Journal of Computers Vol. 35 No. 5, October 2024, pp. 187-201
doi: 10.53106/199115992024103505013

Design of a Machine Vision Based Image Acquisition and Detection System for Surface Defects of Metal Bars

Xiang-Yun Yi[1,2*], Rui-Da He[1], Juan-Hong Jing[1],
Wei-Tao Jia[1], Wen-Ya Shi[1], and Tao Zhang[1]

[1] Department of mechanical engineering, Hebei Institute of Mechanical and Electrical Technology,
Xingtai City 054000, Hebei Province, China

{xiangyun7987, ruida8908, hong798798, taotao987,
wenya9789, xiaotao8978}@163.com

[2] Xingtai City Verification Equipment Intelligent Technology Innovation Center,
Xingtai City 054000, Hebei Province, China

Received 26 September 2024; Revised 17 October 2024; Accepted 19 October 2024

Abstract. This article constructs an automated recognition system for surface defect recognition of metal bars, which includes hardware components, recognition algorithms, and recognition systems. Firstly, a visual recognition hardware system was established, which includes the selection process of cameras and lenses, determining the hardware model and brand. At the same time, considering the reflective characteristics of the rod surface, the light source scheme of this system was explained. In terms of recognition algorithms, Mask R-CNN is used as the metal rod feature network. In response to the problems of information loss and aliasing caused by the reduction of pyramid feature channels and multi-level feature fusion in the feature extraction process of Mask R-CNN, a bottom-up reverse fusion path is proposed to be added to the original feature extraction network, fully utilizing shallow feature information. At the same time, a hierarchical attention mechanism and a difference region attention module are added to the path of obtaining the feature map; Finally, in the experimental stage, the construction of the visual detection system framework was completed, which proved the effectiveness and scientificity of the proposed method in terms of recognition accuracy and speed.

Keywords: defect identification, machine vision, Mask R-CNN

1 Introduction

Metal bars made by casting, forging, rolling, and powder metallurgy processes are important industrial structural materials widely used in various fields such as aerospace, military production, energy manufacturing, construction, chemical engineering, and metallurgy. These metal bars, as structural materials, are often subjected to harsh working conditions such as high temperature, friction, corrosion, and fatigue during use. Therefore, if there are defects in metal bars, it is highly likely to cause production accidents, especially for metal bars produced using casting processes. Due to casting defects such as dense looseness, central shrinkage, and surface cracks generated during the casting process, if the defective bars cannot be detected and removed in a timely manner, it is extremely easy to reduce the reliability of key equipment operation in the aerospace and military fields during use, and even cause component fracture accidents, resulting in casualties and significant economic losses. It can be seen that the material condition of metal bars and their components during the manufacturing process directly determines the performance and safety of the final product, which has a huge impact on the development of modern society and economy. Therefore, it is of great significance to carry out research on improving the level of defect detection technology and quality control capabilities of metal bars and their components.

Due to factors such as manufacturing process control and processing industry, various types of small defects are inevitably introduced into the metal bars obtained during processing. Generally, these defects exist inside or on the surface of the metal bars, and under the combined effect of internal stress and external environment, they are prone to further expansion, resulting in safety production accidents. At the same time, these small defects are generally not easily detectable by the naked eye, and even if they can be detected by the naked eye, manual defect detection cannot achieve continuous and error free defect detection. Therefore, in order to ensure

* Corresponding Author

the safety of the use of metal bars, it is necessary to use automated detection technology to conduct quality inspection during the production process and defect detection before leaving the factory for metal bars and their components. Automated testing technology, as an important branch of material processing engineering, not only plays an important role in controlling product material quality and production process quality, coordinating performance and economic benefits, but also is one of the important technical means to ensure the safety of material and product use. It can provide quality information parameters for manufacturing process informatization, in order to improve production processes and increase efficiency [1].

Manual inspection is the initial method for detecting surface defects in products, which uses sampling inspection. However, this method has low sampling rate, low accuracy, poor real-time performance, low detection efficiency, high labor intensity, and is greatly influenced by human experience and subjective factors, so it has gradually been eliminated. With the development of industrial technology, non-destructive testing has become the main means of surface defect detection, including magnetic particle testing, magnetic flux leakage testing, penetrant testing, eddy current testing, ultrasonic testing and other testing methods. Although these methods have high flexibility, are easy to implement automation, have high detection sensitivity, and ultrasonic testing also has strong penetration ability, the above methods also have unavoidable disadvantages: for example, magnetic particle testing is limited to ferromagnetic materials and has a small detection range; Magnetic flux leakage testing also requires ferromagnetic materials that are insensitive to shallow pits on the surface of the workpiece; Penetration testing has long inspection time, high cost, and serious environmental pollution; Eddy current testing is limited to conductive materials and has a slow detection speed; Ultrasonic waves are prone to errors when detecting parts with rough surfaces. In addition, both manual testing and non-destructive testing have certain labor costs. With the development of machine vision, vision based defect detection technology has gradually been applied to surface defect detection of various products. The surface defect detection methods of machine vision are divided into defect recognition based on traditional algorithms and defect recognition based on deep learning. Traditional algorithm based defect recognition has poor adaptability and requires different algorithms for detection based on different samples. Compared with deep learning detection methods, the level of intelligence is lower [2]. Therefore, this article uses a deep learning based defect detection method for surface defect detection of metal bars.

The main research object of this article is metal bars, which have strong reflective ability under light sources, thus posing a challenge to the accuracy of image acquisition. The work done in this article is as follows:

1) A visual inspection system was built based on the usage scenario, and the selection process of cameras and lenses in the inspection system was elaborated in detail. The brand and model of cameras and lenses were determined, and the lighting scheme was determined based on the reflective characteristics of the rod material.

2) Using Mask R-CNN as the surface defect detection feature network for metal bars, a bottom-up reverse fusion path is added to the original feature extraction network. A hierarchical attention mechanism and a differential region attention module are added to the path of obtaining the feature map to collect defect position sensitive feature maps separately, thereby improving the feature extraction ability of the model.

3) We have built a visual inspection system framework and completed the detection experiment of surface defects on metal bars. The feasibility and scientificity of the method proposed in this paper have been confirmed through experiments.

2 Related Work

In 2021, Xiaohong Li proposed a CNN compressed image method to address the inefficiency of manual surface defect detection. The compressed images significantly improved the accuracy of subsequent defect classification, and this method also had strong generalization ability. The method was successfully applied to machine vision systems, providing a new approach for image processing [3]. Chao Feng developed a steel ball surface defect image acquisition and processing system based on the arc-shaped characteristics of the steel ball surface, providing a reference for detecting arc-shaped surface defects [4]. Suqin Wang proposed a machine vision based surface defect algorithm for defect detection of irregular products in industry. The algorithm combines surface scratch detection algorithm and surface glue overflow algorithm and is applied to defect detection of Bluetooth earphones, with a recognition accuracy of up to 98% [5]. Liming Gao proposes an improved Faster R-CNN algorithm for detecting surface defects in insulated bearings. Firstly, image acquisition is performed on the surface

defects of insulated bearings, and the K-means++ algorithm is used to cluster the defect dataset to obtain anchor boxes suitable for such defects; Then, replace ROI Pooling with ROI Align to avoid localization errors caused by ROI Pooling quantization; Finally, an online difficult case mining strategy is added to the algorithm to improve the detection accuracy of difficult to detect samples. The experimental results show that the improved Faster R-CNN network achieves an accuracy of 91.2% in detecting surface defects on insulated bearings, which is 4.8% higher than the Faster R-CNN network [6].

After image acquisition is completed, it is necessary to classify defects in the image and recognize target features in the image.

In 2020, Yuli Chen used EfficientNet to build a transfer learning network and trained a classification model for cracked and intact eggs. He designed a simple simulated rolling egg breaking device to achieve the goal of rolling eggs colliding with each other, providing samples and datasets for the effectiveness experiment of crack image acquisition and image classification. The final test set achieved a comprehensive accuracy of 97.2%, The detection accuracy of cracked eggs reached 95.9%, and the comprehensive detection accuracy reached 96.7% [7].

Zheng Xiong from Kunming University of Science and Technology proposed an improved surface defect classification algorithm for hot-rolled steel strips based on MobileNetV3 algorithm to address the problem of surface defects in production. By introducing an improved permutation attention mechanism SAMM to replace the SE attention mechanism in the original model, the classification accuracy is improved while reducing the number of parameters. At the same time, the Mish activation function was introduced to replace the Hard Wish activation function in the original model, further improving the classification accuracy. The experimental results showed that the improved algorithm achieved a classification accuracy of 99.44%, an increase of 5.74% compared to the original model, while the parameter size was only 10.35 MB, a decrease of 35.59% compared to the original model. This indicates that the improved classification algorithm is capable of handling the task of classifying surface defects in hot-rolled steel strips [8].

Yan Xia from Southeast University proposed a lightweight deep learning image classification model based on Swin Transformer to balance the accuracy and speed of defect classification in TFT-LCD panel defect detection. Firstly, Token fusion is applied to the feature maps input at each layer of the model to reduce its computational complexity and improve its lightweight level. Secondly, a depthwise separable convolution module is introduced to help the model increase convolution induction bias, thereby alleviating the model's dependence on massive data. Finally, the knowledge distillation method is used to overcome the problem of decreased detection accuracy caused by model lightweighting. The experiments on the self-made TFT-LCD panel defect classification dataset show that the improved model proposed in this paper reduces FLOPs computation by 2.6 G and improves speed index by 17% compared to the baseline model, while the Top-1 Acc accuracy only loses 1.3%. Compared with other mainstream image classification models, it has a more balanced accuracy and speed on both self-made and public datasets [9].

In terms of identifying surface defects on steel, Shiqiang Zhang proposed an improved YOLOv5s steel surface defect detection algorithm based on the low detection efficiency and accuracy of previous steel surface defect detection techniques. The algorithm incorporates a coordinate attention mechanism and a dilated space convolution pooling pyramid, as well as an improved selective kernel attention mechanism, allowing the model to better utilize frequency information in the feature map and enhance its expressive ability [10].

Yuxiang Wang from Nanjing University proposed a metal surface defect detection algorithm based on U^2Net called Attention-U^2Net. Firstly, a U-shaped attention encoding module is designed to increase the weight of the defect area while suppressing background noise during encoding. Then, a loss function with weights is designed and combined with a multi-level supervision method to make the detection algorithm more focused on difficult to detect samples and boundary pixels. Finally, an algorithm that automatically calculates the grayscale threshold based on the image results for result optimization is used to generate high-quality defect prediction images [11].

On the basis of the above research results, this article designs a visual based surface defect automatic detection system for metal bar materials and studies the method of feature recognition in the visual recognition process. Therefore, the chapter arrangement of this article is as follows: Chapter 2 mainly searched and collected the research results of relevant scholars, and based on this, organized the ideas of this article. Chapter 3 discusses the construction process of the detection system. Chapter 4 mainly focuses on the improvement and implementation process of recognition algorithms. Chapter 5 is the simulation experiment section, and the construction of the detection system has been completed. Chapter 6 is the conclusion section, which summarizes the research results of this article and provides prospects.

3 Construction of Detection System

The configuration steps of a complete visual inspection system should be carried out in the following order: first, determine the type and resolution of camera sensors based on the required detection accuracy of the actual scene; second, match the camera with a lens with the appropriate focal length according to the camera lens interface and detection distance; and finally, select the light source to provide illumination scheme for rod inspection [12].

3.1 Selection of Camera and Lens

In actual production practice, the surface quality of special steel bars manufactured on the production line is poor, with a large amount of dust attached, making it difficult to identify surface crack defects. Therefore, when selecting a camera, choosing a color camera can better restore the real situation of the bar surface, which is beneficial for the target detection network to extract features from the image for subsequent recognition and localization. The photosensitive chip uses CMOS sensors, which are widely used in industrial cameras and have a high cost performance ratio. The resolution and line frequency of industrial cameras are important parameters for camera selection. According to the technical requirements of the detection system, the surface defect detection accuracy of the rod is 0.5 mm. In order to ensure the detection accuracy of the image, this article sets the detection accuracy at 0.15 mm. The rod used for the experiment is about 400 mm long, and the camera inevitably captures images of the surrounding environment during shooting. Therefore, the camera's field of view should be appropriately expanded to 500 mm. Therefore, when the camera reaches the lowest horizontal resolution, it requires 2400 pixels, with each pixel corresponding to a detection accuracy of 0.15 mm. Considering the stability of detection, usually 3 to 4 pixels are used to represent a detection accuracy instead of one pixel. In this article, 3 pixels are selected to correspond to a unit length of detection accuracy. Therefore, a minimum of 7200 pixels are required in the horizontal direction, so an industrial line array camera with a resolution of 8192 × 4 was ultimately chosen. To ensure stable acquisition of high-quality raw images, there should be a margin for resolution. Choose MindVision's MV-GEL83BM series camera as the camera, the detailed parameters of the camera are shown in Table 1.

Table 1. Detailed parameter list of the camera

Model	MV-XGLC83C-T4-90
Colour	Color
Sensor type	CMOS
Effective pixels	8192×3
Video output format&maximum line frequency	Bayer 8bit:48K
Pixel size	7*um*×7*um*
Weight	<550*g*

The resolution of the linear array camera selected in this article is 8192 × 3, with a single pixel size of 7.04 μ m and a selected field of view of 500 mm. The optical magnification and lens focal length of the camera can be calculated using the following formula [13]. The calculation formula for target surface size is:

$$\text{Target size} = \frac{7.04 \times 8192}{1000} = 57.67mm \tag{1}$$

The optical magnification is expressed as:

$$\text{Magnification factor} = \frac{\text{Target size}}{\text{field of view}} = \frac{57.67}{500} = 0.115 \tag{2}$$

According to the actual working situation, if the vertical distance of the rod detected by the lens distance is 800mm, the focal length of the lens is expressed as:

$$\text{focal length} = \frac{\text{detection distance} \times \text{Target size}}{\text{field of view}} = 92.272 \quad (3)$$

Therefore, the lens is selected as Schneider lens, and the lens parameters are as Table 2:

Table 2. Lens parameter list

Parameter	Parameter size
Model	Componon-S 5.6/100
Image field	108mm
Interface	V-mount
Focal length	102.3mm
F value	5.6-45
Maximum diameter	47mm
Filter size	*M*35.5*0.5

3.2 Selection of Light Sources and Lighting Schemes

The light source is an important factor affecting the imaging quality, and the image quality determines the complexity of subsequent image processing, which requires the collected surface images of the rod to have high image quality. For the collected surface images of the rod, three requirements must be met. Firstly, the overall brightness of the rod image. If the surface image of the rod is too bright or too dark, it will reduce the key information of cracks on the rod. Secondly, the image of the rod cannot be blurred, and the design principles of the detection system also include the detection accuracy of the system. Image blurring and defocusing will cause the system to lose accuracy. Finally, there needs to be a certain degree of discrimination between cracks and background areas in the bar image.

In the process of image acquisition, based on the illumination method of coaxial light source, different colors of light were applied to collect defects, and more suitable raw data was selected to prepare for image processing. During the experiment, three color light sources were used to collect defects, namely white light, red light, and blue light. In order to objectively demonstrate the impact of each light source on defect images, experiments were conducted on images from three different light sources, and grayscale processing was performed. By observing the distribution of grayscale values in images under different light sources and comparing the characteristics of the original data under different light sources, the highest and lowest grayscale values were removed, and the standard deviation of the remaining data was calculated. According to Table 3, under the red light source, the standard deviation of the image grayscale value is the lowest, which further indicates that the grayscale value uniformity of the white light collected image is the best, which is more conducive to the next batch preprocessing of the image. Due to the uniform distribution of partial image data from red and blue light sources, white light source images will be the main focus in dataset production, with red and blue light source image data as supplementary.

Table 3. Comparison of data under different light sources

Light source	Average gray value	Variance	Standard deviation
Red light source	119	93.6	8.74
White light source	126	12.8	3.17
Blue light source	147	78.3	6.29

Due to the advantages of long service life, small size, low power consumption, fast response speed, high reliability, uniform and stable light emission, and easy integration of LEDs, and considering the complex background of the rod surface and the small area occupied by cracks in the image, in order to better preserve the original surface information, a widely applicable white LED light source was ultimately chosen [14].

In forward illumination, the light source and camera are located on the same side of the object being measured, mainly used to detect important details, defects, and scratches on the surface of the object. Compared with flat plates, rods have a cylindrical structure and are affected by their surface curvature, resulting in uneven

brightness between the center and edges when the light source shines on the surface of the rod. This leads to the highest brightness in the middle area of the image and darker edges. At the same time, the large amount of flake like oxide on the surface also changes the reflection angle of light, resulting in uneven brightness of the image. Overall, the light source layout scheme is shown in Fig. 1.

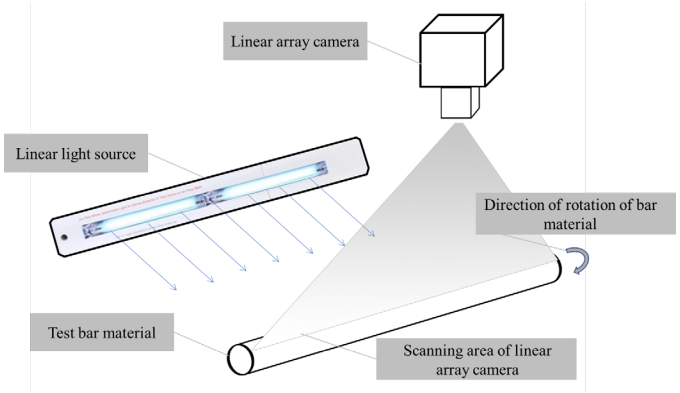

Fig. 1. Overall scheme of detection system

The above process completed the construction of the visual system for bar inspection, taking into account the structural characteristics of metal bars, and designed linear light sources and supplementary lighting schemes.

4 Construction of Defect Detection Model

This article uses Mask R-CNN as the surface defect detection feature network for metal bars. However, Mask R-CNN [15] cannot perfectly fit the usage scenario in this article. The object detection part in the Mask R-CNN network adopts the Faster R-CNN framework, which has two drawbacks: first, Faster R-CNN only predicts on the shared feature layer, but at this time, the shared feature layer only has high abstract semantic features. The higher the level of abstraction, the less instance detail information is retained. In the continuous pooling and convolution process, the detail information of small targets may be lost, resulting in poor small target detection performance, or even missed detection, ultimately affecting the accuracy of instance segmentation in the second stage. Secondly, predictions were made in both the RPN stage and the Fast R-CNN stage, resulting in a larger model structure and slower training and prediction speeds.

In the feature extraction process, Mask R-CNN has the characteristics of reducing pyramid feature channels and multi-level feature fusion. When detecting surface defects on metal bars, it can cause information loss and aliasing effects. Therefore, this article adds a bottom-up reverse fusion path [16] to the original feature extraction network of Mask R-CNN, fully utilizing shallow feature information and reducing information loss. Secondly, a hierarchical attention mechanism and a differential region attention module are added to the path of obtaining feature maps. The roles of the two modules are: the hierarchical attention mechanism is responsible for collecting defect location sensitive feature maps [17], and the differential region attention module [18] is responsible for obtaining the differences between the high-level feature map and the current level feature map, which are used to fuse and update distinguishable features, thereby improving the feature extraction ability of the model. The overall structure of the model is shown in Fig. 2.

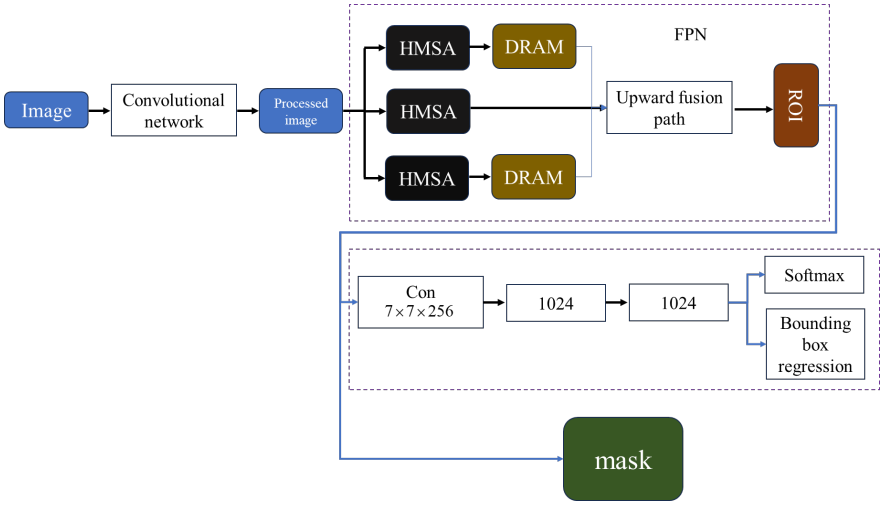

Fig. 2. Overall network architecture diagram

4.1 Improvement of Feature Extraction Network

Based on the original FPN structure, this paper introduces a bottom-up reverse fusion path L_1, L_2, L_3. Feature map B undergoes 128 1 × 1 convolutions to produce L_1, while L_1 undergoes 128 3 × 3 convolutions to yield P_1. Additionally, L_1 undergoes a double downsampling operation, and the resulting output is then combined with the adjacent high-level feature map F_2, which has undergone 256 1 × 1 convolutions, to produce L_2. L_2 undergoes 256 3 × 3 convolutions to form feature map P_2. Meanwhile, L_2 continues to undergo a double downsampling operation, and the resulting output is then combined with the adjacent high-level feature map F_3, which has undergone 256 1 × 1 convolutions, to yield L_3. L_3 undergoes 256 3 × 3 convolutions to produce feature map P_3 [19]. The structure is illustrated in Fig. 3.

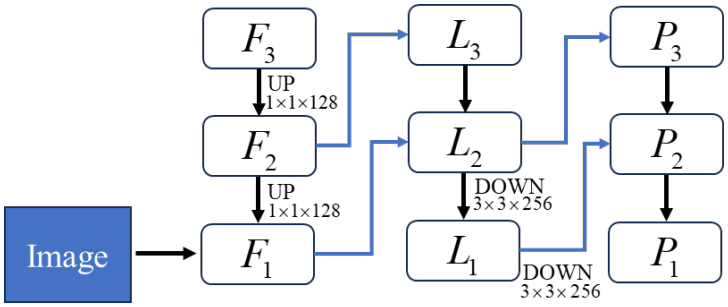

Fig. 3. Schematic diagram of improved fusion path structure

4.2 Integration of Attention Mechanism

Add Hierarchical Multi Scale Attention (HMSA) and Difference Region Attention Module (DRAM) to the upper horizontal path and on this path respectively, fully utilizing shallow feature information and adjacent semantic information.

The main idea of the HMSA feature extraction structure is to scale the original image to 0.5 times and 2 times its original size, and input 3 images as a set of training data to the network. For each scale, a feature matrix with preference properties is learned, and then the extracted features at different scales are fused together. It is worth noting that in the hierarchical method proposed in this article, all relative attention feature information between

adjacent scale sets is learned, rather than learning attention feature information of a fixed scale. That is, this structure does not only extract feature information of 0.5 times, 2 times, and the original image, but also extracts feature information of all scales between 0.5 times and 2 times. The network structure is shown in Fig. 4.

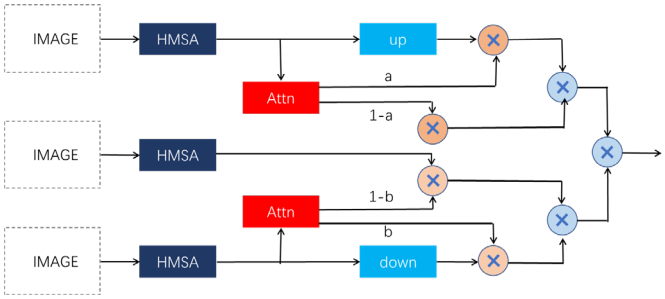

Fig. 4. Schematic diagram of HMSA network structure

In the process of model training, the richer the image data, the better the model training effect. Therefore, data augmentation methods are used to scale the original image to obtain three sizes of training data, including the original image, which are used as inputs for the three branches of the HMSA feature extraction network model. However, in the process of predicting results, in order to ensure that the test images are not distorted, only 0.5 times and itself are used for prediction, because the network has already extracted feature information at all scales between 0.5 times and 2 times the images during training, so there is no need to upsample the original images again during inference prediction. The prediction phase has been reduced from 3 branches to 2 branches. On the one hand, only downsampling is performed to prevent image distortion and avoid unnecessary noise during the upsampling process. On the other hand, reducing the number of branches from 3 to 2 can reduce the computational burden caused by network forward propagation, improve the inference speed of the model, and ensure both accuracy and speed.

DRAM aims to extract features ignored by the current level by calculating the difference between the current level and the previous level. Given the i-th feature F_i and the i+1-th feature F_{i+1}, re scale the high-level features using the upsampling method in FPN, and then use a compatibility function based on cosine similarity to calculate the difference between adjacent layers of features. Meanwhile, a new 1 × 1-convolutional layer α is adopted to embed into F_i to ensure that these two modules are independent of each other and do not interfere with each other. Another convolutional layer β is used to encode differences and obtain difference driven attention maps. Express this process as follows:

$$F_i = w_\beta U\left(F_{i+1}\right)\left(\varphi\left(1-\mathrm{sim}\left(w_\alpha F_i, U\left(F_{I+1}\right)\right)\right)\right) \tag{4}$$

$$w_\beta \in r^{c\times c} \tag{5}$$

$$w_\alpha \in r^{c\times 1} \tag{6}$$

Here, cosine similarity is used to describe the consistency between two feature levels, and the Sigmoid activation function is used to unify the output between (0,1). This module can clearly explain the difference between two adjacent feature levels to discard or retain high-level features. It provides a gate mechanism to discard or retain high-level features, which reduces the impact of interpolation and aliasing to a certain extent. Finally, use a 3 × 3 convolutional layer to recombine the input features and expand their receptive field [20].

$$\left(w_\alpha, F_i, F_{i+1}\right) = \varphi \cdot \mathrm{sim}\left(w_\alpha F_i, U\left(F_{i+1}\right)\right) \tag{7}$$

The final result of integrating the differential features is:

$$Fl_i = w_\beta F_i \oplus L_i \tag{8}$$

Fl_i is the refined feature level obtained through DRAM, and $\oplus$ represents element wise addition. Finally, the refined feature $\{L_1, L_2, L_3\}$ will be obtained, and the network structure is shown in Fig. 5.

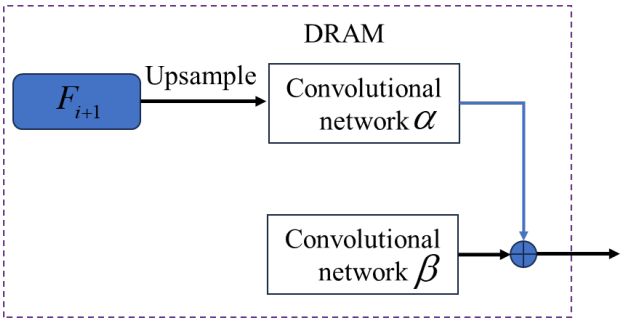

Fig. 5. DRAM network architecture

4.3 Data Enhancement

The sample sources of the metal bar dataset in this article include on-site collected data, network images, and borrowed datasets from others. In the existing data, there are relatively more surface defect data for pipes. However, considering the similarity in surface defects between metal rods and metal pipes of similar materials, there is no difference in visual recognition between the two. Therefore, a small part of the dataset in this article comes from the pipe surface defect dataset of Tianjin University of Commerce. There are various types of surface defect detection for metal bars, including scratches (abrasions), cracks, rust, and other types. These defects have different shapes and sizes. Among them, the crack characteristics are described as straight or curved cracks along the rolling direction of the metal bar, which may appear as linear or crack shapes on the surface of the bar. Cracks may lead to a decrease in the surface strength of the bar, which has a negative impact on the mechanical properties and durability of subsequent use. Scratch (abrasion) phenomenon usually refers to linear or curved marks formed on the surface of the rod, which are usually caused by the relative motion between external objects and the surface of the rod, resulting in scratches, abrasion or wear. Scratches usually present irregular and uneven shapes, which may affect the appearance, quality, and performance of materials. Corrosion phenomena include rust spots, oxide layers, etc., presenting different colors and textures of coverings [21]. The data image is shown in Fig. 6.

(a) Rust (b) Flaw (c) Scratch

Fig. 6. Typical surface defects of pipes

After the production of the dataset, a total of 67 flaw images, 138 scratch images, and 62 rust images were collected. The image resolution was set to 1024 × 1024, and each image contained one or more surface defects. Overall, the number of images in the dataset is relatively small, so data augmentation is used to increase the number of defective images. Based on conventional methods of image augmentation, commonly used methods

include image mirroring, image rotation, changes in the number of visual system channels, visual system color jitter, and adding noise to recognized images. The image mirroring method includes three mirroring methods: horizontal, vertical, and diagonal. Horizontal mirroring swaps the left and right parts of the image, vertical mirroring swaps the upper and lower parts, while diagonal mirroring is equivalent to sequentially performing horizontal and vertical mirroring. Image rotation involves rotating around the center point at a certain angle. Based on this, this chapter first expands the dataset through mirroring and rotation methods, and the enhancement effect is shown in Fig. 7.

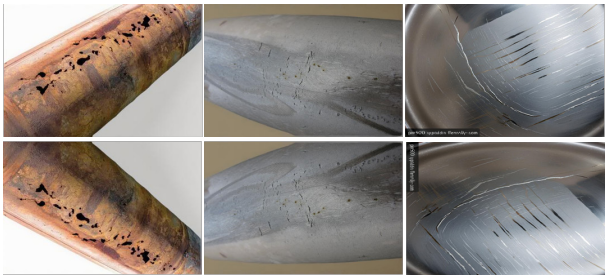

Fig. 7. Example of rotation operation for partial defects

This chapter mainly completed the major improvement of the recognition model, adding attention mechanism to the model to enhance the recognition effect of defects. Then, based on the existing defect images, the number of images in the dataset was increased through data processing methods.

5 Construction of Defect Recognition System and Analysis of Experimental Results

For the recognition of surface defects on metal bars, a complete recognition system needs to be built, which mainly completes the display function. This chapter will study the software framework of the visual inspection system for surface defects on metal bars, and use PyQt software to develop and design the system interface. The system software designed in this article mainly includes four parts, namely parameter setting module, software input module, bar surface defect detection module, and output module [22]. The framework diagram of the visual inspection system software for surface defects of metal bars is shown in Fig. 8.

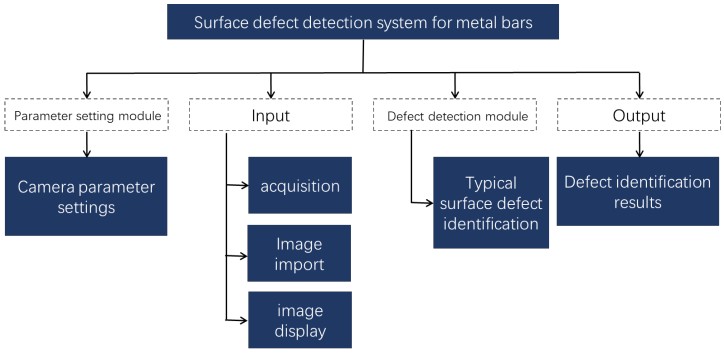

Fig. 8. Framework diagram of metal bar surface defect recognition system

The system includes the frame rate, resolution, exposure time, and gain settings of the camera; The input module mainly selects the method of image acquisition, which can achieve camera capture of images, local import of images, and display of input images; The bar surface defect detection module can meet the functions of bar sur-

face defect detection and multi type defect detection on bar surfaces; The output module can display the surface defect detection results of the bar material.

The parameter setting module mainly completes the setting of camera frame rate, resolution, exposure time, and gain. Considering the performance of the camera and computer, the camera frame rate can be set within the range of 1-20 frames per second, and the resolution can be selected within the maximum resolution range according to actual shooting needs. The exposure time and gain are mainly used to adjust the exposure level of the images captured by the camera, and are adjusted according to the on-site environment.

The input module includes functions such as capturing images with a camera, importing images locally, and displaying input images. By capturing images with a camera, real-time images can be captured and processed, or images that need to be processed can be selected through local import. The images captured by the camera or imported locally will be displayed in the display area of the input module.

The surface defect detection module of bar materials mainly completes the control of metal bar material defect detection and multi type defect detection process on the surface of bar materials.

The first detection mode is surface defect detection of bar materials, mainly using the detection algorithm introduced in Chapter 4 of this article to process the input image and compare it to determine whether it is the corresponding bar material surface defect. The second detection mode is multi type defect detection on the surface of bar materials, which is an application of the multi type defect detection algorithm for bar materials introduced in Chapter 4 of this article. The trained model is used to detect the input image and check whether there are scratches (abrasions), cracks, rust and other defects in the image.

The main function of the output module is to display the processing results of the input image. In rod defect detection, the defect type is displayed, and in multi type defect detection on the surface of the rod, the detection result image can be directly displayed.

The design of the detection interface is shown in Fig. 9.

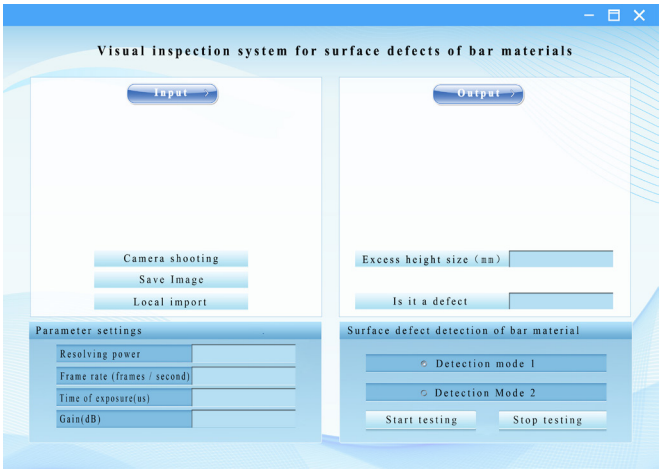

Fig. 9. Framework diagram of metal bar surface defect recognition system

5.1 Testing System Experimental Environment Configuration

The experimental environment is configured as Ubuntu system, equipped with Nvidia RTX4090 graphics card, using Python 3.6.6, and PyTorch deep learning framework. Starting from a size of 4 × 4, use the DiffGrad optimizer and set the weight parameters to 0.9 and 0.99.

5.2 Surface Defect Detection Experiment on Bar Material

Use the improved recognition model in this article to identify three typical defects of metal bar materials, namely cracks, scratches, and rust, and conduct recognition experiments to verify the defect recognition results of the bar

materials. The real experimental scene constructed is shown in Fig. 10.

Fig. 10. Construction of visual inspection system

To verify the effectiveness of the surface defect detection of the bar material proposed in this article, three types of bar material samples as shown in Fig. 6 were selected for testing experiments. Firstly, the visual inspection system developed in this article was validated, and the detection result interface is shown in Fig. 11.

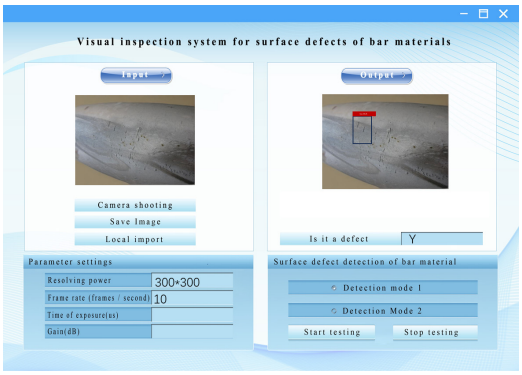

Fig. 11. Surface defect identification results

From the above figure, it can be seen that after the camera captures the image of the rod, the defect detection results can be seen in the output box on the right side of the system, and the experimental system can complete the predetermined function.

5.3 Analysis of Algorithm Recognition Results

This article elaborates theoretically on the beneficial effects of introducing a bottom-up fusion path based on the original framework structure, as well as adding Hierarchical Multi Scale Attention (HMSA) and Difference Region Attention Module (DRAM) to this path for feature extraction in network models. In this section, ablation experiments were conducted on the surface defect dataset of bar materials to demonstrate the impact of the two improved parts on the model feature extraction network through experimental data.

In order to verify the performance of the method proposed in this article, comparative ablation experiments were conducted with three models: the model with only added reverse fusion path, the model with only added hierarchical attention mechanism and differential region attention module. For ease of description, the model with added reverse fusion path was referred to as Mark R-CNN-RP, the model with only added hierarchical attention mechanism was referred to as Mark R-CNN-HMSA, and the model with only added differential region atten-

tion module was referred to as Mark R-CNN-DRAM. The improved algorithm in this article is called i- Mark R-CNN. Then, three models were used to identify the three typical features of cracks, scratches, and rust. The recognition results were compared in terms of recognition accuracy and missed detection rate. The recognition results are shown in Fig. 12.

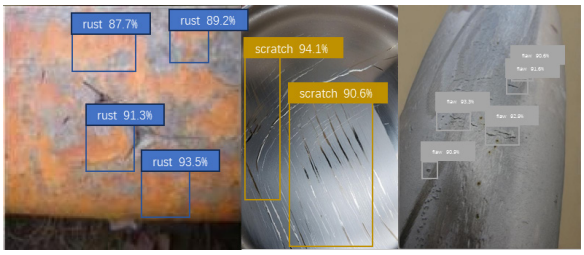

Fig. 12. Comparison of test results

The recognition results of each algorithm for the three types of defects are shown in Table 4, Table 5, Table 6.

Table 4. Rust recognition results

Defect category	Actual quantity	Inspection quantity	Number of missed detections	Loss	Noise factor	Recognition accuracy
Mark R-CNN-RP	62	57	5	8.06%	6.45%	93.55%
Mark R-CNN-HMSA	62	59	3	4.83%	5.72%	94.28%
Mark R-CNN-DRAM	62	61	1	1.61%	7.01%	92.99%
i- Mark R-CNN	62	62	0	0	1.04%	98.96%

Table 5. Flaw recognition results

Defect category	Actual quantity	Inspection quantity	Number of missed detections	Loss	Noise factor	Recognition accuracy
Mark R-CNN-RP	67	64	3	4.47%	7.17%	92.83%
Mark R-CNN-HMSA	67	60	7	10.44%	8.25%	91.75%
Mark R-CNN-DRAM	67	62	5	7.46%	4.38%	95.62%
i- Mark R-CNN	67	62	5	7.46%	2.19%	97.81%

Table 6. Scratch recognition results

Defect category	Actual quantity	Inspection quantity	Number of missed detections	Loss	Noise factor	Recognition accuracy
Mark R-CNN-RP	138	133	5	3.62%	6.27%	93.73%
Mark R-CNN-HMSA	138	129	9	6.52%	8.01%	91.99%
Mark R-CNN-DRAM	138	104	34	24.63%	17.46%	82.54%
i- Mark R-CNN	138	137	1	0.72%	1.01%	98.99%

After identifying typical defects, the following results can be obtained from the above statistical results: Firstly, for the identification of the three types of defects, the identification method proposed in this paper can ensure an overall recognition accuracy of over 97%, which basically meets the requirements of modern production technology for recognition accuracy. Analysis shows that there are many missed detections when identifying crack defects in this paper. After analysis, it is believed that due to lighting problems, some small cracks are identified as scratches. Therefore, in the process of identifying cracks, the recognition ability of the model needs to be increased. In addition, the overall recognition speed of this paper does not exceed 0.98 seconds, which meets the requirements.

6 Conclusion

This article proposes a vision based automatic detection scheme for common surface defects of metal bars. Based on the research and analysis of successful systems and detection methods at home and abroad, the characteristics of bar surface defects and online operation are considered. According to the development of machine vision technology and image processing theory, a solution for detecting bar surface defects using a linear array camera is designed. We have conducted in-depth research and simulation of detection algorithms for key technologies and difficulties in visual inspection technology for surface defects of bar materials. The main research results of this paper are summarized as follows:

1) The hardware composition, working principle, and partial component manufacturing process of the bar surface detection system are provided, and the hardware selection principles and parameter selection criteria of the detection system are analyzed in detail. A lighting grayscale compensation scheme was designed and developed to address the uneven grayscale characteristics of bar surface imaging, fully utilizing the dynamic range of CMOS. Considering the harsh factory environment and the presence of multiple sources of interference, a bandpass filter with the same wavelength as the lighting source was used to address light pollution, which improved imaging clarity and provided rich details of the original image for subsequent recognition systems.

2) Designed an online visual inspection system for surface defects of bar materials based on linear array cameras and linear array light sources to meet the requirements of high-speed and high-resolution detection of bar materials; At the same time, in order to display various defects with high contrast, the imaging model of the surface of the rod was calculated, and the imaging mechanism was analyzed based on the light intensity distribution. The required lens focal length was calculated according to the resolution requirements for rod detection, and a suitable lens model was selected. In order to filter out other wavelengths and only allow laser wavelength light to enter the lens, a filter was designed and selected to ensure high-quality image acquisition.

3) Focusing on the surface of bar materials, the principle of genetic algorithm in image segmentation was studied and experiments were conducted. Genetic algorithm is not easily trapped in local optima and can effectively find the global optimal value of parameter space. It has fast convergence speed and strong adaptability, which verifies the effectiveness of genetic algorithm.

4) We conducted on-site debugging and implementation verification of the red steel surface defect detection system, which includes both hardware and software components. The installation and debugging of the linear array camera and linear array light source were carried out, and they were arranged at a small angle, so that the laser reflected light shining on the surface of the rod could enter the camera lens as much as possible, forming clear and high contrast images. Software debugging includes image size, exposure time, and the effectiveness of detection algorithms. The recognition experiment of surface defects on bar materials in the comparative experiment was completed through visual collection.

Although this article has achieved certain results in the theoretical and experimental research of the visual inspection system for surface defects of bar materials, due to limitations in conditions, further research is needed in the following areas to truly apply the system to industrial production lines:

1) Due to the harsh environment of high temperature, dust, and high interference in bar production, in order for the system to have on-site application value, dust prevention and insulation design must be carried out, and in-depth anti-interference research must be conducted.

2) Image preprocessing should undergo detailed algorithm design, but it takes up a long processing time and is difficult to ensure the speed requirements of the production line. Therefore, it is necessary to further optimize the preprocessing algorithm and achieve lightweight design of the image preprocessing algorithm.

3) The defect detection algorithm in the article is mainly aimed at scratch defects. For other defects, a richer defect dataset should be gradually established. Therefore, the next step is to detect narrow and fine defects to ensure effective detection of various jaw defects.

4) The experimental system is only a simple scheme layout and has not been made into a fixed mechanical structure. On site debugging is greatly affected by human factors, so further improvement is needed to increase the flexibility of system adjustment and facilitate experimentation and installation.

7 Acknowledgement

Research on the Application of Machine Vision in Intelligent Manufacturing (ZD24001).

References

[1] L.-H. Li, B. Liu, M. Zhong, Z.-N. Lian, C.-L. Yang, Research on Water Immersion Focused Ultrasonic Testing Technology of Metal Bars, Nondestructive Testing Technology 45(6)(2021) 17-21+43.

[2] Y.-H. Yue, T. Sun, Y.-C. Wang, Y.-H. Cao, X.-J. Lu, L.-F. Qin, Research on Visual Detection Method of Ash Content in Tailing Coal Water Based on GA-BP Neural Network, Journal of Yancheng Institute of Technology (Natural Science Edition) 36(3)(2023) 28-33.

[3] X.-H. Li, X.-X. Wang, An image Tamper Detection Method for CNN Mobile Devices Based on Pruning Compression, Computer Simulation 38(3)(2021) 83-86+91.

[4] C. Feng, Y. Jiang, X.-M. Zhao, Micro Defect Detection of Steel Ball Surface Based on Spatial Omni-Directional Light Source, Laser & Optoelectronics Progress 59(2)(2022) 311-320.

[5] S.-Q. Wang, Q. Ren, M. Shi, D.-M. Zhu, Product surface defect detection and segmentation based on anomaly detection, Journal of Graphics 43(3)(2022) 377-386.

[6] L.-M. Gao, S.-H. Jia, G.-L. Zhang, Y. Li, M.-Q. Yang, Detection Method for Surface Defects of Insulated Bearings Based on Improved Faster R-CNN, Bearing (4)(2023) 1-8. DOI:10.19533/j.issn1000-3762.2023.04.001.

[7] Y.-L. Chen, F.-C. Sun, X.-H. Guo, D.-W. Guo, L.-K. Liu, T.-T. Zhang, Research on Egg Crack Detection Based on EfficientNet, Journal of Chengdu University (Natural Science Edition) 39(4)(2020) 385-389.

[8] Z. Xiong, W.-G. Che, Y.-L. Bao, X.-T. Liu, Improved MobileNetV3 hot-rolled strip surface defect classification algorithm, Journal of Shanxi University of Technology (Natural Science Edition) 39(5)(2023) 30-37.

[9] Y. Xia, C. Luo, Y.-J. Zhou, L. Jia, A lightweight deep learning model for TFT-LCD circuits defect classification based on swin transformer, Optics and Precision Engineering 31(22)(2023) 3357-3370.

[10] S.-Q. Zhang, W.-Y. Shi, S.-W. Zhang, T.-T. Wang, Steel Surface Defect Detection Based on Improved YOLOv5 Algorithm, Science Technology and Engineering 23(35)(2023) 15148-15157.

[11] Y.-X. Wang, H.-W. Ge, Metal surface defect detection algorithm based on U^2-Net, Journal of Nanjing University (Natural Science) 59(3)(2023) 413-424.

[12] H.-M. Wang, X.-W. Zhu, X.-F. Yang, K. Xu, J.-L. Yu, Design of brush head feeding system based on visual detection and positioning, Journal of Jiangsu University (Natural Science Edition) 44(4)(2023) 431-436.

[13] C.-C. Chang, L.-Z. Pan, Y.-M. Xu, L.-Q. Wu, H. Tao, D. Liu, F. Chen, C. Liu, J.-Q. Zhu, Application and Progress of Computational Optical Imaging in Inertial Confinement Fusion, Acta Optica Sinica 43(22)(2023) 9-41.

[14] P. Xu, Y. Chao, J.-J. Zhu, Z. Cao, Design and optimization method of supplementary lighting system in visual inspection system, Laser Journal 44(6)(2023) 78-85.

[15] N. Zhang, Y.-X. Yuan, C.-L. Han, Y.-L. Li, Research on crack detection and localization algorithm for advancing face in coalmine roadways based on Mask R-CNN, Journal of Mining & Safety Engineering 40(5)(2023) 925-932.

[16] D.-D. Wang, X.-D. Zhang, Z.-G. Fan, R. Sun, A reverse fusion instance segmentation algorithm based on RGB-D, Journal of Graphics 42(5)(2021) 767-774.

[17] J.-F. Zhang, S. Zhang, A Improved Generation Adversarial Network Image Inpainting Method based on Attention Mechanism and Texture Structure Stratification, Journal of Mudanjiang Normal University (4)(2022) 22-28. DOI: 10.13815/j.cnki.jmtc(ns).2022.04.006.

[18] S.-J. Xu, M. Du, Z.-X. Duan, M.-H. Li, J.-Q. Han, Multi-Scale Regional Attention InfoGAN License Plate Recognition Network, Journal of Xi'an Jiaotong University 57(8)(2023) 206-218.

[19] R.-H. Zhang, X.-Q. Li, L.-Z. Dang, Y.-F. Wang, Y.-P. Hong, Design of Image Sensor FPN Noise Removal Algorithm Based on Vivado HLS, Integrated Circuits and Embedded Systems 23(12)(2023) 55-58+62.

[20] K.-Y. Zhang, N.-Y. Zhang, Z.-H. Jiang, W.-K. Dai, X. Zou, Performance analysis of different activation function in deep learning model, Cyber Security And Data Governance 42(S1)(2023) 149-156.

[21] J.-C. Zheng, Q.-L. Xie, T. Fang, J.-M. Han, C. Dong, A Visual Inspection Method for Pipe Surface Defect Based on Feature Point, Journal of Anhui University of Technology (Natural Science) 39(1)(2022) 21-24.

[22] Y.-D. Yang, P. Liu, G.-L. Zhou, Q.-J. Chen, Y. Li, Design and implementation of the blade profile detection system based on computer vision, Chinese Journal of Scientific Instrument 44(6)(2023) 213-222.