Vol. 14 (2026): Proceedings of the 2nd International Conference on Artificial Intelligence, Modern Engineering and Environmental Sustainability (IC-AIMEES 2026)
E-ISSN: 3105-0522 P-ISSN：3105-0514
DOI: https://doi.org/10.70267/ ic-aimees.20260291298

A Review of Surface Defect Detection Technologies Based on Machine Vision

Ruihong Zhang*

School of Software, Dalian University of Technology, Dalian 116620, China

**Corresponding author: Ruihong Zhang.*

Abstract

Machine vision-based surface defect detection offers the advantages of being non-contact, non-destructive, and highly automated; consequently, it is widely applied across various industrial production processes. This article provides a brief overview of commonly used methods for surface defect detection, evaluation indicators for detection results, and key challenges currently faced. Defect detection methods are categorized into three types: traditional image processing methods, traditional machine learning methods, and deep learning methods. The core principles and representative studies of each method are reviewed, and their respective advantages and limitations are analyzed. We briefly describe the method for evaluating detection results, examine the few-shot learning problem encountered in practical applications, and provide an outlook on feasible pathways for addressing this issue in the future.

Keywords

surface defect detection, machine vision, machine learning, deep learning, few-shot learning

1. Introduction

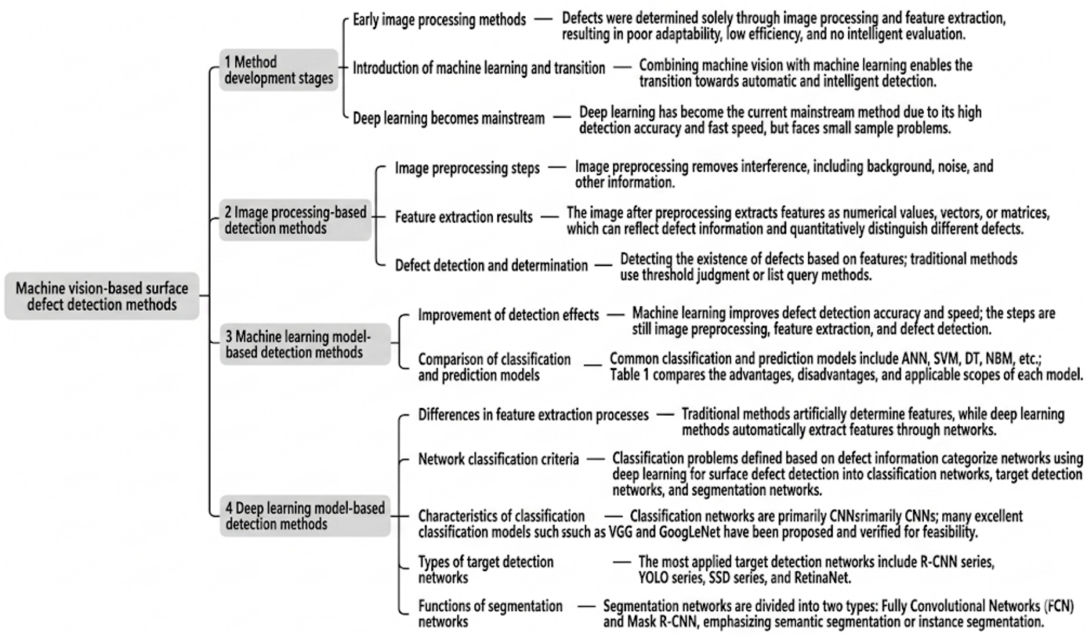
Surface defect detection is a quality control technique that utilizes machine vision technology to automatically identify and assess defects on the surfaces of objects, such as spots, pits, and scratches. This technology is widely used in industrial fields where appearance requirements are stringent, such as for metal surfaces, glass surfaces, and electronic components. It primarily employs image processing algorithms combined with a multi-source collaborative imaging system, utilizing dark-field, bright-field, and transmitted-light illumination techniques to enhance the imaging of surface defects on materials of varying compositions. Machine vision-based surface defect detection is crucial for effectively controlling product quality in automated production. Meanwhile, visual defect detection technology is widely applied across various industrial scenarios, distinguished by its outstanding advantages—including low cost, non-contact operation, non-destructive nature, and safety and reliability.

Surface defect detection primarily addresses three issues: first, determining whether defects exist on the product surface; second, identifying the specific type of defect; and third, locating the position and size of the defect. These three problems progress in a hierarchical manner; however, in actual testing, a single method—after simple adjustments—can simultaneously resolve all three types of problems [1]. This paper reviews currently employed detection methods, quantitative evaluation metrics for detection results, and common issues encountered during the detection process—along with their corresponding solutions—with a particular focus on strategies for addressing the few-shot learning problem.

2. Surface Defect Detection Methods Based on Machine Vision

With the advancement of applying computer vision techniques to image processing, methods for identifying surface defects using machine vision have emerged. Technologies for surface defect recognition based on machine vision can be broadly summarized into three stages: Initially, methods relied solely on image processing; subsequently, these evolved into approaches that combined image processing with machine learning. Today—following the rapid advancements in artificial intelligence—the most sophisticated and widely adopted methods integrate deep learning with machine vision. The complexity of these three methods increases progressively, but their detection effectiveness also improves accordingly. Figure 1 provides an overview of various machine vision-based surface defect detection methods; the core concepts and relevant research content of each method are introduced individually below.

Figure 1: Framework of Surface Defect Detection Methods Based on Machine Vision



2.1 Surface Defect Detection Methods Based on Image Processing

This method is the earliest surface defect detection method. The core idea is relatively simple, which is mainly divided into three steps: The first step is image preprocessing, removing interference in the image, including removing background debris, reducing blur and noise in the image, and making defect regions more visible; The second step is feature extraction, that is, extract the features that can reflect the defect from the processed image, such as the color, shape, size, etc. of the defect, and convert these features into information that can be recognized by the computer; The third step is defect identification. By setting a fixed standard (such as a threshold) or using a rule-based matching methods, it is used to judge whether the extracted features are defects.

However, because this is an early method, the technology has many shortcomings. The literature [2] explains the traditional calculation of this technology. The main limitations are as follows: 1. Image preprocessing: When suppressing high-density noise, it is difficult for traditional filtering algorithms to balance noise suppression and image detail protection. 2. Defective target segmentation: The traditional threshold method lacks the ability to segment weak targets under complex backgrounds and is sensitive to noise. 3. Feature extraction: Traditional methods are difficult to take into account the global and local structural information of the data set, sensitive to noise, and the number of nearest neighbors of some algorithms (such as LPP) is difficult to determine. In addition, the literature also elaborates on new algorithms to solve these problems. For details, see Table 1.

Table 1: Novel Algorithms for Addressing Typical Problems in Machine Vision Technologies Based on Image Processing

Category	Algorithm	Characteristics and Advantages
Filtering	Weighted Median and Mean Filtering Algorithm	The weights can be adaptively determined according to the noise density, effectively suppressing noise while preserving image details.
Object Segmentation	Pixel Search-Based Object Segmentation Algorithm	Addresses the challenge of weak object segmentation through image partitioning and pixel-wise traversal.
Object Segmentation	Improved Data Field-Based FCM Clustering Segmentation Method	Improves the data field, optimizes the membership function, and introduces neighborhood constraints, thereby enhancing the segmentation accuracy and noise robustness for weak objects in complex backgrounds.
Feature Extraction	Fuzzy k-Nearest Neighbor-Based LPP Feature Extraction Algorithm	Adaptively determines the neighborhood range and the number of nearest neighbors (k), thereby improving the performance of the LPP algorithm.
Feature Extraction	Multi-Scale Global–Local Structure Preserving Embedding Feature Extraction Framework	Integrates methods such as LPP, LLE, and PCA, reducing sensitivity to noise while effectively preserving both the global and local structural information of the dataset.

The advantage of machine vision technology based on image processing is that it is simple to operate and easy to understand. The disadvantage is that it is difficult to cope with complex situations. For example, due to the challenges of complex texture, uneven lighting and variable defect morphology during fabric defect inspection, even with advanced algorithms, the detection performance is still poor [3].

2.2 Surface Defect Detection Methods Based on Machine Learning Models

With the development of machine learning techniques, machine learning has been applied to defect detection, which has improved the accuracy and detection speed of defect detection to a certain extent; The steps are still image preprocessing, feature extraction and defect detection. Among them, defect detection is completed through the machine learning model. The core difference is that defect detection is no longer a simple threshold judgment, but allows computers to learn to distinguish between defects and normal areas through machine learning models. This change can significantly improve the accuracy and speed of defect detection.

There are four commonly used machine learning models, namely Artificial Neural Network (ANN), Support Vector Machine (SVM), Decision Tree (DT) and Naive Bayes Model (NBM). The comparison of classifiers in surface defect detection methods based on traditional machine learning models is shown in Table 2, which compares the advantages and disadvantages and scope of application of each model [1, 4].

Table 2: Comparison of Classifiers Used in Traditional Machine Learning-Based Surface Defect Detection Methods

Model	Advantages	Disadvantages	Applicable Scenarios
Artificial Neural Network (ANN)	Possesses strong nonlinear information processing capabilities and self-learning ability.	Lacks interpretability and requires relatively high training costs.	Various defect detection and classification tasks.
Support Vector Machine (SVM)	Exhibits strong generalization capability.	Detection performance is sensitive to parameter settings and kernel functions. A single SVM can only solve binary classification problems.	Situations with relatively small training datasets.
Decision Tree	Highly interpretable and computationally efficient.	Prone to overfitting.	Situations where different categories of training samples are relatively balanced.
Naïve Bayes Model	Requires estimation of only a small number of parameters, is less sensitive to missing data, and demonstrates high robustness.	Assumes statistical independence among features, an assumption that is often difficult to satisfy in practice.	Situations involving a certain number of defects with unknown patterns or forms.

On the whole, the advantage of this method is that it has higher accuracy and fast detection speed than simple image processing, and it can handle some products with relatively complex surfaces; However, the disadvantage is that it is still necessary to extract the defect characteristics manually, and the extensive manual feature engineering is still required. Moreover, when encountering products with many types of defects and extremely complex surfaces, the detection effect is still not ideal.

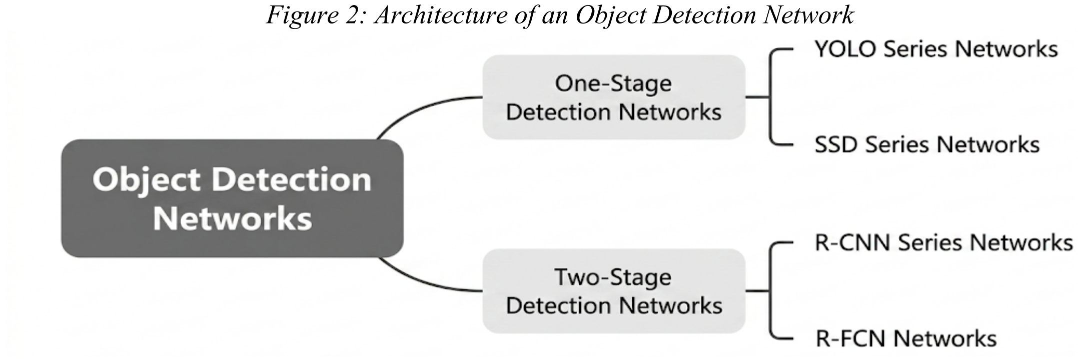
2.3 Surface Defect Detection Methods Based on Deep Learning Models

In recent years, with the breakthrough of deep learning technology, especially the wide application of convolutional neural networks (CNN), surface defect detection methods based on deep learning have become the mainstream methods of surface defect detection. Traditional machine learning methods rely heavily on manually designed features, while deep learning greatly reduces the dependence on manual experience through end-to-end automatic learning of hierarchical features. It allows the computer to automatically extract defect features from images through deep learning models, and can also complete the classification, localization and segmentation of defects at the same time, which greatly improves the detection accuracy and efficiency. According to the actual needs of testing, commonly used deep learning models are divided into three categories, which correspond to the three problems of surface defect detection [1], The following will introduce these three types of models and the problems they can solve one by one.

The first category is a model based on classification networks, and the deep network used for defect classification is mainly convolutional neural networks (CNN). This method mainly solves the problem of "defect types". For example, the literature [5] proposes a multi-scale CNN detection network based on deep learning for the surface defect detection of aluminum profiles of different sizes. It uses 10 kinds of aluminum profile surface defect images to train and evaluate the network to verify the accuracy of the network's aluminum profile defect detection. The advantage of this practice is that it can automatically extract the characteristics of each defect and accurately distinguish the types of defects. However, the disadvantage is that the training requires a large number of samples and the training time is long.

The second category is a model based on target detection, which mainly solves the problem of "defect localization". In recent years, the most widely applied object detection networks include the Region-Convolutional Neural Network (R-CNN) series, the YOLO (You Only Look Once) series, the SSD (Single Shot MultiBox Detector), and the Region-based Fully Convolutional Network (R-FCN). Figure 2 below illustrates the structural frameworks of these four mainstream object detection networks [6–9].

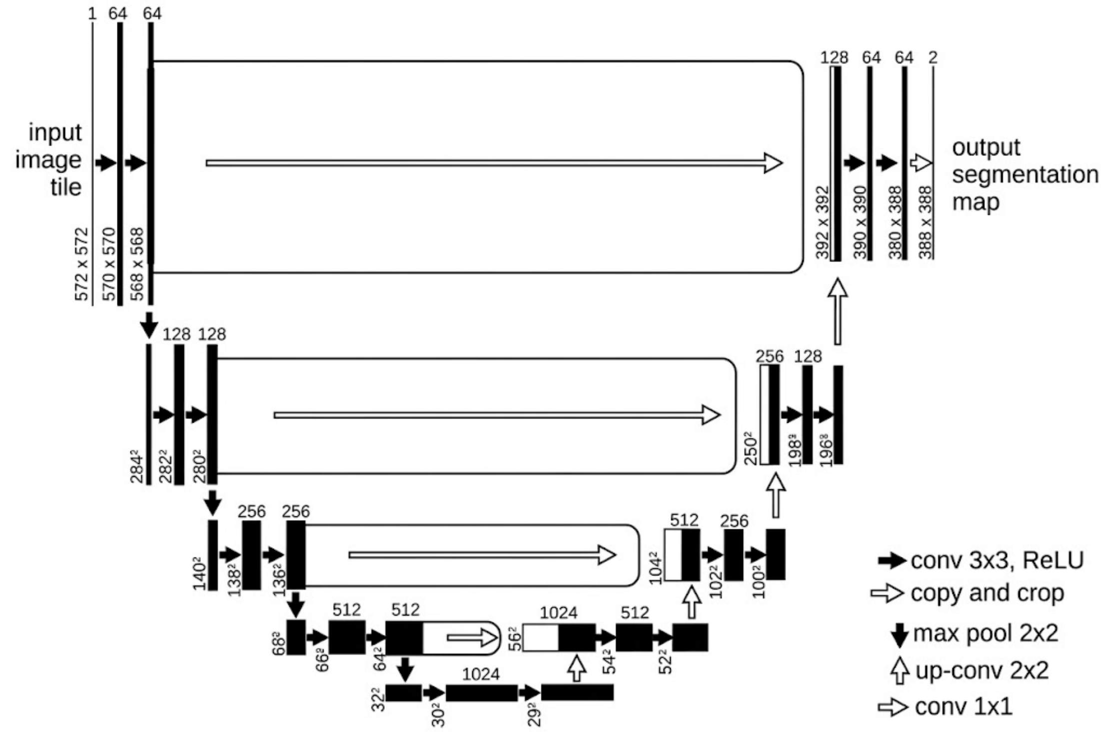
Figure 2: Architecture of an Object Detection Network



The following is an example of completing the actual task through the YOLO model or the Faster R-CNN model. Literature [10] in response to the challenges of various types of strip steel surface defects, complex morphology, large background interference, etc., the structure of the YOLO target detection network has been improved. The introduction of deeper convolutional modules and attention mechanisms enhances the feature extraction ability and the adaptability of the model to different defects, and significantly improves the generalization performance of the network. The improved network was applied to the detection tasks of six typical strip steel surface defects (such as scratches, oxidation spots, holes, etc.). Through a large number of experimental verifications, the results show that the model not only has high detection accuracy, but also has a fast inference speed, which can effectively meet the dual requirements of real-time and reliability in the industrial field; Literature [11] In view of the key issue in the monitoring of the safety status of shield tunnels - the detection of tunnel lining defects, a defect identification and localization method based on the complex background of Faster R-CNN is proposed. By building a deep learning model suitable for the tunnel environment, this method effectively responds to the identification challenges of complex background interference, illumination variations and different defect forms. Research systematically evaluates the performance of the model from multiple dimensions such as detection accuracy, detection efficiency and localization accuracy. The experimental results show that while maintaining a high detection accuracy, the proposed method significantly improves the detection speed and achieves pixel-level precise localization, which fully proves that the model has the significant advantages of fast response and efficient detection in practical engineering applications.

The third category is a model based on the segmentation network. Semantic Segmentation Network is an end-to-end detection network. Semantic segmentation classifies each pixel in the image and marks different objects in the image into different colors, so as to obtain the label map required for the research problem [12]. Next, we will mainly introduce the U-Net. It is a network improved on the Full Convolutional Network (FCN) [13], which is called U-Net because of its U-shaped structure. The network consists of a symmetrical compression path (encoder) and an expansive path (decoder). It achieves accurate pixel-level classification by combining high-resolution features and contextual information. Figure 3 shows the basic structure of the U-Net.

Figure 3: Architecture of the U-Net Network



Literature [12] In order to reduce the leakage rate of artificial surface defect detection and improve the detection effect of surface crack defects in small samples, a detection method based on the semantic segmentation network U-Net is proposed. First, the data of small sample surface crack defect data is enhanced, and then the BN layer is added to the U-Net for batch normalization processing and enhance the generalization ability of the network. Finally, the U-Net and the residual network are combined to avoid network degradation and gradient vanishing, deepen the depth of the network to enhance the feature extraction ability, improve the network detection effect, and conduct verification experiments by adding different residual modules. Experiments show that the defect detection effect of the U-Net is improved after adding the BN layer, and the detection effect is further improved after adding 3, 5 and 7 residual modules. The results show that the U-Net with 7 residual modules suffered from overfitting, and the effect of adding 5 residual modules is the best, indicating that optimal performance can be achieved by appropriately increasing the network depth.

The advantages of this method are high detection accuracy, strong adaptability, can handle the defect detection of various complex products, and especially have good detection effect on small sample surface defect data sets. It is the most widely used method at present; The disadvantage is that it requires a large number of samples for training, and the model structure is complex and difficult to operate. It takes a long time to train the model.

3. Quantitative Evaluation Metrics for Defect Detection Results

When the model is well trained, its performance needs to be evaluated. At present, the defect detection method based on machine vision continues to follow the evaluation indicators in general target classification, detection and segmentation. How to choose the appropriate evaluation indicators as the basis for model judgment is a key issue, and the measurement method needs to be made according to actual needs. The following focuses on the most commonly used quantitative evaluation indexes, including Precision, Recall and Accuracy.

First, we introduce the confusion matrix, which is frequently employed in model evaluation; the relationships among its key parameters—TP (True Positive), FN (False Negative), FP (False Positive), and TN (True Negative)—are illustrated in Table 3 [14]. The formulas for calculating Precision, Recall, and Accuracy are as follows:

- Precision=TP/(TP+FP)
- Recall=TP/(TP+FN)
- Accuracy=(TP+TN)/(TP+FN+FP+TN)

Accuracy refers to the proportion of correctly detected samples relative to the total number of samples tested. For instance, if 100 products are tested and 95 of the detection results are correct, the accuracy is 95%. Precision refers to the proportion of truly defective samples among those classified as "defective." For instance, if 10 products are identified as defective, and 8 of them are genuinely defective, the precision is 80%. Recall refers to the proportion of all truly defective samples that are correctly detected. For instance, if there are 10 genuinely defective products but only 7 are detected, the recall is 70%.

Based on the three categories of metrics mentioned above, the F1-Score was derived; this represents the harmonic mean of Precision and Recall, defined by the formula F1 = 2 × (Precision × Recall) / (Precision + Recall), and serves to comprehensively reflect the model's ability to strike a balance between precision and recall.

In practical industrial scenarios, a single metric often proves insufficient to comprehensively reflect model performance; consequently, multi-dimensional evaluations—typically incorporating ROC curves [15] and P-R curves [16]—are frequently employed to provide a more comprehensive evaluation of model performance. Figures 4 and 5 illustrate the typical shapes of ROC and P-R curves, respectively.

Table 3: Confusion Matrix

	Predicted Positive	Predicted Negative
Actual Positive	True Positive (TP)	False Negative (FN)
Actual Negative	False Positive (FP)	True Negative (TN)

Figure 4: Typical ROC Curve

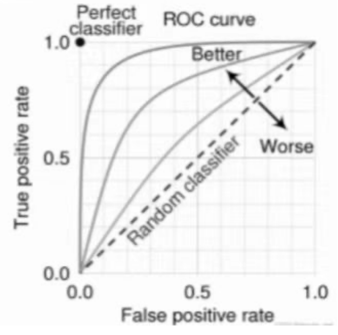
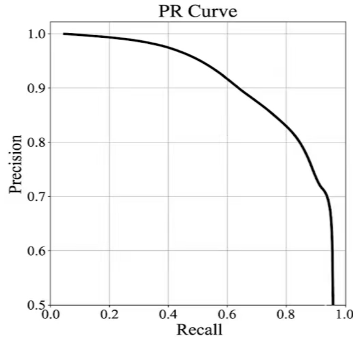


Figure 5: Typical Precision–Recall (P–R) Curve



4. Key Issues and Future Perspectives in Defect Detection

Based on a comprehensive review of the three aforementioned methods, deep learning-based defect detection yields the best overall performance; however, the most significant challenge encountered in practical applications is the few-shot learning problem. In many industrial applications, defect samples are difficult to obtain. For example, some precision instruments and high-end castings have extremely low defect rates during production, making it very difficult to collect a large number of defect samples. Deep learning models require a large volume of samples to complete training; insufficient samples result in a decline in the model's detection accuracy, rendering it unable to meet practical detection requirements. Addressing this issue, future research could proceed along the following lines:

(1) The first method is data augmentation, which involves employing technical means to increase the quantity of existing samples—for instance, by rotating, scaling, or cropping images of defects to generate new samples, thereby providing the model with a larger volume of training data.

(2) The second approach is transfer learning, which involves utilizing a pre-trained model—combined with a small number of new defect samples—to rapidly complete the training of a new model, thereby eliminating the need to train from scratch.

(3) The third type is unsupervised learning. This method does not require samples of defective items; training can be completed using only samples of normal products. By enabling the model to learn the characteristics of normal productss, it determines whether the item being inspected is consistent with the standard—if it is not consistent, it is classified as defective.

5. Conclusions

Through the summary and comparative analysis of the research results in the field of surface defect detection in recent years, in order to complete the task of defect detection, from the first image processing method, to the later machine learning method, to today's mainstream deep learning method, the detection accuracy and robustness have been continuously improved. However, even the most advanced deep learning technology still has shortcomings, especially in industrial sites with small samples, unbalanced categories and high real-time requirements. In the future, we may be able to solve the problem by improving the algorithm. Through the above summary and analysis, I hope to provide valuable reference.

References

[1] Yang, Z. Q., Zhang, M. X., Chen, Y. S., et al. (2023). Research progress on surface defect detection methods based on machine vision. Modern Manufacturing Engineering, (4), 143–156. https://doi.org/10.16731/j.cnki.1671-3133.2023.04.020

[2] Zhao, J. A. (2016). Research on theories and methods for workpiece surface defect detection based on image processing (Master's thesis, Southeast University).

[3] Han, J. Y., Cao, J. T., Wang, H. N., et al. (2022). A review of fabric defect detection methods based on computer vision. Journal of Liaoning Petrochemical University, 42(1), 70–77.

[4] Yang, J. F., Qiao, P. R., Li, Y. M., et al. (2019). A review of machine learning classification problems and algorithms. Statistics & Decision, 35(6), 36–40. https://doi.org/10.13546/j.cnki.tjyjc.2019.06.008

[5] Wei, R., & Bi, Y. (2019). Research on recognition technology of aluminum profile surface defects based on deep learning. Materials, 12(10), 1–14.

[6] Girshick, R., Donahue, J., Darrell, T., et al. (2015). Region-based convolutional networks for accurate object detection and segmentation. IEEE Transactions on Pattern Analysis and Machine Intelligence, 38(1), 142–158.

[7] Redmon, J., Divvala, S., Girshick, R., et al. (2016). You only look once: Unified, real-time object detection. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR) (pp. 779–788). IEEE.

[8] Liu, W., Anguelov, D., Erhan, D., et al. (2016). SSD: Single shot multibox detector. In B. Leibe, J. Matas, N. Sebe, & M. Welling (Eds.), Computer Vision – ECCV 2016 (pp. 21–37). Springer.

[9] Dai, J., Li, Y., He, K., et al. (2016). R-FCN: Object detection via region-based fully convolutional networks. In Advances in Neural Information Processing Systems 29 (NIPS 2016) (pp. 379–387). Curran Associates, Inc.

[10] Li, J., Su, Z., Geng, J., et al. (2018). Real-time detection of steel strip surface defects based on an improved YOLO detection network. IFAC-PapersOnLine, 51(21), 76–81.

[11] Xue, Y., & Li, Y. (2018). A fast detection method via region-based fully convolutional neural networks for shield tunnel lining defects. Computer-Aided Civil and Infrastructure Engineering, 33(8), 638–654.

Funding

This research received no external funding.

Conflicts of Interest

The authors declare no conflict of interest.

Acknowledgment

This paper is an output of the science project.

Open Access