normal category). However, the multi-class classification method often uses one the basic network performs two classifications of defects and normal samples, and then shares the feature extraction part on the same network to modify or increase the classification branch of the defect category. This method is equivalent to preparing a pre-training weight parameter for the subsequent multi-target defect classification network. This weight parameter is obtained through binary classification training between normal samples and defective samples. Xie et al. [27] first trained the first ND (normal defective) CNN model for two classifications (normal images and all other defective images), which alleviated the problem of data The problem of balance. After training the NDCNN model, change the output vector to a 6-dimensional vector to train the ID (inter-defect) CNN model to make it suitable for multi-class defect label problems. The model uses defect images based on the NDCNN weights Fine-tuned, thereby reducing the sample size requirement and saving training time. Fusaomi et al. [28] proposed a ssnet (Net with SVMs to classify sample images) network, which has two classification branches, the first two classification branch is used to classify normal samples and NG samples. The network model uses AlexNet for feature extraction. Its classifier uses SVM. The second branch of the network is used for 7 categories of defect classification. Multi-category classification uses this two-branch structure, which can be sufficient Use the characteristics of the uneven number of defective samples and normal samples to explore the differences between the characteristics of the two.

IV. CONCLUSION

With the development of artificial intelligence technology, the current research focus of surface defect detection based on machine vision has shifted from classic image processing and machine learning methods to deep learning methods, and in many industrial scenarios, problems that cannot be solved by traditional methods in the past have been solved. This paper systematically summarizes, compares, and analyzes the research progress of deep learning algorithms in the field of surface defect detection. At the same time, it looks forward to the research trend of surface defect detection based on deep learning, hoping to provide relevant researchers with detailed and effective references

REFERENCES

References


