A Survey: Industrial Anomaly Detection based on Data Mining

Jinrui Li

Abstract

Industrial defect detection plays a crucial role in modern manufacturing. Identifying and addressing inferior products contributes to enhancing product quality, strengthening product competitiveness, and increasing customer satisfaction. Existing surveys of industrial defect detection are relatively scarce and struggle to reflect the latest development trends. Therefore, this article provides a more detailed and in-depth survey of industrial defect detection technologies. The article first reviews the development history of industrial defect detection methods. It then covers three aspects: the concept of general anomalies, concepts related to image anomaly detection, and industrial defects, providing an overview of industrial defect detection in these areas. It also summarizes the current state of development, as well as the advantages and disadvantages of each aspect. Additionally, the article identifies the limitations of industrial detection methods in practical industrial applications. Finally, it looks forward to the future development trends and potential research directions in this field, aiming to inspire future research.

Keywords: Data mining, industrial design, abnormal detection, product quality

1 Introduction

With the rapid development of global manufacturing and the continuous improvement of automation levels, high-efficiency and high-quality production have become the key to corporate competitiveness [1]. In this context, industrial defect detection has become increasingly important as an important link in ensuring product quality [2]. Traditional defect detection methods mostly rely on manual visual inspection. Still, this method is inefficient [3], has accuracy limited by the experience and subjective judgment of the inspector, and is easily affected by fatigue and distraction. Therefore, finding more efficient and accurate industrial defect detection methods has become important in industrial automation and intelligence.

With the rapid improvement of computing power and the development of artificial intelligence technology, especially the successful application of deep learning in image processing [4] and pattern recognition [5], industrial defect detection possibilities have been provided. Deep learning technology [6], especially the application of convolutional neural networks (CNN) [7][8], has shown its excellent image recognition capabilities in many fields. These technologies can automatically learn complex features from large amounts of image data, greatly improving the accuracy and efficiency of detection, especially showing great potential in dealing with complex or subtle defects.

Nonetheless, applying deep learning technology to industrial defect detection still faces many challenges. The first is the reliance on large amounts of labeled data, which is often difficult to meet in industrial applications. The second is the generalization ability of the model in different production environments, which requires the algorithm not only to perform well in a specific environment but also to adapt to changing industrial production conditions. In addition, the demand for real-time detection also places higher requirements on the computational efficiency of the algorithm. To solve these problems, researchers have begun exploring new learning paradigms such as semi-supervised, transfer, and self-supervised learning and try to apply these methods to industrial defect detection.

This article aims to provide a comprehensive review of industrial defect detection technology based on deep learning. We first reviewed the development history of industrial defect detection methods, from early manual detection to today's deep learning technology. This development process not only the progress of computing technology andening of pattern recognition theory. Subsequently, this article briefly introduces the specific application of deep learning technology in industrial defect detection, including common network architectures, learning strategies, and application cases of these methods in actual industrial environments. We particularly focus on comparing the effectiveness of different deep learning methods in dealing with various industrial defects and analyze the advantages and limitations of these methods.

Furthermore, we discussed the challenges faced by industrial defect detection in practical applications, such as insufficient data, model generalization capabilities, and real-time detection requirements, and proposed possible solutions and research directions. Finally, this article looks forward to the future development trends in this field, including the combination of deep learning technology with other cutting-edge technologies (such as edge computing cloud computing), as well as potential research directions such as adaptive learning and autonomous decision-making in intelligent manufacturing, aiming to provide Future research on industrial defect detection provides deeper insights.

2 Related Methodology

2.1 Anomaly Detection

2.1.1 Theoretical Framework of Anomaly Detection

In industrial defect detection, anomaly detection is a crucial link, identifying products or processes that do not meet expected production standards or quality requirements. Abnormalities can manifest as size, color, texture, or functionality deviations. In theory, anomaly detection is identifying rare items, events, or observations in data that deviate significantly from most of the data [9][10]. The core challenge of this detection lies in defining "normal" and "abnormal," which depends not only on the specific application scenario but also on the characteristics of the data set.

2.1.2 Anomaly Detection Methodology

In deep learning, neural network-based anomaly detection methods have made significant progress. These methods can be broadly classified into supervised learning, unsupervised learning, and semi-supervised and self-supervised learning. Supervised learning requires large amounts of labeled data to train a model to distinguish between normal and abnormal conditions, such as using convolutional neural networks for image classification. Unsupervised learning methods, without labeled data, attempt to learn the normal patterns of the data and identify anomalies that deviate from these patterns, such as autoencoders [11] and clustering algorithms [12]. Semi-supervised [13][14] and self-supervised learning [15] combine the advantages of labeled and unlabeled data and improve the accuracy of unsupervised learning through a small amount of labeled data.

2.1.3 Challenges and coping strategies

The main challenges facing industrial defect detection include high-dimensional data processing, environmental noise and changes, and the need for real-time detection. Processing high-dimensional data requires algorithms that can effectively extract key features. Environmental noise and lighting changes in industrial environments may affect detection accuracy. In addition, since industrial applications usually require fast response, this places requirements on the computational efficiency of the algorithm. To address these challenges, researchers are exploring more efficient data processing methods, improving algorithms to adapt to environmental changes, and optimizing algorithms to meet real-time processing needs.

2.1.4 Development trends and future directions

The current main development trends of deep learning in anomaly detection include ensemble learning and the use of deep network structures, such as deep autoencoders and deep generative models, to improve the sensitivity and accuracy of detection. In addition, the application of cross-domain transfer learning is also increasing, using models pre-trained in other domains to improve the model's performance in specific industrial applications. At the same time, researchers are also focusing on improving the interpretability of models and optimizing algorithms to meet the needs of real-time processing, which is crucial for industrial applications.

2.2 Image Anomaly Detection

2.2.1 Theoretical Framework for Image Anomaly Detection

Image anomaly detection plays a vital role in industrial defect detection [16][17][18]. This process involves identifying patterns or features in image data that deviate from normal production or quality standards. Image anomalies usually manifest as irregular changes in the size, shape, color, texture, and other visual characteristics of products or materials. In theory, this detection focuses on identifying a small number of abnormal images from many normal images. Image anomaly detection is more complex than general data anomaly detection because it needs to process high-dimensional data and distinguish subtle visual differences while excluding the influence of external factors such as lighting or angle changes.

2.2.2 Methodology for Image Anomaly Detection

Image anomaly detection methods can be roughly divided into traditional image processing technology and deep learning-based methods. Before deep learning became popular, traditional methods, such as threshold segmentation, edge detection, and texture analysis, were widely used. These methods are often based on hand-designed features and heuristics. However, with the development of deep learning technology, methods based on convolutional neural networks (CNN) have become the mainstream of image anomaly detection. These deep learning methods can automatically learn complex feature representations from large amounts of data, improving detection accuracy and efficiency. In addition to this, there are autoencoder-based reconstruction error methods that identify anomalies by learning the compression and reconstruction of normal images and generative adversarial network (GAN)-based methods [19][20] that use generative networks to simulate normal data distributions, thereby identifying abnormal patterns.

2.2.3 Image Anomaly Detection Challenges

Image anomaly detection faces several challenges in practical applications. The first is the huge computational burden brought by processing high-resolution images, which challenges the processing speed and efficiency of the algorithm. Secondly, noise, lighting changes, and other interference factors in the industrial environment will affect detection accuracy. In addition, due to the diversity and complexity of industrial image data, the model's generalization ability becomes a key issue. Limitations in training data can cause model performance to degrade when faced with real-world variability. Therefore, developing efficient and robust algorithms that adapt to these changes becomes particularly important.

2.2.4 Development Trend of Image Anomaly Detection

Future image anomaly detection technology will likely achieve breakthroughs in several key areas. The first is the improvement of algorithm efficiency, which includes accelerating the training and inference process of deep learning models so that they can process high-resolution images in real-time. The second is enhancing model generalization ability and robustness, which may be achieved by introducing more advanced network architectures, such as deeper convolutional networks, ensemble learning methods, and attention mechanisms. In addition, improving detection accuracy and model interpretability are also important directions for future development. To this end, researchers may explore combining transfer learning [21][22] and reinforcement learning strategies [23][24], as well as developing advanced data augmentation and simulation techniques to improve the quality and diversity of training data.

2.3 Industrial Defects

2.3.1 Definition and Classification of Industrial Defects

An industrial defect refers to any deviation from predetermined product quality standards that occurs during the manufacturing process. These defects may be caused by factors such as poor quality of raw materials, operating errors on the production line, aging or failure of mechanical equipment, or management errors during the manufacturing process. We can divide the defects into several categories in more detail: surface defects such as scratches, cracks, dents, rust, oil stains, color differences, etc.; dimensional defects, including product sizes that are too large or too small. Typically caused by worn or improperly calibrated equipment; structural defects, such as voids, inclusions, or unfused areas within the material, which can significantly affect the strength and durability of the product; and functional defects, which often manifest as The performance of the product during use does not meet expectations, such as circuit failure in electronic equipment or power transmission problems in mechanical equipment. Properly identifying and classifying these defects is critical to ensuring product quality and production efficiency.

2.3.2 Detection Methods for Industrial Defects

Traditional industrial defect detection methods mainly rely on manual visual inspection and rule-based automated inspection systems. Although manual inspection is flexible, it is inefficient and susceptible to the operator's subjective judgment and fatigue. With the advancement of technology, automated inspection technology has gradually developed, including non-destructive inspection technologies such as visual inspection systems based on image processing, ultrasonic inspection based on sound waves, and radiation inspection based on electromagnetic waves. Based on these technologies, the application of deep learning has brought revolutionary progress to automated detection. The advantages of deep learning, especially convolutional neural networks (CNN), in image recognition and classification, make it effective in identifying defects from complex image data. In addition, with the continuous development of technology, automated inspection systems can detect the presence of defects and classify and evaluate the type and severity of defects.

2.3.3 Challenges

In the practical application of industrial defect detection, many challenges still need to be solved. First, data acquisition and quality control are a major challenge. For machine learning-based methods, high-quality and large amounts of labeled data are essential, but obtaining such data in actual production processes is often difficult and costly [25][26]. Secondly, the generalization ability and adaptability of the model is another important challenge. Due to the diversity of production conditions, a model trained from one production environment may not directly apply to another. In addition to the demand for real-time detection, the detection system must have the ability to respond quickly while ensuring accuracy. This requires detection algorithms that are accurate, efficient, and capable of performing real-time analysis without slowing down the production line.

2.3.4 Future development trends

Looking ahead, industrial defect detection will likely see significant advancements in several key aspects. With the continuous development of deep learning technology, it is expected that more advanced network architectures and learning algorithms will emerge, further improving detection accuracy and efficiency. For example, the model's ability to identify complex defects can be improved by integrating deep learning models or using more complex network structures, such as residual networks or self-attention mechanisms. At the same time, toto improve the model's generalization ability and adaptability, researchers may explore more transfer learning and meta-learning methods. In addition, with the development of edge computing and cloud computing technologies, combining these technologies with industrial defect detection can achieve more efficient data processing and flexible resource allocation. In terms of hardware, higher-performance computing platforms, and more accurate sensors will also further enhance the capabilities of detection systems. In addition, as augmented reality (AR) [27] and virtual reality (VR) [28] technologies mature, applying these technologies to the process of industrial defect detection and repair can provide more intuitive visual assistance, thereby improving the efficiency and accuracy of inspection and repair.

3 Discussion and Outlook

3.1 Discussion

This article reviews industrial defect detection technology based on deep learning, revealing its importance and application potential in modern manufacturing. Deep learning, especially the superior performance of convolutional neural networks in image processing, has become the key to improving the accuracy and efficiency of defect detection. However, these technologies also face many challenges. First, the success of deep learning models relies heavily on large amounts of high-quality labeled data, which is often difficult to obtain in actual production environments. Second, the model's generalization ability is another key issue, especially in diverse and changing industrial environments. In addition, the demand for real-time detection places higher requirements on the algorithm's computational efficiency and response speed.

Although deep learning shows great potential in industrial defect detection, it is not a panacea. Further research is still needed to improve detection accuracy and reliability in some complex or subtle defect detection scenarios. For example, small cracks or subtle color changes may be difficult for traditional deep-learning models to identify accurately. In addition, the interpretability of deep learning models is also an important research direction, especially in safety-critical applications such as aerospace and medical device manufacturing.

3.2 Outlook

In the future, industrial defect detection technology based on deep learning will continue to develop and may achieve breakthroughs in the following aspects. First, algorithmic innovation will be a key factor driving the development of this field. By introducing more advanced deep learning models, such as variational autoencoders (VAE) and generative adversarial networks (GAN), as well as emerging learning paradigms, such as transfer learning, meta-learning, and self-supervised learning, the model can be further improved in limited or performance in the case of unlabeled data, while enhancing its adaptability and generalization capabilities.

Secondly, developing real-time detection technology will make industrial defect detection more efficient and reliable. With advances in computing hardware, such as more powerful GPUs, dedicated deep learning accelerators, and algorithm optimization, processing high-resolution images and complex data in real-time will become more feasible.

Another important development direction is to improve the interpretability and user-friendliness of the model. In industrial applications, models not only need to provide accurate detection results but also need to be able to explain their decision-making processes to engineers and operators. This is critical to increasing user trust and acceptance of automated systems.

Finally, it is expected that future industrial defect detection will not only be limited to traditional defect identification but will also expand to production process optimization and predictive maintenance. Through continuous analysis of production line data, early warning of potential problems can be achieved, thereby reducing downtime and improving production efficiency and product quality.

4 Conclusion

This article reviews industrial defect detection technology based on deep learning and discusses its applications and challenges in modern manufacturing. We first reviewed the development history of industrial defect detection, from the initial manual detection to modern automation and intelligent technology, especially the introduction of deep learning technology, marking a major change in this field. The advantages of deep learning, especially convolutional neural networks, in image recognition and classification, provide an effective tool for identifying and classifying various industrial defects, significantly improving the accuracy and efficiency of detection.

However, we also point out the challenges of applying deep learning to industrial defect detection, including reliance on large amounts of labeled data and model generalization.

References

[1] Chen Y, Ding Y, Zhao F, et al. Surface defect detection methods for industrial products: A review[J]. Applied Sciences, 2021, 11(16): 7657.

[2] Saberironaghi A, Ren J, El-Gindy M. Defect detection methods for industrial products using deep learning techniques:

A review[J]. Algorithms, 2023, 16(2): 95.
[3] Gao Y, Lin J, Xie J, et al. A real-time defect detection method for digital signal processing of industrial inspection applications[J]. IEEE Transactions on Industrial Informatics, 2020, 17(5): 3450-3459.
[4] Kovásznay L S G, Joseph H M. Image processing[J]. Proceedings of the IRE, 1955, 43(5): 560-570.
[5] Theodoridis S, Koutroumbas K. Pattern recognition[M]. Elsevier, 2006.
[6] LeCun Y, Bengio Y, Hinton G. Deep learning[J]. nature, 2015, 521(7553): 436-444.
[7] Gu J, Wang Z, Kuen J, et al. Recent advances in convolutional neural networks[J]. Pattern recognition, 2018, 77: 354-377.
[8] O'Shea K, Nash R. An introduction to convolutional neural networks[J]. arXiv preprint arXiv:1511.08458, 2015.
[9] Ahmed M, Mahmood A N, Hu J. A survey of network anomaly detection techniques[J]. Journal of Network and Computer Applications, 2016, 60: 19-31.
[10] Bhuyan M H, Bhattacharyya D K, Kalita J K. Network anomaly detection: methods, systems, and tools[J]. Ieee communications surveys & tutorials, 2013, 16(1): 303-336.
[11] Louizos C, Swersky K, Li Y, et al. The variational fair autoencoder[J]. arXiv preprint arXiv:1511.00830, 2015.
[12] Duran B S, Odell P L. Cluster analysis: a survey[M]. Springer Science & Business Media, 2013.
[13] Zhu X J. Semi-supervised learning literature survey[J]. 2005.
[14] Van Engelen J E, Hoos H H. A survey on semi-supervised learning[J]. Machine learning, 2020, 109(2): 373-440.
[15] Jaiswal A, Babu A R, Zadeh M Z, et al. A survey on contrastive self-supervised learning[J]. Technologies, 2020, 9(1): 2.
[16] Czimmermann T, Ciuti G, Milazzo M, et al. Visual-based defect detection and classification approaches for industrial applications—A survey[J]. Sensors, 2020, 20(5): 1459.
[17] Ren Z, Fang F, Yan N, et al. State of the art in defect detection based on machine vision[J]. International Journal of Precision Engineering and Manufacturing-Green Technology, 2022, 9(2): 661-691.
[18] Bai X, Fang Y, Lin W, et al. Saliency-based defect detection in industrial images by using phase spectrum[J]. IEEE Transactions on Industrial Informatics, 2014, 10(4): 2135-2145.
[19] Creswell A, White T, Dumoulin V, et al. Generative adversarial networks: An overview[J]. IEEE signal processing magazine, 2018, 35(1): 53-65.
[20] Goodfellow I, Pouget-Abadie J, Mirza M, et al. Generative adversarial networks[J]. Communications of the ACM, 2020, 63(11): 139-144.
[21] Weiss K, Khoshgoftaar T M, Wang D D. A survey of transfer learning[J]. Journal of Big Data, 2016, 3(1): 1-40.
[22] Zhuang F, Qi Z, Duan K, et al. A comprehensive survey on transfer learning[J]. Proceedings of the IEEE, 2020, 109(1): 43-76.
[23] Li Y. Deep reinforcement learning: An overview[J]. arXiv preprint arXiv:1701.07274, 2017.
[24] Arulkumaran K, Deisenroth M P, Brundage M, et al. Deep reinforcement learning: A brief survey[J]. IEEE Signal Processing Magazine, 2017, 34(6): 26-38.
[25] Huang C C, Lin X P. Study on machine learning based intelligent defect detection system[C]//MATEC Web of Conferences. EDP Sciences, 2018, 201: 01010.
[26] Bertolini M, Mezzogori D, Neroni M, et al. Machine Learning for industrial applications: A comprehensive literature review[J]. Expert Systems with Applications, 2021, 175: 114820.
[27] Carmigniani J, Furht B. Augmented reality: an overview[J]. Handbook of augmented reality, 2011: 3-46.
[28] Wohlgenannt I, Simons A, Stieglitz S. Virtual reality[J]. Business & Information Systems Engineering, 2020, 62: 455-461.