ISSN: 2688-8645

A Review of Surface Defect Detection of Metal Workpieces Based on Machine Vision

Shiwei Xu

College of Artificial Intelligence, Wenzhou Polytechnic, Wenzhou, 325035, China

Abstract

Surface defects are an inevitable problem in the production process of metal workpieces, and surface defect detection methods based on machine vision overcome the problems of low detection efficiency, high false detection and leakage rate of manual inspection methods, and are gradually being applied in industrial production. In this paper, the surface defect detection methods of metal workpieces based on machine vision in recent years were sorted into two categories: conventional machine vision-based detection methods and deep learning-based detection methods, and the basic principles of various typical methods were outlined. The development routes of deep learning-based detection methods and the advantages and disadvantages of various methods were highlighted, and the problems and solution ideas in industrial practical applications were analyzed. Finally, the existing methods of surface defect detection based on machine vision were summarized and the future development trend was prospected.

Keywords

Machine vision; Metal workpiece; Surface defect detection; Deep learning.

1. Introduction

With the rapid development of China's manufacturing industry, the demand for the quantity and variety of metal workpiece products is increasing day by day. Consumers and manufacturers have put forward higher requirements for the quality of metal workpieces, which need to meet the normal use performance, but also good surface quality. Therefore, the surface quality of metal workpieces has become one of the important competitive indicators in the market, and the role of quality control of metal workpiece surfaces in industrial production is becoming more and more significant [1]. However, in the actual production process, due to the influence of the process, production equipment and site environment and other factors, resulting in various defects on the surface of metal workpieces, which not only directly affect the appearance of the workpiece itself, but also affect the performance and commercial value of the workpiece. Therefore, in the production and processing of metal workpieces must be quality inspection of its surface, in order to timely detection of defects and control, so as to reduce the production of defective workpieces and improve the economic efficiency of enterprises.

In the past, small and medium-sized enterprises were limited by capital and technology, and mostly used manual visual inspection and laser scanning inspection [2]. Among them, manual visual inspection is carried out by inspectors through naked-eye observation for quality inspection and defect judgment, which mainly relies on inspectors' experience and lacks quantitative testing standards, poor reliability, low testing efficiency, high labor intensity, and easy to produce leakage and misdetection. Laser scanning inspection is a detection technology developed with the maturity and perfection of laser technology, which is highly sensitive and can meet the real-time requirements of product surface quality inspection, but it has insufficient ability to distinguish defects with poor contrast, and the system's recognition accuracy is not high. At the same time, because its optical system structure is quite complex, the detection signal is easily disturbed by the external environment, and the stability and maintainability of the system is poor.

With the development of image processing technology, machine vision-based surface defect detection methods have gradually replaced manual inspection methods and are practiced in industrial production inspection links. Machine vision inspection technology is a non-contact automatic inspection technology, with the advantages of safety and reliability, high detection accuracy, can operate in complex production environment for a long time, etc., is an effective way to achieve factory production automation and intelligence.

This paper analyzes machine vision-based methods for detecting surface defects on metal workpieces, outlines the development history and principles of various inspection methods, and compares them. Finally, the outlook and summary of the machine vision-based metal workpiece surface defect detection technology are presented.

2. Machine Vision-based Detection of Surface Defects on Metal Workpieces

There are two research methods to apply machine vision technology to metal workpiece surface defect detection: conventional machine vision and deep learning vision. Conventional machine vision is based on digital signal analysis and processing theory, and then machine learning methods are used to get the desired results. With the improvement of computer hardware performance and breakthroughs in artificial intelligence algorithms, the research hotspot has changed to a deep learning-based approach. Compared with conventional techniques, deep learning vision can achieve higher accuracy rates for tasks such as image classification, semantic segmentation, and object detection [3]. Although the research hotspots in the field of machine vision in recent years are in the direction of deep learning, conventional machine vision methods for some specific types of surface defects do not have disadvantages in terms of detection speed and accuracy, but on the contrary some ideas and techniques are worth learning from, so the research on deep learning vision detection is of reference value.

2.1. Surface Defect Detection Methods Based on Conventional Machine Vision

Surface defect detection based on conventional machine vision is mainly divided into two parts: image pre-processing and defect detection. Image pre-processing includes algorithms such as image denoising and image segmentation, which is the pre-work of defect detection. The defect detection part mainly uses image feature extraction algorithms to complete the detection of defects, and its algorithmic process can be summarized as follows.

(1) Selection of regions of interest, selecting regions that may contain objects.

(2) Feature extraction of regions that may contain objects.

(3) Detection and classification of the extracted features.

Three typical image feature extraction algorithms are as follows.

2.1.1 VJ Detector

VJ (Viola Jones) Detector[4] uses a sliding window to check whether the target exists in the window or not, the detector seems to be simple and stable, but the time complexity is extremely high due to the large amount of computation. To solve this problem, VJ Detector greatly improves the detection speed by combining the following three techniques:

(1) The integral map: a fast computation method for features.

(2) AdaBoost: an effective classifier learning method.

(3) The design of cascade structures: an efficient classification strategy.

HOG Detector

The HOG (Histogram of Oriented Gradients) Detector [5] was proposed in 2005 as an important improvement to the then Scale Invariant Feature Transform and Shape Contexts, in order to balance feature invariant (including translation, scale, illumination, etc.) and nonlinear (distinguishing different object classes), HOG Detector improves detection accuracy by computing overlapping local contrast normalization on a dense grid of uniformly spaced cells, so HOG Detector detector is an algorithm for feature histogram extraction based on local pixel blocks, and it works well under both local deformation of the target and under the influence of HOG Detector has laid an important foundation for many later detection methods, and related techniques are widely used in major applications of computer vision.

2.1.2 DPM Detector

As the winner of VOC 2007-2009 Object detection Challenge, DPM [6] (Deformable Parts Model) is a well-deserved SOTA (State Of The Art) algorithm among the conventional algorithms for object detection.DPM was proposed in 2008 and has made many improvements compared to HOG, so the algorithm can be seen as an extension of HOG. The DPM algorithm consists of a main filter (Root-filter) and several auxiliary filters (Part-filters), and improves the detection accuracy through Hard negative mining, Bounding box regression and Context priming) techniques to improve detection accuracy. As SOTA of conventional object detection algorithm, DPM method is fast and can adapt to object deformation, but it cannot adapt to large rotations, so its stability is poor.

Conventional object detection algorithms based on manual extraction of features have three main drawbacks.

(1) Insufficient recognition and accuracy, which may produce multiple correctly identified results.

(2) computationally intensive and slow.

(3) The extracted features tend to be valid only for a particular defect type, so they are only suitable for products with clear contours and single defects, and not for products with complex backgrounds.

2.2. Deep Learning Based Surface Defect Detection Method

Conventional machine vision-based methods for surface defect detection rely on manual feature extraction, and progress is slow and performance is low. Until the rise of CNN(Convolutional Neural Networks) in 2012, deep learning gradually became a mainstream technology in the field of computer vision. Computer vision generally includes three major tasks such as image classification, object detection, and image segmentation. Among them, image classification is to find out a classification result for the whole image; object detection is to obtain the position and class information of the object in the image; image segmentation is to further obtain the outline of the object on the basis of object detection.

According to the task requirements of common metal workpiece surface defect detection, it can generally correspond to the object detection task, which is to label the location and class of defects in the image. There are two main technical development lines for CNN-based object detection: anchor-based and anchor-free methods, and the anchor-based methods include One-stage and Two-stage detection algorithms, see Figure 1. Generally, Two-stage detection algorithms achieve higher accuracy than One-stage, while One-stage detection algorithms are faster than Two-stage.

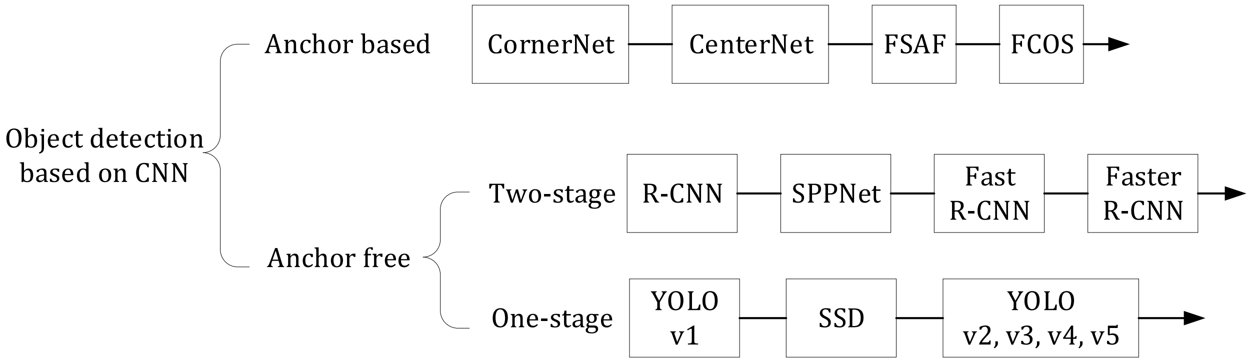

Figure 1. The roadmap of CNN-based surface defect detection algorithm

2.2.1 Anchor-Based Two-stage Surface Defect Detection Algorithm

The two-stage detection algorithm is divided into the following two main stages:

Stage1: generation of region proposals from images.

Stage2: generation of final object borders from region proposals.

The two-stage detection algorithms are mainly represented by the R-CNN (Region-CNN) [7] family.R-CNN first selects the possible object frames from a set of object candidate frames by the selective search algorithm Selective Search, then resize the images in these selected object frames to a certain fixed size image and feed them to a CNN model (a CNN model trained on the ImageNet dataset, such as AlexNet) to extract features, and finally the extracted features are fed to an SVM classifier to predict whether there is a target to be detected in the image in that object box, and further to predict which class the detected target specifically belongs to.The R-CNN algorithm achieves very significant results on the VOC-07 dataset, with an average accuracy was improved from 33.7% (DPM-V5, SOTA for conventional machine vision detection) to 58.5%. Compared with the conventional algorithm, the deep learning-based detection algorithm has made a qualitative leap in accuracy. However, the disadvantage of R-CNN is that the redundant computation of overlapping frames (more than 2000 candidate frames for a single image) features makes the detection of the whole network slow.

To reduce the redundant computation due to a large number of overlapping frames, K. He et al. proposed SPPNet [8].SPPNet proposes SPP (Spatial Pyramid Pooling Layer). Its main idea is for a pair of images divided into blocks of images at several scales, and then the features extracted from each block are fused together to take into account features at multiple scales.SPPNet is more than 20 times faster than R-CNN, but does not address the process complexity of multi-stage training.

Fast R-CNN [9] and Faster R-CNN [10] are proposed based on RCNN and SPPNet.Faster RCNN is the first end-to-end, closest to real-time performance deep learning detection algorithm.The main innovation of this network is the proposed region selection network for generating candidate frames, which can greatly improve the generation speed of detection frames. The network first inputs an image into the convolutional network to generate a feature mapping of that image. Region Proposal Network is applied on the feature mapping to return object proposals and corresponding scores. The Rol pooling layer is applied to correct all the proposals to the same size. Finally, the proposals are passed to the fully connected layer to generate the bounding box of the target object.

2.2.2 Anchor-Based One-stage Surface Defect Detection Algorithm

The One-stage detection algorithm does not require a region proposal, and directly generates the class probability and position coordinate values of the object, and directly obtains the final detection result after one stage, so it has a faster detection speed.

YOLO v1 [11] is the first One-stage detection algorithm which is very fast.The idea of this algorithm is to divide the image into multiple grids and then predict the bounding box for each grid simultaneously and give the corresponding probability. For example the center of a target to be detected falls within one of the cells divided in the image, then that cell is responsible for predicting the location and class of that target. yolo v1 detection is very fast and achieves an mAP of 52.7% on the VOC-07 dataset, achieving a real-time performance of 155 fps. Compared to the Two-stage detection algorithm, although the detection speed of the YOLO v1 algorithm is greatly improved, the accuracy is relatively taught low, especially for some small object detection problems.

The SSD algorithm [12] proposes multi-reference and multi-resolution detection techniques. SSD has several different detection branches and different detection branches can detect multiple scales of targets, so SSD has a great improvement in the accuracy of multi-scale object detection and is much better for small targets.

Subsequent studies basically belong to the YOLO family. YOLO v2 [13] uses DarkNet19 as the feature extraction network, which is faster than the VGG-16 used by YOLO v1. Combined with methods such as Word Tree, this expands YOLO v2 to thousands of detection categories. YOLO v3 [14] replaces the feature extraction network with DarkNet53, replaces Softmax with Logistic for object classification, and uses three different scales of feature maps to detect objects with different sizes. yolo v3 has a higher accuracy than SSD some, slightly inferior than Faster R-CNN, but at least two times faster than SSD, RetinaNet and Faster R-CNN.

YOLO v4 [15], on the other hand, introduces tricks from the latest research in the field of deep learning in recent years. On the input side, Mosaic data augmentation, cmBN, and SAT self-adversarial training are introduced. On the feature extraction network, YOLO v4 combines various new approaches, including CSPDarknet53 , mish activation function, dropblock. In the detection head, the SPP module is introduced, drawing on the FPN+PAN structure. In the prediction phase, CIOU is adopted as the bounding box loss function of the network, while replacing NMS with DIOU_NMS, etc. YOLO v5, for which the source code is published but no paper has yet been published, is similar to YOLO v4 in that it heavily integrates the latest tricks in computer vision, thus significantly improve detection performance. Compared to YOLO V4, YOLO V5 is slightly inferior in performance, but it is far more flexible and faster than YOLO V4, and has an extreme advantage in rapid deployment of models.

The major drawbacks of anchor-based surface defect detection algorithm are listed as follows.

(1) Sensitivity to anchor hyperparameters. The size, number, and aspect ratio of the anchor have a significant impact on detection performance. If a fixed anchor is used the universality of the detector is greatly compromised, so the size and aspect ratio of the anchor has to be reset for different tasks.

(2) Sample imbalance. In order to de-match the real boxes, a large number of anchors need to be generated, but most of the anchors are marked as negative samples during training and do not make full use of the fore-ground, so it causes the problem of extreme sample imbalance.

(3) High resource consumption during training. The network needs to calculate the IoU (Intersection over Union) of all anchors with real frames during training, which consumes a lot of computer memory and time.

2.3. Anchor-Free Surface Defect Detection Algorithm

Anchor-based detection algorithms are computationally complex due to too many anchors, where a large number of hyperparameters can affect the model performance. The recent anchor free technique, on the other hand, discards anchor and accomplishes detection by identifying key points, which greatly reduces the number of network hyperparameters.

ISSN: 2688-8645

CornerNet [16] is a pioneering work in the anchor free technical line, which proposes a novel approach to object detection by transforming the network's detection of target bounding boxes into the detection of a pair of key points (i.e., the upper left and lower right corners), by detecting objects as pairs of key points without designing anchor boxes as a priori boxes.

CenterNet [17] is another anchor free algorithm, which has a very simple structure and discards the idea of two key points in the upper left and lower right corners, but directly detects the center point of the target, other features such as size, 3D position, orientation, and even pose can be regressed using the image features at the center point location, which is truly anchor free.On the COCO dataset, CenterNet achieves 47.0% AP. compared to anchor-based one-stage and two-stage detection algorithms, CenterNet has a lot of speed and accuracy improvement.

The FSAF [18] network proposes an FSAF module for training the anchor free branch in the feature pyramid, allowing each object to automatically select the most appropriate feature. In this module, the size of the anchor box no longer determines which features are selected for prediction, making the size of the anchor an irrelevant variable and automating the model to learn to select features.

FCOS [19] network is a pixel-by-pixel object detection algorithm based on FCN (Fully Convolutional Networks), which implements the solution of anchor free and proposal free, and proposes the idea of Center ness. The algorithm completely avoids the complex operation of anchor by removing the anchor and saves a large amount of memory occupation during training, reducing the total training memory occupation space by a factor of about 2. FCOS outperforms existing one-stage detectors, while FCOS can also be used as an RPN in two-stage detectors Faste r R-CNN, and both outperform the existing one-stage detectors to a large extent. and are largely superior to RPN.

3. Problems and Solutions

Machine vision-based metal workpiece surface defect detection technology has achieved satisfactory results in theoretical research, but there are still some difficulties when applying these technology in industrial applications for practical use. The existing problems and solution ideas are summarized as follows.

3.1. Small Sample Problem

In the actual image acquisition scenario, the defect samples are few. The most critical problem faced by metal workpiece surface defect detection is the small sample problem compared to the over 14 million image samples in the ImageNet dataset. In many real industrial scenarios there are even only a few dozen defect samples. In addition, the variety and forms of metal workpiece surface defects make the extraction of defect features less efficient, while the model cannot correctly identify newly generated defect types, which is insufficient for training using deep learning methods.

To solve the small sample problem, the main solution ideas are.

(1) Data widening [20]. Various image processing operations such as mirroring, rotating, panning, warping, and distorting can be applied to the original defective samples to generate more samples. Or we can fuse individual defects to superimpose on defect-free samples to form defect samples.

(2) Transfer learning [21]. Due to the large number of parameters in deep learning neural networks, direct training of the network using small samples can lead to overfitting easily. With the help of transfer learning, the common feature data and weight information in the pre-trained model can be used to fine-tune the detection model making it suitable for a specific scenario.

ISSN: 2688-8645

3.2. Real-time Issues

The real-time performance of many surface defect detection algorithms is still some distance away from the actual application, such as the two-stage detection algorithm represented by Faster R-CNN, which is difficult to meet the industrial application scenario in real-time. In practice, the one-stage detection algorithm, represented by the YOLO series, is generally chosen to ensure real-time detection while giving up a certain accuracy rate.

In order to improve the execution speed of the detection algorithm, the main solution idea is model lightweighting [22]. Using lightweight networks, such as MobileNet, the backbone structure of the original original detection algorithm is optimized to simplify the model parameters.

4. Conclusion

Throughout the research on machine vision in the field of metal workpiece surface defect detection, conventional machine vision methods and deep learning methods have their own advantages, as shown in the detailed comparison in Table 1.

Table 1. Comparison of conventional machine vision and deep learning for metal workpiece surface defect detection

Comparison Items	Conventional machine vision	Deep learning
approach	Use feature extraction algorithms (e.g. VJ Detector, HOG Detector, DPM Detector) to extract defective features in the image and then classify	End-to-end feature extraction and classification based on convolutional neural networks
essence	Manual extraction of features	Automatic feature learning from a large number of sample images
prerequisite	High requirements for image quality with clear contours and single defects	A sufficiently large sample of data
adaptivity	Poor, often requiring algorithm parameters adjustment when image quality becomes poor or defect types vary	Good, able to cope with some degree of image quality variation and different defect types

In contrast to conventional machine vision methods that handle the defect detection task in multiple steps, deep learning-based methods unify the feature extraction and classification process into an end-to-end job. With the rapid development of deep learning, not only many improved models based on CNN have been proposed, but also the Transformer model in the field of natural language processing has also started to be applied in the field of computer vision and achieved good results. Therefore, absorbing the improvement ideas of these models and gradually improving the overall effect of the algorithm while taking into account the accuracy and real-time, and being able to be practically applied in industrial scenarios, will be the further research direction for metal workpiece surface defect detection.

Acknowledgments

Research Projects of Wenzhou Polytechnic (WZY2020013).

References

[1] Tang Bo, Kong Jianyi, Wu Shiqian. Review of surface defect detection based on machine vision [J]. Journal of Image and Graphics, 2017, 22(12): 1640–1663.

[2] WANG Q, KIM M-K, SOHN H, et al. Surface flatness and distortion inspection of precast concrete elements using laser scanning technology[J]. Smart Structures & Systems, 2016, 18(3): 601–623.

[3] WALSH J, MAHONY N O, CAMPBELL S, et al. Deep Learning vs. Conventional Computer Vision[C]// Computer Vision Conference (CVC) 2019. , 2019.

[4] VIOLA P, JONES M. Rapid object detection using a boosted cascade of simple features[C]// Proceedings of the 2001 IEEE Computer Society Conference on Computer Vision and Pattern Recognition. CVPR 2001. , 2001: I–I.

[5] DALAL N, TRIGGS B. Histograms of oriented gradients for human detection[C]// 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05). , 2005: 886–893 Vol 1.

[6] FELZENSZWALB P, MCALLESTER D, RAMANAN D. A discriminatively trained, multiscale, deformable part model[C]// 2008 IEEE Conference on Computer Vision and Pattern Recognition. , 2008: 1–8.

[7] GIRSHICK R, DONAHUE J, DARRELL T, et al. Rich Feature Hierarchies for Accurate Object Detection and Semantic Segmentation[C]// 2014 IEEE Conference on Computer Vision and Pattern Recognition. , 2014: 580–587.

[8] HE K, ZHANG X, REN S, et al. Spatial Pyramid Pooling in Deep Convolutional Networks for Visual Recognition[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2015, 37(9): 1904–1916.

[9] GIRSHICK R. Fast R-CNN[C]// 2015 IEEE International Conference on Computer Vision (ICCV). , 2015: 1440–1448.

[10] REN S, HE K, GIRSHICK R, et al. Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2017, 39(6): 1137–1149.

[11] REDMON J, DIVVALA S, GIRSHICK R, et al. You Only Look Once: Unified, Real-Time Object Detection[C]// 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). , 2016: 779–788.

[12] LIU W, ANGUELOV D, ERHAN D, et al. SSD: Single Shot MultiBox Detector[C]// European Conference on Computer Vision. , 2016.

[13] REDMON J, FARHADI A. YOLO9000: Better, Faster, Stronger[C]// 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). , 2017: 6517–6525.

[14] REDMON J, FARHADI A. YOLOv3: An Incremental Improvement[J]. arXiv e-prints, 2018.

[15] BOCHKOVSKIY A, WANG C-Y, LIAO H-Y M. YOLOv4: Optimal Speed and Accuracy of Object Detection[J]. arXiv:2004.10934 [cs, eess], 2020.

[16] LAW H, DENG J. CornerNet: Detecting Objects as Paired Keypoints[J]. International Journal of Computer Vision, 2020, 128.

[17] DUAN K, BAI S, XIE L, et al. CenterNet: Keypoint Triplets for Object Detection[C]. , 2019: 6568–6577.

[18] ZHU C, HE Y, SAVVIDES M. Feature Selective Anchor-Free Module for Single-Shot Object Detection[C]. , 2019: 840–849.

[19] TIAN Z, SHEN C, CHEN H, et al. FCOS: Fully Convolutional One-Stage Object Detection[C]. , 2019: 9626–9635.

[20] TAO X, WANG Z, ZHANG Z, et al. Wire Defect Recognition of Spring-Wire Socket Using Multitask Convolutional Neural Networks[J]. IEEE Transactions on Components, Packaging and Manufacturing Technology, 2018, 8(4): 689–698.

[21] REN R, HUNG T, TAN K C. A Generic Deep-Learning-Based Approach for Automated Surface Inspection[J]. IEEE Transactions on Cybernetics, 2017, PP(99): 1–12.

[22] LI Y, HUANG H, XIE Q, et al. Research on a Surface Defect Detection Algorithm Based on MobileNet-SSD[J]. Applied Sciences, 2018, 8: 1678.