INTERNATIONAL JOURNAL OF ADVANCED INNOVATIVE TECHNOLOGY IN ENGINEERING

Published by Global Advanced Research Publication House Journal

Home page: www.ijaite.co.in

Steel Plate Defect Detection Using Machine Learning and Deep Learning: A Review

[1]Komal Raju Nimsarkar, [2]Supriya Sawwashere, [3]Shrikant V. Sonekar, [4]Ashutosh Lanjewar, [5]Mirza Moiz Baig

[1,2,3,4,5]Department Of Computer Science and Engineering, JD College Of Engineering and Management Katol Road, Nagpur, Maharashtra, India

[1]nimsarkar4@gmail.com, [2]sssawwashere@jdcoem.ac.in, [3]srikantsonekar@gmail.com, [4]aolanjewar39@gmail.com, [5]mirzammb@gmail.com

Article History

Received on: 10 June 2024

Revised on: 28 June 2024

Accepted on: 15 July 2024

Keywords: Steel Plate Defect Detection, Machine Learning, Deep Learning, Feature Extraction, Computer Vision.

e-ISSN: 2455-6491

DOI: 10.5281/zenodo.15289187

Production and hosted by
www.garph.org

ABSTRACT

Steel plate defect detection is a critical quality control process in manufacturing industries to ensure product reliability and safety. Recent advancements in machine learning (ML) and deep learning (DL) have significantly improved the accuracy and automation of defect detection systems. This paper reviews the current approaches and techniques used in steel plate defect detection, highlighting traditional machine learning models such as Random Forest and XGBoost, as well as deep learning architectures like Convolutional Neural Networks (CNNs) and MobileNetV2. We compare different methods, discuss their strengths and limitations, and outline future research directions. This review aims to provide insights into how intelligent systems can be further enhanced for real-world industrial applications.

1. INTRODUCTION

Steel plates are widely used in industries such as construction, automotive, and shipbuilding, where material integrity is critical. Manual inspection methods are often labor-intensive, time-consuming, and prone to human error. With the growing adoption of smart manufacturing, automated defect detection systems powered by machine learning (ML) and deep learning (DL) have emerged as promising solutions. These systems aim to identify surface anomalies like cracks, dents, inclusions, and scratches with higher accuracy and speed.

Traditionally, the control of surface quality is conducted manually, and workers are trained in order to identify the complicated surface defects. Nevertheless, this kind of control is inefficient and time-consuming, and its accuracy of detection is affected by the experience, energy and subjectivity of inspectors. With the aim of overcoming the shortcomings of manual inspection, the automatic detection of surface defects on the basis of machine vision came into being. In the last decade, many approaches [1]-[2] have been utilized for automatic detection of surface defects on steel. The major principle is to utilize the shape or pixel values of steel surface to predict defects; nevertheless, it is time-consuming and complex to set threshold and obtain feature parameters.

This review paper presents a comprehensive overview of various ML and DL-based techniques applied to steel plate defect detection. We aim to summarize the key developments, compare their performances, and highlight the future trends in this rapidly evolving field.

2 Steel Plate Scratch Detection System

The steel plate scratch defect detection system built in this paper is shown in Fig. 1, which mainly includes visual monitoring device, communication network and monitoring center. The visual monitoring device includes CCD (charge coupled device) cameras, light source and power supply. The cameras are set up above the steel plate and are perpendicular to the steel plate plane, and the light source is installed directly above the steel plate. The communication network transmits the image data to the monitoring center. The received images are processed and analyzed by PC (person computer) through intelligent analysis software and various embedded image processing algorithms [9]. Detection results are recorded and stored for qualitative and quantitative analysis of entire steel plate.

Analysis of Features of Roll Marks

Roll marks are a set of uneven defects with periodicity. They are generally due to roll fatigue, insufficient hardness, or foreign matter on the surface of the rolls during rolling operations. The morphological features of roll marks on the same batch are stable and similar; the morphological features of roll marks on different batches can vary due to the repeated rolling of the steel plates. Roll marks are present on both the upper and lower surfaces of plates, mainly at the operating side and the middle position of the plates. They are bright spots observable by the naked eye, and dark spots in the image captured by a camera.

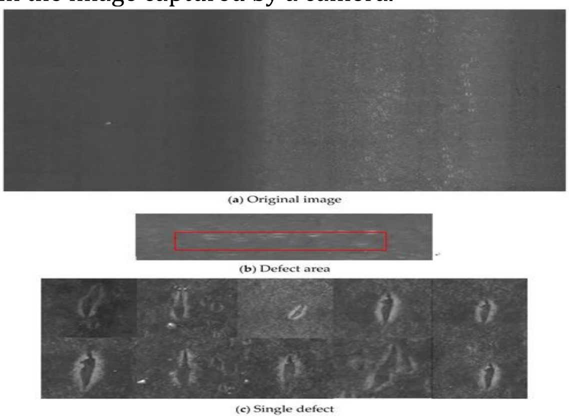

Figure 1: Images of rolling marks

Figure 1 shows an original image of roll marks. Figure 1b shows the defect area and that the roll marks in the frame are arranged periodically. Figure 2 contains details of a single defect, which shows that the morphological features are similar.

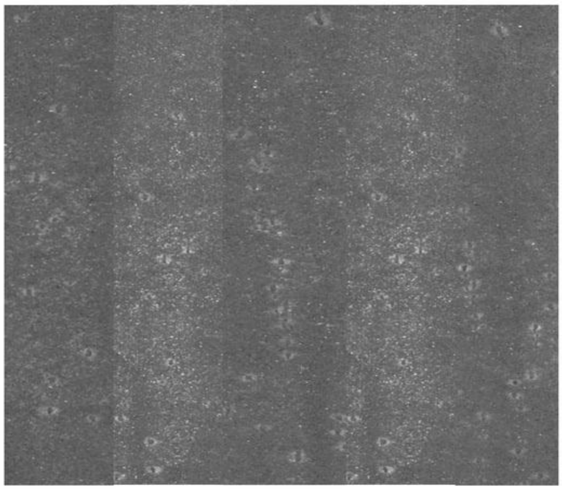

Figure 2: Roll marks on the same batch

Figure 3: Roll marks on different batches.

Roll mark defects are not well-detected because of the greatly different morphological features of roll marks on different batches. The traditional CNN classifies defects by extracted morphological features [6]. Therefore, a CNN can easily misclassify roll marks due to their unfixed morphological features. Consequently, the classification accuracy is not high.
However, as roll mark defects have strong periodicity, their time-sequenced characteristics are suitable for handling by LSTM.

2. RELATED WORK

A. *Machine Learning*

Machine Learning is the general term for when computers learn from data. Machine learning is the application/subset of artificial intelligence. Machine learning centers on the advancement of PC programs, and the essential point is to enable PCs to adapt consequently without human intervention. There are various algorithms that machines can learn. The information that you feed to a machine learning algorithm can be input-output sets or just inputs. Supervised learning calculations require input-output sets (i.e., they require the output). Unsupervised learning requires just the input information (not the output).

Supervised Learning Techniques: Supervised learning techniques learn a function to map the given input to discrete/continuous output space. For example, in a webpage classification problem, the learner approximates a function mapping the feature vector into "Politics" or "Cricket" by reducing the error on training input-output examples. Since the function outputs a discrete value, it is referred to as classification. Consider another example, where the input review text has to be mapped to a score between 1 and 10. In this case, the output is a continuous value, it is referred to as regression. Active Learning is a supervised learning paradigm, which allows querying the user/oracle to obtain the label for an input data point. Typically, it is used in cases where there is a fixed budget for getting labels. The querying must be done on the most informative data points concerning the task at hand.

Unsupervised Learning Techniques: Unsupervised learning techniques model the input data. The input does not include class labels as in the case of supervised learning. For example, to understand user browsing behavior, It would be like to group users based on their online browsing patterns and model each group. It is also called clustering. The grouping is performed concerning an objective function. For example, in the k-means algorithm clustering is done such that the data points within a cluster are more similar and the data points across clusters are less similar. Other unsupervised learning techniques include association rule mining, and blind separation techniques for feature extraction such as principal component analysis, independent component analysis, etc. [9].

Semi-Supervised Learning: Semi-supervised learning (SSL) combines both labeled and unlabeled data to learn an appropriate function for prediction. Typically, it uses a small amount of labeled data and a large amount of unlabeled data. It's a relatively new paradigm and is gaining importance because getting access to completely labeled data is hard and costly. The use of unlabeled data provides regularization, and it has been shown to improve performance. Transductive learning works only on the labeled and unlabeled training data, and cannot handle unseen data. On the other hand, inductive learners can handle unseen data.

B. Classification

Classification is a data mining technique that typically involves three phases, a learning phase, a testing phase, and an application phase. A learning model or classifier is built during the learning phase. It may be in the form of classification rules, a decision tree, or a mathematical formula. Since the class label of each training sample is provided, this approach is known as supervised learning. In unsupervised learning (clustering), the class labels are not known in advance. In the testing phase, test data are used to assess the accuracy of the classifier. If the classifier passes the test phase, it is used for the classification of new, unclassified data tuples. This is the application phase. The classifier predicts the class label for these new data samples. For classification algorithms, the two major problems in classifying a data stream are the infinite length and the concept drift. The first one makes the traditional multi-pass classification algorithms incapable of classifying a data stream for their requirement of infinite storage and a large amount of training time. The second one makes the most static stream classification algorithms incapable of classifying a data stream with concept drifts for the underlying changes that occurred in the stream. For a time, changing data stream, an incremental updating manner of the classifier is very important. A temporal model is used to capture the evolutions of the stream. In general, the classification process is always accompanied by the course of model construction and test. The classification model keeps changing with the progression of the stream. If a static classifier is used to classify an evolving data stream, its accuracy of it will drop greatly. For a sudden burst of concept drift in a time-changing stream, an up-to-date model always provides better accuracy. But for relatively stable time-changing streams, models built with long-term samples will be great [10].

In recent years, advancements in computer technology have expanded the use of vision techniques in defect detection. This involves traditional machine vision algorithms and deep learning. Hu et al.3 introduced the AdaBoost algorithm to enhance the accuracy of detecting uncommon defects on steel plate surfaces. This algorithm improves accuracy by increasing combinable weak classifiers through a filtering mechanism.

Congzhe You et al. (2024) proposed the deep learning model enhanced steel surface defect detection algorithm based on YOLOv8 was introduced to enhance the accuracy of small target detection. This algorithm incorporates an attention-free mechanism to calculate attention-weight, aiding in the extraction of specific feature regions. Additionally, improvements were made to the SPPF module to expand the receptive field and enhance target detection optimization. Experimental evaluations on the NEU-DET dataset demonstrated significant enhancements over the original YOLOv8 algorithm. The improved algorithm exhibited a 9.3 percentage point increase in precision, a 10-percentage point increase in recall, a 4.6 percentage point increase in mAP@0.5, and a remarkable 21.2 percentage point increase in mAP@0.5:0.95. Significant progress has also been made in analyzing the surface data of aluminum sheets. The enhanced algorithm has shown a 6% increase in precision compared to the original YOLOv8 algorithm. Additionally, recall has improved by 3.2%, mAP@0.5 has increased by 4.1%, and mAP@0.5:0.95 has seen a notable rise of 17.4%.

Lu, J. et al. (2024) introduces SS-YOLO (YOLOv7 for Steel Strip), an enhanced lightweight YOLOv7 model. This method replaces the CBS module in the backbone network with a lightweight MobileNetv3 network, reducing the model size and accelerating the inference time. The D-SimSPPF module, which integrates depth separable convolution and a parameter-free attention mechanism, was specifically designed to replace the original SPPCSPC module within the YOLOv7 network, expanding the receptive field and reducing the number of network parameters. The parameter-free attention mechanism SimAM is incorporated into both the neck network and the prediction output section, enhancing the ability of the model to extract essential features of strip surface defects and improving detection accuracy. The experimental results on the NEU-DET dataset show that SS-YOLO achieves a 97% mAP50 accuracy, which is a 4.5% improvement over that of YOLOv7. Additionally, there was a 79.3% reduction in FLOPs(G) and a 20.7% decrease in params.

Wang, S. et. al (2021) proposes a method combining improved ResNet50 and enhanced faster region convolutional neural networks (faster R-CNN) to reduce the average running time and improve the accuracy. Firstly, the image input into the improved ResNet50 model, which add the deformable revolution network (DCN) and improved cutout to classify the sample with defects and without defects. If the probability of having a defect is less than 0.3, the algorithm directly outputs the sample without defects. Otherwise, the samples are further input into the improved faster R-CNN, which adds spatial pyramid pooling (SPP), enhanced feature pyramid networks (FPN), and matrix NMS. The final output is the location and classification of the defect in the sample or without defect in the sample. By analyzing the data set obtained in the real factory environment, the accuracy of this method can reach 98.2%. At the same time, the average running time is faster than other models. [1].

Liu, Yang et. al. (2019) proposed to detect periodic defects, such as roll marks, according to the strong time-sequenced characteristics of such defects. Firstly, the features of the defect image are extracted through a CNN network, and then the extracted feature vectors are inputted into an LSTM network for defect recognition. The experiment shows that the detection rate of this method is 81.9%, which is 10.2% higher than a CNN method. In order to make more accurate use of the previous information, the method is improved with the attention mechanism. The improved method specifies the importance of inputted information at each previous moment, and gives the quantitative weight according to the importance. The experiment shows that the detection rate of the improved method is increased to 86.2%. [2].

Li, Hanlin et. al. (2023) proposed RepBi-PAN fusion network into YOLOv5, enhancing the detection capability for large targets in complex backgrounds. To mitigate issues related to the premature introduction of shallow features and decrease in Precision, we optimized the model structure by incorporating the DenseNet structure into the backbone for improved feature extraction. Additionally, we introduced the Normalized Attention Module (NAM) to enhance the detection capability for small targets. Experimental results demonstrate the effectiveness of the enhanced model, showing a 4.1% increase in mean average precision (mAP), a 3.2% improvement in precision, and a 2.4% enhancement in recall. The improved algorithm outperforms the original in complex backgrounds and recognizing small targets, addressing limitations of the Rep-Bi network. Compared to other YOLO algorithms, our approach achieves optimal values for recall and mAP while maintaining a smaller model size. In comparison with YOLOv8, the improved model surpasses all V8 models, being only 0.5% below the precision of the largest YOLOv8x model. Simultaneously, the improved model is smaller and has fewer parameters compared to all models in the YOLOv8 series, with slightly higher GFLOPs than the smaller v8 models. [3].

Vira Fitriza et. al. (2020) proposed a deep learning CNN with Xception architecture to detect steel defects from images taken from high-frequency and high-resolution cameras. There are two techniques used, and both produce respectively 0.94% and 0.85% accuracy. The Xception architecture used in this case shows optimal and stable performance in the process and its results. [4].

Weidong Zhao et. al. (2021) proposed an urrent detection algorithms for NEU-DET dataset detection accuracy are low, so we choose to verify a steel surface defect detection algorithm based on machine vision on this dataset for the problem of defect detection in steel production. A series of improvement measures are carried out in the traditional Faster R-CNN algorithm, such as reconstructing the network structure of Faster R-CNN. Based on the small features of the target, we train the network with multiscale fusion. For the complex features of the target, we replace part of the conventional convolution network with a deformable convolution network. The experimental results show that the deep learning network model trained by the proposed method has good detection performance, and the mean average precision is 0.752, which is 0.128 higher than the original algorithm. Among them, the average precision of crazing, inclusion, patches, pitted surface, rolled in scale and scratches is 0.501, 0.791, 0.792, 0.874, 0.649, and 0.905, respectively. The detection method is able to identify small target defects on the steel surface effectively, which can provide a reference for the automatic detection of steel defects [5].

Yang, L., et. al. (2023) proposed an automatic detection method for steel plate scratch. irstly, the steel plate image is decomposed by channel and the enhanced image is obtained by the improved MSR (Multi-Scale Retinex) enhancement algorithm. Then, the phase consistency is detected after the Log Gabor wavelet transform and the scratch areas are obtained by the threshold segmentation and intersection of three channels. Finally, the scratch position is identified and the scratch characteristics such as width and length can be calculated. The results show that the minimum error of the characteristics measurement is only 2.28% in the experimental environment and 4.15% in the field environment, and the mean running time is 0.2826 s in the experimental environment and 0.3193 s in the field environment. It verifies that the proposed method is effective and practical [6].

Wang, D. et. al. (2023) first classified the common edge defects and then made a dataset of edge defect images on this basis. Subsequently, edge defect recognition models were established on the basis of LeNet-5, AlexNet, and VggNet-16 by using a convolutional neural network as the core. Through multiple groups of training and recognition experiments, the model's accuracy and recognition time of a single defect image were analyzed and compared with recognition models with different learning rates and sample batches. The experimental results showed that the recognition model based on the AlexNet had a maximum accuracy of 93.5%, and the average recognition time of a single defect image was 0.00 35 s, which could meet the industry requirement. The research results in this paper provide a new method and thought for the fine detection of edge defects in hot rolling strips and have practical significance for improving the surface quality of hot rolling strips. [7].

Zhang,M. et. al. (2021) proposed the image enhancement algorithm based on adaptive threshold gray transformation to enhance the quality of steel surface defect image, and then the image was processed by Gabor filter and image segmentation [8].

Kun Liu et. al. (2020) established a specific template for each defect image, and the test image was decomposed into structural component and texture component. By calculating the index gradient similarity between template and texture component, various defects on the steel plate surface can be detected [9].

Yue Wu et. al. (2021) introduced the advantages of residual structure and feature fusion of YOLOv3 model into the Faster R-CNN model and realized the classification of different defects of steel plate. However, these methods were qualitative analysis, and they all processed the ideal images without considering the image samples under non-ideal conditions such as uneven illumination and blurred target. With the development of the scale and technology of steel industry, there is an increasing demand for quantitative analysis of steel plate surface defects [10].

C. Defect Recognition Algorithm Based on Traditional Machine Learning

The traditional machine learning approach was an epoch-making advancement from manual inspection, and usually starts with the manual design of feature extraction rules, followed by feature extraction, and finally feeds the extracted features into the classifier to achieve the classification of defects. Because of the reliance on manually designed feature extraction rules, it leads to poor robustness and generalization ability of the algorithm and is susceptible to interference and the influence of noise, thus reducing the detection accuracy. The most traditional methods basically only provide a defect classification function and do not perform defect localization or segmentation, which is an incomplete defect recognition process. The machine learning algorithms used for steel surface defect recognition can be broadly classified into texture feature-based methods, shape feature-based methods, and color feature-based methods. However, in the field of steel surface defect detection, since color features mainly refer to grayscale features of the image, and the methods used to extract grayscale features are statistically based, the color feature-based methods were classified here under the texture feature-based methods.

D. Texture Feature-Based Methods

Texture feature-based methods are the most common methods in the field of steel defect detection, which reflects the homogeneity phenomenon in the image and can reflect the organization and arrangement characteristics of the image surface through the grayscale distribution of pixels and their nearby spatial neighborhood [1]. As shown in Figure4, it can be subdivided into statistical-based methods, filter-based methods, structure-based methods, and model-based methods. These four methods can be used in combination or in conjunction with each other to achieve a higher performance. Regarding the literature on texture-based feature methods, these are shown in Table 1.

Statistical-based methods are used to measure the spatial distribution of pixel values, usually by using the grayscale distribution of image regions to describe texture features such as heterogeneity and directionality. Its common statistical methods include histogram, co-occurrence matrix, local binary patterns, etc. In 2015, Chu et al. [2] proposed a feature extraction method based on smoothed local binary patterns, which is insensitive to noise and invariant to scale, rotation, translation, and illumination, so the algorithm can maintain a high classification accuracy for the identification of strip surface defects. In 2017, Truong and Kim [3] proposed an automatic thresholding technique, which is an improved version of the Otsu method with an entropy weighting scheme that is able to detect very small defect areas. Luo et al. [4] proposed a selective local binary pattern descriptor, which was used to extract defect features, and then combined it with the nearest neighbor classifier (NNC) to classify strip surface defects; this algorithm pursued the comprehensive performance of recognition accuracy and recognition efficiency. The following year, Luo et al. [5] also proposed an improved generalized complete local binary pattern descriptor and two improved versions of the improved complete local binary pattern descriptor (ICLBP) and improved the complete noise-invariant local structure pattern (ICNLP) to obtain the surface defect features of the hot rolled steel strip, and then used the nearest neighbor classifier to achieve defect recognition classification, thus achieving high recognition accuracy. Zhao et al., in 2018 [6], designed a discriminative manifold regularized local descriptor algorithm to obtain steel surface defect features and complete matching by the manifold distance defined in the subspace to achieve the classification of defects in images. In 2019, Liu et al. [7] proposed an improved multi-block local binary pattern algorithm to extract the defect features and generate grayscale histogram vectors for steel plate surface defect recognition, and this work was able to recognize images at 63 FPS with a high detection accuracy at the same time.

Filter-based methods are also called spectrum-based methods and can be divided into spatial domain-based methods, frequency domain methods, and space–frequency domain methods. They aim to treat the image as a two-dimensional signal, and then analyze the im- age from the point of view of signal filter design. The filter-based methods include curvelet transform, Gabor filter, wavelet transform, and so on. Xu et al. [8] achieved the multiscale feature extraction of surface defects of a hot-rolled steel strip by curvilinear wave transform and kernel locality preserving projections (KLPP), thus generating high-dimensional feature vectors before dimensionality reduction, and finally, defect classification by SVM. In 2015, Xu et al. [9] designed a scheme that introduced Shearlet transform to provide effective multi-scale directional representation, where the metal surface image is decomposed into multiple directional sub bands by Shearlet transform, thus synthesizing high-dimensional feature vectors, which were used for classification after dimensionality reduction. Doo-chul CHOI et al. [10] used a Gabor filter combination to extract the candidate defects and preprocessed them with the double threshold method to detect whether there were pinhole defects on the steel plate surface. In 2018 [11], the classification of surface defects of a hot-rolled steel strip was achieved by extracting multidirectional shear wave features from the images and performing gray-level co-occurrence matrix (GLCM) calculations on the obtained features to obtain a high-dimensional feature set, before finally using principal component analysis (PCA) for dimensionality reduction followed by SVM for defect classification. Liu et al. [12] improved the contour wave transform based on the contour wave transform and the non-downsampled contour wave transform, and combined the multi-scale subspace of kernel spectral regression for feature extraction to achieve a relatively good recognition speed and the algorithm is applicable to a wide range of metallic materials.

The core goal of structure-based methods is to extract texture primitives, followed by the generalization of spatial placement rules or modeling, which is based on texture primitive theory. Texture primitive theory indicates that texture is composed of some minimal patterns (texture primitive) that appear repeatedly in space according to a certain rule. This method is applicable to textures with obvious structural properties such as texture primitives such as density, directionality, and scale size. In 2014, Song et al. [13] used saliency linear scanning to obtain oiled regions and then used morphological edge processing to remove oil interference edges as well as reflective pseudo-defect edges to enable the recognition of various defects in silicon steel. In 2016, Shi et al. [14] reduced the effect of interference noise on defect edge detection by improving the edge detection Sobel algorithm, thus achieving accurate and efficient localization of rail surface defects. Liu et al. [15] proposed an enhancement operator based on mathematical morphology (EOBMM), which effectively alleviated the influence of uneven illumination and enhanced the details of strip defect images. In 2016, [16] applied morphological operations to extract features of railway images and used Hough transform and image processing techniques to detect the track images obtained from the real-time camera to accurately recognize defect areas and achieve real-time recognition.

Model-based methods construct a representation of an image by modeling multiple attributes of a defect [17]. Some of the more common model-based approaches in the field of industrial product surface defect recognition are Markov models, fractal models, Gaussian mixture models, and low-rank matrix models, etc. In 2013, Xv et al. [18] introduced an environment-based multi-scale fusion method CAHMT based on the hidden Markov tree model HMT to achieve multi-scale segmentation of strip surface defects, which greatly reduced the error rate of fine-scale segmentation and the complexity of the algorithm. In 2018 [19], a saliency detection model of double low-rank sparse decomposition (DLRSD) was proposed to obtain the defect foreground image. Finally, the Otsu method was used to segment the steel plate surface defects, which improved the robustness to noise and uneven illumination. In 2019, [20] detected strip surface defects based on a simple guidance template. By sorting the gray level of the image, the sorted test image was subtracted from the guidance template to realize the segmentation of the strip surface defects. In the same year, Wang et al. [21] constructed a compact model by mining the inherent prior of the image, which provided good generalization for different inspection tasks (e.g., hot-rolled strip, rails) and had good robustness.

E. Shape Feature-Based Methods

Shape feature-based methods are also very effective defect detection methods. These methods obtain image features through shape descriptors, so the accuracy of the shape description becomes the key to the merit of the image defect recognition algorithm. A good shape descriptor should have the characteristics of geometric invariance, flexibility, abstraction, uniqueness, and completeness. The commonly used shape descriptors can be divided into two categories: one is the contour shape descriptor, which is used to describe the outer edge of the object area, and the other is the area shape descriptor, which is used to describe the whole object area. The common methods based on contour shape descriptors are Fourier transform and Hough transform, etc. For the method using Fourier transform, it mainly uses the closure and periodicity of the region boundary to convert the two- dimensional problem into a one-dimensional problem. For example, Yong-hao et al [22] enables the detection of longitudinal cracks on the surface of the continuous casting plate in a complex background by calculating the Fourier magnitude spectrum of each sub-band to obtain features with translational invariance. In addition, Hwang et al. [23] used linear discriminant analysis using short-time Fourier transform pixel information generated from ultrasound guided wave data to achieve defect detection on 304SS steel plates. The Hough transform methods use the global features of the image to connect the edge pixels to form a regionally closed boundary. For example, Wang et al. in 2019 [24] achieved the detection of product surface defects by using the fast Hough transform in the region of interest (ROI) extraction stage to detect the boundary line of the light source. Regional shape features include the length and width, elongation, area ratio, and other aggregate shape parameter methods, which is a simple shape expression method. In addition, moments are a more reliable and complex region shape feature including geometric moments, central moments, etc. As Hu invariant moments [25], moment expressions are commonly used to describe the shape of steel surface defect regions. As Hu et al. [26] used both Fourier descriptors and moment descriptors to extract the shape features of steel strip surface defect images, in addition to the grayscale features and geometric features of the images, and finally support vector machine (SVM) was used to classify the defects in the steel strip surface images. For shape feature extraction, it must be built on image segmentation and is extremely dependent on the accuracy of image segmentation. For both methods, based on texture features and shape features, they can also be used in combination. For example, Hu et al. [27] proposed a classification model based on the hybrid chromosome genetic algorithm (HCGA) and combined geometric, shape, texture and grayscale features to identify and classify steel strip surface defects.

Table 1: Comparative Analysis of Different Approaches

Author	Classifiers/Models	Accuracy
Ahmet Feyzioglu et al. (2023)	RF, LR, DT, SVM	74–79%
Yu Cheng et al. (2021)	CNN	80.25%
Renjie Tang et al. (2020)	YOLOv3, Faster R-CNN	71–72%
Weidong Zhao et al. (2021)	Faster R-CNN, RetinaNet	60–75%

CHALLENGES IN STEEL PLATE DEFECT DETECTION

Despite advancements, several challenges remain:

- Data Scarcity: High-quality, labeled defect datasets are limited.
- Class Imbalance: Some defects occur rarely, making classification difficult.
- Small Defect Size: Fine cracks and small surface defects are harder to detect.
- Real-Time Requirements: Many industrial systems require fast and efficient models for real-time inspection.
- Generalization: Models trained on specific datasets may not generalize well across different steel types or production conditions.

CONCLUSION

Steel plate defect detection using machine learning and deep learning has made remarkable progress, moving from traditional feature-based models to automated feature extraction using CNNs. MobileNetV2, with its lightweight yet powerful architecture, has demonstrated superior performance, making it highly suitable for industrial applications. However, challenges like data scarcity, real-time requirements, and generalization issues still need to be addressed. Continued research in transfer learning, data augmentation, and real-time deployment will further advance the field, making automated defect detection an indispensable part of modern manufacturing processes.

FUTURE SCOPE

Future research should focus on:

Transfer Learning: Fine-tuning pre-trained models on steel defect datasets can improve performance with limited data.
Data Augmentation: Techniques such as rotation, flipping, and noise addition can help address data scarcity.
Anomaly Detection: Implementing unsupervised or semi-supervised learning to detect unknown defects.

Explainable AI (XAI): Developing models that provide interpretable predictions to build trust in industrial applications.
Edge Deployment: Optimizing models like MobileNetV2 for deployment on edge devices for real-time factory automation.

CONFLICT OF INTEREST

The authors declare that they have no conflict of interest.

FUNDING SUPPORT

The author declare that they have no funding support for this study.

REFERENCES

[1] Congzhe You And Haozheng Kong, "Improved Steel Surface Defect Detection Algorithm Based on YOLOv8, IEEE Access, VOLUME 12, 2024.DOI: 10.1109/ACCESS.2024.3429555

[2] Lu, J., Yu, M. & Liu, J. Lightweight strip steel defect detection algorithm based on improved YOLOv7. Sci Rep 14, 13267 (2024). https://doi.org/10.1038/s41598-024-64080-x

[3] Wang, S.; Xia, J.; Ye, L.; Yang, B. Automatic Detection and Classification of Steel Surface Defect Using Deep Convolutional Neural Networks. Metals 2021, 11, 388. https://doi.org/10.3390/met11030388

[4] Liu, Yang, Ke Xu, and Jinwu Xu. 2019. "Periodic Surface Defect Detection in Steel Plates Based on Deep Learning" Applied Sciences 9, no. 15: 3127. https://doi.org/10.3390/app9153127

[5] Li, Hanlin & Liu, Ming & Yin, Yanfang & Sun, Weiliang. (2023). Steel Surface Defect Detection Based on Multi-Layer Fusion Networks. 10.21203/rs.3.rs-3802774/v1.

[6] Vira Fitriza Fadli, Iwa Ovyawan Herlistiono (2020) Steel Surface Defect Detection using Deep Learning, International Journal of Innovative Science and Research Technology, Volume 5, Issue 7, July – 2020

[7] Weidong Zhao, Feng Chen, Hancheng Huang, Dan Li, Wei Cheng, "A New Steel Defect Detection Algorithm Based on Deep Learning", Computational Intelligence and Neuroscience, vol. 2021, Article ID 5592878, 13 pages, 2021. https://doi.org/10.1155/2021/5592878

[8] Yang, L., Huang, X., Ren, Y. et al. Study on steel plate scratch detection based on improved MSR and phase consistency. SIViP 17, 119–127 (2023). https://doi.org/10.1007/s11760-022-02211-5

[9] Wang, D.; Xu, Y.; Duan, B.; ang, Y.; Song, M.; Yu, H.; Liu, H. Intelligent Recognition Model of Hot Rolling Strip Edge Defects Based on Deep Learning. Metals 2021, 11, 223. https://doi.org/10.3390/met11020223

[10] Zhang,M., Sun, W., Wang, J.: Steel surface defect detection based on adaptive threshold gray transformation. In: 2021 IEEE International Conference on Power, Intelligent Computing and Systems (ICPICS), Shenyang, Liaoning, China, pp. 280–283. (2021)

[11] Liu, K., Luo, N., Li, A., Tian, Y., Sajid, H., Chen, H.: A new self-reference image decomposition algorithm for strip steel surface defect detection. IEEE Trans. Instrum. Meas. 69(7), 4732–4741 (2020)

[12] Wu,Y.: Research on detection method of steel plate surface defects based on deep learning," Master. dissertation, Dept. Mechatronic Engineering, Donghua University, Shanghai, China, (2021)

[13] Basson, E. World Steel in Figures 2022. Available online: https://worldsteel.org/steel-topics/statistics/world-steel-in-figures-20 22/ (accessed on 30 April 2022).

[14] Jain, S.; Seth, G.; Paruthi, A.; Soni, U.; Kumar, G. Synthetic data augmentation for surface defect detection and classification using deep learning. J. Intell. Manuf. 2020, 33, 1007–1020.

[15] He, Y.; Song, K.; Meng, Q.; Yan, Y. An end-to-end steel surface defect detection approach via fusing multiple hierarchical features. IEEE Trans. Instrum. Meas. 2019, 69, 1493–1504.

[16] Luo, Q.; Fang, X.; Su, J.; Zhou, J.; Zhou, B.; Yang, C.; Liu, L.; Gui, W.; Lu, T. Automated Visual Defect Classification for Flat Steel Surface: A Survey. IEEE Trans. Instrum. Meas. 2020, 69, 9329–9349.

[17] He, Y.; Wen, X.; Xu, J. A Semi-Supervised Inspection Approach of Textured Surface Defects under Limited Labeled Samples. Coatings 2022, 12, 1707.

[18] Ma, S.; Song, K.; Niu, M.; Tian, H.; Yan, Y. Cross-scale Fusion and Domain Adversarial Network for Generalizable Rail Surface Defect Segmentation on Unseen Datasets. J. Intell. Manuf. 2022, 1–20.

[19] Wan, C.; Ma, S.; Song, K. TSSTNet: A Two-Stream Swin Transformer Network for Salient Object Detection of No-Service Rail Surface Defects. Coatings 2022, 12, 1730.

[20] Li, Z.; Tian, X.; Liu, X.; Liu, Y.; Shi, X. A Two-Stage Industrial Defect Detection Framework Based on Improved-YOLOv5 and Optimized-Inception-ResnetV2 Models. Appl. Sci. 2022, 12, 834.

[21] Song, K.; Wang, J.; Bao, Y.; Huang, L.; Yan, Y. A Novel Visible-Depth-Thermal Image Dataset of Salient Object Detection for Robotic Visual Perception. IEEE/ASME Trans. Mechatron. 2022.

[22] Sun, G.; Huang, D.; Cheng, L.; Jia, J.; Xiong, C.; Zhang, Y. Efficient and Lightweight Framework for Real-Time Ore Image Segmentation Based on Deep Learning. Minerals 2022, 12, 526.

[23] Neogi, N.; Mohanta, D.K.; Dutta, P.K. Review of vision-based steel surface inspection systems. EURASIP J. Image Video Process. 2014, 2014, 50.

[24] Verlag Stahleisen GmbH, Germany. Available online:www.stahleisen.de(accessed on 13 June 2022.

[25] Viola, P.; Jones, M.J. Robust real-time face detection. Int. J. Comput. Vis. 2004, 57, 137–154.

[26] Dalal, N.; Triggs, B. Histograms of oriented gradients for human detection. In Proceedings of the 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05), San Diego, CA, USA, 20–25 June 2005; IEEE: Piscatevi, NJ, USA, 2005; 1, pp. 886–893.

[27] Felzenszwalb, P.F.; Girshick, R.B.; McAllester, D.; Ramanan, D. Object detection with discriminatively trained part-based models. IEEE Trans. Pattern Anal. Mach. Intell. 2010, 32, 1627–1645.

[28] Krizhevsky, A.; Sutskever, I.; Hinton, G.E. Imagenet classification with deep convolutional neural networks. Commun. ACM 2017. 60, 84–90.

[29] Goodfellow, I.; Pouget-Abadie, J.; Mirza, M.; Xu, B.; Warde-Farley, D.; Ozair, S.; Courville, A.; Bengio, Y. Generative adversarial networks. Commun. ACM 2020, 63, 139–144.

[30] Simonyan, K.; Zisserman, A. Very deep convolutional networks for large-scale image recognition. arXiv 2014, arXiv: preprint,1409,1556.

[31] Szegedy, C.; Liu, W.; Jia, Y.; Sermanet, P.; Reed, S.; Anguelov, D.; Rabinovich, A. Going deeper with convolutions. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Boston, MA, USA, 7–12 June 2015; pp. 1–9.

[32] He, K.; Zhang, X.; Ren, S.; Sun, J. Spatial pyramid pooling in deep convolutional networks for visual recognition. IEEE Trans. Pattern Anal. Mach. Intell. 2015, 37, 1904–1916.

[33] Girshick, R.; Donahue, J.; Darrell, T.; Malik, J. Rich feature hierarchies for accurate object detection and semantic segmentation. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Columbus, OH, USA, 23–28 June 2014; pp. 580–587.

[34] He, K.; Zhang, X.; Ren, S.; Sun, J. Deep residual learning for image recognition. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Las Vegas, NV, USA, 27–30 June 2016; pp. 770–778.

[35] Girshick, R. Fast r-cnn. In Proceedings of the IEEE International Conference on Computer Vision, Santiago, Chile, 7–13 December 2015; pp. 1440–1448.

[36] [33] Liu, W.; Anguelov, D.; Erhan, D.; Szegedy, C.; Reed, S.; Fu, C.Y.; Berg, A.C. Ssd: Single shot multibox detector. In European Conference on Computer Vision; Springer: Cham, Switzerland, 2016; pp. 21–37.

[37] Iandola, F.N.; Han, S.; Moskewicz, M.W.; Ashraf, K.; Dally, W.J.; Keutzer, K. SqueezeNet: AlexNet-level accuracy with 50x fewer parameters and<0.5 MB model size. arXiv 2016, arXiv:preprint. 1602.07360.

[38] Ren, S.; He, K.; Girshick, R.; Sun, J. Faster r-cnn: Towards real-time object detection with region proposal networks. Adv. Neural Inf. Process. Syst. 2015, 28.

[39] Redmon, J.; Divvala, S.; Girshick, R.; Farhadi, A. You only look once: Unified, real-time object detection. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Las Vegas, NV, USA, 27–30 June 2016; pp. 779–788.

[40] Lin, T.Y.; Goyal, P.; Girshick, R.; He, K.; Dollár, P. Focal loss for dense object detection. In Proceedings of the IEEE International Conference on Computer Vision, Venice, Italy, 22–29 October 2017; pp. 2980–2988.

[41] [38] Cai, Z.; Vasconcelos, N. Cascade r-cnn: Delving into high quality object detection. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Los Alamitos, CA, USA, 18–22 June 2018; pp. 6154–6162.

[42] Huang, G.; Liu, Z.; Van Der Maaten, L.; Weinberger, K.Q. Densely connected convolutional networks. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Honolulu, HI, USA, 21–26 July 2017; pp. 4700–4708.

[43] Howard, A.G.; Zhu, M.; Chen, B.; Kalenichenko, D.; Wang, W.; Weyand, T.; Adam, H. Mobilenets: Efficient convolutional neural networks for mobile vision applications. arXiv 2017, arXiv:preprint 1704.04861.

[44] Zhang, X.; Zhou, X.; Lin, M.; Sun, J. Shufflenet: An extremely efficient convolutional neural network for mobile devices. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Los Alamitos, CA, USA, 18–22 June 2018; pp. 6848–6856