

Research Article

Detecting Gear Surface Defects Using Background-Weakening Method and Convolutional Neural Network

Liya Yu ¹, Zheng Wang ¹, and Zhongjing Duan ²

¹School of Mechanical Engineering, Guizhou University, Guiyang 550025, China

²Key Laboratory of Advanced Manufacturing Technology of Ministry of Education, Guizhou University, Guiyang 550025, China

Correspondence should be addressed to Zheng Wang; zhengwang0216@163.com

Received 15 July 2019; Accepted 17 September 2019; Published 19 November 2019

Academic Editor: Pawel Malinowski

Copyright © 2019 Liya Yu et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

A novel, efficient, and accurate method to detect gear defects under a complex background during industrial gear production is proposed in this study. Firstly, we first analyzed image filtering and smoothing techniques, which we used as a basis to develop a complex background-weakening algorithm for detecting the microdefects of gears. Subsequently, we discussed the types and characteristics of gear manufacturing defects. Under the complex background of image acquisition, a new model S-YOLO is proposed for online detection of gear defects, and it was validated on our experimental platform for online gear defect detection under a complex background. Results show that S-YOLO has better recognition of microdefects under a complex background than the YOLOv3 target recognition network. The proposed algorithm has good robustness as well. Code and data have been made available.

1. Introduction

In recent years, the demand for online quality inspection of mechanical parts under high-efficiency, high-precision manufacturing conditions has continued to grow with the rapid development of the manufacturing industry. Considering that a gear is a transmission part with a wide range of applications in the machinery industry, gear quality is particularly important in production. The development of the gear industry currently faces great challenges. Complex backgrounds, such as oil stains and dust particles, cannot be avoided in the gear manufacturing line. Identifying ways to accurately and efficiently identify gear surface defects in complex backgrounds and improve the quality inspection accuracy and production efficiency of gear production lines is important to advance the level of the manufacturing industry.

Traditional testing standards mainly detect the appearance size [1] and shape error [2] of parts, among which the error is maintained between 0.12 mm and 0.23 mm. In this paper, the gear defect is located by a deep learning algorithm, which lays a foundation for more precise quality inspection

such as the subsequent dimension measurement. The traditional detection of gear manufacturing defect detection is based mainly on machine vision [3, 4], in which the contour extraction algorithm is often used to extract the image features of a single gear. After extracting the features, the gear is detected and checked via template matching. This method not only processes the image at a slow speed but also has low detection efficiency because only one gear sample can be detected in each feature image. In the case of insufficient illumination or complex background, the traditional visual detection method relies heavily on the light source, and the background-weakening effect is poor. As a result, detection accuracy is greatly reduced.

With the rapid development of deep learning in daily life [5–7] and industrial fields [8, 9], many scholars attempt to apply deep learning methods for detecting part defects [10]. To ensure the quality of online defect detection, the network must exhibit fast positioning speed and high classification accuracy. At present, the mainstream target recognition networks include You Only Look Once (YOLOv3) [11], FAST-RCNN [12, 13], SSD, and FPN [14]. No complicated computation is required because

the YOLOv3 target detection network uses an end-to-end method to regress features. Previous research shows that YOLOv3 is faster than SSD, FPN, and other target recognition networks. However, the direct application of the YOLOv3 method to detect gear defects cannot satisfy the high accuracy requirements in industrial production. Therefore, the *k*-means clustering method is adopted to obtain the most suitable anchor to improve the positioning and detection accuracy of YOLOv3 for detecting gear defects. The background is weakened and denoised via image filtering and smoothing under the complex background of gear manufacturing, thereby improving the accuracy and detection efficiency of online gear defect detection. The proposed algorithm provides reference for the gear manufacturing industry to improve production efficiency, enhance product quality, and strengthen quality control capabilities.

A defect detection algorithm based on the deep learning algorithm of YOLOv3 for 62 gear line surface manufacturing is proposed in this study, which has the following main contributions:

- (1) By analyzing image filtering and smoothing technology aimed at the microdefects of gears under a complex background, this study proposes a complex background-weakening algorithm based on image filtering and smoothing, which weakens the background noise of oil and dust, among others
- (2) This study designs and opens source gear defect datasets for common defects, including missing teeth, broken teeth, surface scratches, and normal gear
- (3) This study proposes an improved network for online gear defect detection called S-YOLO. This network is created by combining the types and characteristics of defects during the actual manufacturing of gears under the complex background of image acquisition on the factory production line. S-YOLO improves detection accuracy

The main structure of this paper is as follows. The second section mainly describes the related works on gear manufacturing defects and sorts the techniques of gear running fault and fatigue damage defect detection. The third section proposes a background-weakening algorithm for the complex background in gear manufacturing. The fourth section introduces the deep learning network target detection model, which is based on the YOLOv3 model for improvement and model training. The fifth section designs and manufactures an online detection platform for industrial defects. The sixth section designs and makes the gear defect dataset and compares and analyzes the experimental results. The final section summarizes the research content.

2. Related Work

In the research on fault diagnosis during gear runtime, Mączak and Jasiński [15] discussed the simulation model of the helical gearbox and analyzed a phenomenon during

the tooth-meshing process in the presence of manufacturing and assembly errors. This work proposed a kind of gear fault diagnosis method based on the model. The detection method is simple, and the detection speed is fast. However, the effect of gear detection in large-volume motion on the production line is unknown. Gandarias et al. [16] took pressure reading as a standard image processing technique with the new high-resolution pressure sensor. It connects the tactile sensor with the robot detector with high resolution and realizes the image recognition of the contact object via a convolutional neural network (CNN) and migration learning. Lu et al. [17] applied the improved CNN model to an embedded system composed of signal acquisition and processing circuits and proposed a method for on-site motor fault diagnosis. A heterogeneous computing framework was proposed, and an integrated embedded system was designed based on the analysis of different motor signals. This method uses artificial intelligence technology to provide a solution for the field motor fault diagnosis on small, flexible, and convenient handheld devices. Cheng and Hu [18] proposed a method based on a physical model to detect the damage quantification of the planetary gear set. The performance of the feature in the damage evolution tracking was analyzed via the double-sample test method, and the state monitoring of the planetary gear transmission system was realized. Nabih et al. [19] experimentally verified the dynamic model of the single-stage gear transmission system and analyzed the effect of the perforation on TE. The results proved that a simple perforation model can reproduce the actual vibration caused by the failure of the perforation surface. Younes et al. [20] proposed a vibration acoustic signal analysis theory. The theory uses the feature extraction and classification of acoustic signals to accurately identify the defects of gears and bearings, but its algorithm cannot identify the exact location of the defects.

In research on gear defect detection through data acquisition and signal processing during gear operation, Zhao et al. [21] proposed a gearbox health evaluation framework based on R/C (run-up/coast-down) signal analysis by studying the mechanical vibration information. A feature enhancement scheme based on sparse guidance was proposed to extract the weak phase jitter associated with gear defects and detect the damage position of the gear. Kidar et al. [22] provided the crack characteristics in the vibration signal through the numerical model of the data. The analysis of the phase estimated using the Hilbert method and the signal parameters estimated via the sliding window-based rotation invariant technique were compared to achieve the detection of gear cracks. A sensor position optimization method based on finite element analysis and spectrum analysis was proposed in [23]. The existing two nonlinear models of mechanical rotating parts were solved, and the dynamic response of the whole system under defect excitation was used to determine the predictive maintenance for defect detection in the optimal sensor location. The defect of mechanical rotating parts was accurately detected. Moreno et al. [24] proposed various signal processing strategies for the

detection and quantification of early gear defects. A comparison among the early detection capabilities of the microphone, accelerometer, and LDV sensors verified that the acoustic signal was the first method to detect the initial progressive crack of the gear (detecting a 1.3 mm long crack). Using a microphone signal had obvious advantages, but the result was sensitive to speed and torque. The pitting of gears was tested, and the vibration data was recorded in [25]. The application of vibration-based time, frequency, cepstrum, wavelet transform, and other methods in each set of experimental data, pitting fault, and the progress of pitting failure in gears were reviewed as well.

In research on detecting small defects of gears, Liu et al. [26] aimed to address the high cost, low efficiency, slow speed, and low precision of manual detection of automobile bevel gear surface defects and dimensional measurement. They studied and analyzed the three effective algorithms—neighborhood means difference method, circular approximation method, and fast rotation positioning method. A comprehensive bevel gear quality detection system was developed based on multicamera vision technology, which could simultaneously detect and measure the size of bevel gear surface defects. Fedala et al. [27] aimed to improve the detection and recognition ability of gear defects by extracting the features of the angular frequency domain of angular acceleration sampling, transmission error, and instantaneous angular velocity. SVM was then used to classify and realize gear fault detection under normal and non-stationary states. To isolate the defect signal from the measured signal, Djebala et al. [28] proposed a gear defect detection method based on wavelet multiresolution analysis and Hilbert transform. Experiments show that, in contrast with the commonly used analysis tools, this new method can isolate defect frequency, which enables the detection of small or combined defects. Focusing on the internal meshing gear defects, Zhang and Fan [29] proposed a universal formula for the identification and conducted the closed defects of the N-lobed noncircular gears (N-LNG) positioning function. The closed condition of the positioning function was satisfied by introducing two correction parameters: proportional and controllable. The controllable correction parameters were further verified and improved on the basis of the relationship between the inner pitch curve and the curvature radius of the outer pitch curve of the inner meshing of N-lobed noncircular gears. The method was applied in several numerical examples, and the simulation results showed that the method can effectively identify and conduct the closed defects of the N-LNG positioning function.

In the field of gear defect detection, many scholars conducted relevant theoretical research on gear operation faults, surface defects, and other aspects. However, research on surface manufacturing defects during the manufacturing of gears and high-speed online defect detection with numerous parts requires further improvement.

3. Complex Background-Weakening Algorithm

Substantial oil, dust, and other debris accumulate on the conveyor during gear production, and they complicate the back-

ground of the gear image sample to be tested. Accurately identifying the minor manufacturing defects on the gears, such as scratches and pinion broken teeth, is difficult. Such defects are called background noise. The images collected by the camera also generate noise due to the randomness of the photon flux and the fact that the gears are in motion on the conveyor belt. If the real pixel value $g_{r,c}$ is disturbed by the noise $n_{r,c}$, the gray value obtained is as follows:

$$\hat{g}_{r,c} = g_{r,c} + n_{r,c}. \quad (1)$$

Noise $n_{r,c}$ is assumed to be smooth in the whole picture; that is, the noise is independent of the position of the pixels on the image. This noise, which is called stationary noise, is equally distributed for each pixel in the picture.

Two methods are commonly used to weaken the two kinds of noise in the picture collected during gear production: time-domain average denoising [30] and spatial average denoising [31]. Time-domain averaging captures and averages multiple images of the same scene. If n images are collected, then time-domain average is obtained as follows:

$$g_{r,c} = \frac{1}{n} \sum_{i=1}^n \hat{g}_{r,c,i}, \quad (2)$$

where $\hat{g}_{r,c,i}$ denotes the grayscale value at position (r, c) on the i image. The time-domain average method effectively reduces noise, and the variance of the noise is reduced to original $1/n$. To suppress noise, the method must collect images in the same scene. For online defect detection, the acquisition of multiple images in the same scene improves the accuracy of defect identification. However, it greatly increases the running time of the algorithm, thereby reducing the overall detection efficiency.

Therefore, the spatial average is used for denoising by taking a filter with a pixel of $(2n + 1) \times (2m + 1)$ and traversing the same image. Depending on the operation and the filter, the filtering algorithm includes mean filtering, block filtering, Gaussian filtering, and median filtering. Among them, mean filtering and Gaussian filtering are the most commonly used filtering algorithms. Mean filtering can be expressed as

$$h_{r,c} = \begin{cases} \frac{1}{(2n + 1) \times (2m + 1)} & |r| \leq n \wedge |c| \leq m \\ 0 & \text{Other} \end{cases}, \quad (3)$$

where (r, c) denotes the pixel position of the image and m and n are the parameters that determine the length and width of the filter.

If the original image matrix is

$$A = \begin{bmatrix} \alpha_{1,1} & \alpha_{1,2} & \cdots & \alpha_{1,k} \\ \vdots & & \ddots & \vdots \\ \alpha_{j,1} & \alpha_{j,2} & \cdots & \alpha_{j,k} \end{bmatrix}, \quad (4)$$

then the filtered matrix is

$$B = \begin{bmatrix} \beta_{1,1} & \beta_{1,2} & \cdots & \beta_{1,k} \\ \vdots & \vdots & \ddots & \vdots \\ \beta_{j,1} & \beta_{j,2} & \cdots & \beta_{j,k} \end{bmatrix}. \quad (5)$$

In the actual operation process, the input image is usually a square, so $m = n$. The pixel $\alpha_{p,q}$ in matrix A is then processed through the mean filter with the size of $(2n + 1) \times (2n + 1)$ to obtain $\beta_{p,q}$ in B :

$$\beta_{p,q} = \begin{bmatrix} \alpha_{p-n,q-n} & \cdots & \alpha_{p-n,q+n} \\ \vdots & \ddots & \vdots \\ \alpha_{p+n,q-n} & \cdots & \alpha_{p+n,q+n} \end{bmatrix}_{(2n+1)(2n+1)} \times \frac{1}{(2n+1)^2} \begin{bmatrix} 1 & \cdots & 1 \\ \vdots & \ddots & \vdots \\ 1 & \cdots & 1 \end{bmatrix}_{(2n+1)(2n+1)}. \quad (6)$$

As shown in Formulas (3) and (6), the averaging filter actually averages the pixels in the effective calculation range and assigns them to the middle value of the filtering window. For the oil stain and dust background of the gear production workshop, the mean filter averages pixel values, such as oil and dust, with the surrounding background pixels. It blurs the oil, dust, and other small particles. It also highlights the position and feature information of the gear in the whole image to prepare for subsequent feature extraction.

Although the mean filter weakens small particles, such as oil stains and dust in the background, most of the stationary noises in the image due to the principle of lens imaging appear in the form of high-frequency fluctuation of gray value. The suppression of high-frequency noise via filtering is not satisfactory. Therefore, to maximize the suppression of the influence of high-frequency stationary noise, the Gaussian filter is used for secondary image smoothing. As such, the eigenvalue of the processed image becomes easy to extract. The 1D Gaussian filter can be expressed as

$$g_{\sigma}(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/(2\sigma^2)}. \quad (7)$$

The two-dimensional Gaussian filter applied to image processing can be expressed as

$$g_{\sigma}(r, c) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-(r^2+c^2)/(2\sigma^2)} = g_{\sigma}(r)g_{\sigma}(c). \quad (8)$$

After the first mean filtering of complex background images, the effect of oil, dust, and other abrupt noises in the background is weakened. The high-frequency noise in the complex background of the image is weakened after the second Gaussian filtering, and the gear body in the relative

Type	Filters	Size	Output
Convolutional	32	3 x 3	256 x 256
Convolutional	64	3 x 3 / 2	128 x 128
Convolutional	32	1 x 1	128 x 128
Convolutional	64	3 x 3	
Residual			
Convolutional	128	3 x 3 / 2	64 x 64
Convolutional	64	1 x 1	64 x 64
Convolutional	128	3 x 3	
Residual			
Convolutional	256	3 x 3 / 2	32 x 32
Convolutional	128	1 x 1	32 x 32
Convolutional	256	3 x 3	
Residual			
Convolutional	512	3 x 3 / 2	16 x 16
Convolutional	256	1 x 1	16 x 16
Convolutional	512	3 x 3	
Residual			
Convolutional	1024	3 x 3 / 2	8 x 8
Convolutional	512	1 x 1	8 x 8
Convolutional	1024	3 x 3	
Residual			
Avgpool	Global		
Connected	1000		
Softmax			

FIGURE 1: YOLOv3 basic network Darknet-53 [11].

image is highlighted, allowing for the easy extraction of the gear body's features.

4. Improved Construction and Training of YOLOv3 Network

4.1. Characteristics of YOLOv3 Network Structure. The YOLOv3 network model uses an end-to-end network architecture implemented in a CNN. The basic network structure is shown in Figure 1.

Its network first divides the input image into $S \times S$ grids and the image by clustering. If the center point of an object in the image falls in the YOLO-divided grid, then the grid is responsible for predicting the object. Each grid is responsible for predicting B bounding boxes and the confidence of the bounding boxes. The confidence reflects the probability of containing objects in the bounding box predicted by the network model and the accuracy of the predicted position of the bounding box, which can be expressed as

$$\text{Confidence} = \text{Pr}(\text{Object}) \times \text{IOU} \frac{\text{truth}}{\text{pred}}, \quad (9)$$

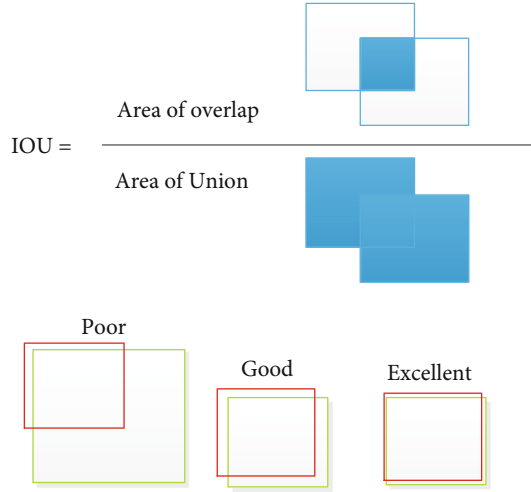


FIGURE 2: IOU evaluation diagram.

where IOU (Intersection over Union) represents the intersection ratio of the real target bounding box and the predicted target bounding box, which can be represented by Figure 2. If an object exists in the grid, $\text{Pr}(\text{Object}) = 1$, then

$$\text{Confidence} = \text{IOU} \frac{\text{truth}}{\text{pred}}. \quad (10)$$

Otherwise, $\text{Pr}(\text{Object}) = 0$, that is,

$$\text{Confidence} = 0. \quad (11)$$

In the YOLOv3 network, each bounding box predicts five values, including (x, y, w, h) and confidence, where x, y represents the coordinates of the center point of the predicted bounding box and w, h are the width and the height of the bounding box. Confidence is the IOU that predicts the bounding and the real bounding boxes.

Each grid predicts the probability of C condition categories, that is, the probability of the mesh containing objects belonging to a certain category. $\text{Pr}(\text{Class}_i | \text{Object})$. Finally, the conditional probability is multiplied by confidence, and the probability that a certain type of object appears in the box and the degree of fit of the bounding box to the object are obtained:

$$\begin{aligned} & \text{Pr}(\text{Class}_i | \text{Object}) \times \text{Pr}(\text{Object}) \times \text{IOU} \frac{\text{truth}}{\text{pred}} \\ & = \text{Pr}(\text{Class}_i) \times \text{IOU} \frac{\text{truth}}{\text{pred}}. \end{aligned} \quad (12)$$

In the design of loss function, the YOLO network takes the form of a weighted summation of the partial loss functions. By weighing the coordinate error, IOU error, and classification error and summing them, the total loss function is calculated and can be expressed as

$$\text{loss} = \sum_{i=0}^{s^2} \text{coordErr} + \text{iouErr} + \text{clsErr}. \quad (13)$$

The loss of the predicted center coordinates is expressed as

$$\lambda_{\text{coord}} \sum_{i=0}^{s^2} \sum_{j=0}^B \ell_{ij}^{\text{obj}} [(x_i - x_i^\wedge)^2 + (y_i - y_i^\wedge)^2]. \quad (14)$$

The loss of the width and the height of the predicted bounding box is expressed as follows:

$$\lambda_{\text{coord}} \sum_{i=0}^{s^2} \sum_{j=0}^B \ell_{ij}^{\text{obj}} \left[(\sqrt{w_i} - \sqrt{w_i^\wedge})^2 + (\sqrt{h_i} - \sqrt{h_i^\wedge})^2 \right] \quad (15)$$

where λ_{coord} denotes the weight factor of the coordinate error in the overall loss function.

The loss made to the forecast category is expressed as

$$\sum_{i=0}^{s^2} \ell_i^{\text{obj}} \sum_{j=0}^B [(p_i(c) - p_i^\wedge(c))^2]. \quad (16)$$

The loss of confidence in the prediction is expressed as follows:

$$\sum_{i=0}^{s^2} \sum_{j=0}^B \ell_{ij}^{\text{obj}} [(c_i - c_i^\wedge)^2] + \lambda_{\text{noobj}} \sum_{i=0}^{s^2} \sum_{j=0}^B \ell_{ij}^{\text{obj}} [(c_i - c_i^\wedge)^2], \quad (17)$$

where C is the confidence score; \hat{C} is the intersection of the predicted bounding box and the basic fact, when an object exists in a cell; and ℓ_{ij}^{obj} is equal to 1; otherwise, it is 0; λ_{noobj} represents the confidence weight when no object exists in the bounding box [10].

4.2. Improved YOLOv3 Network. The original YOLOv3 network uses a CNN, so the image is extracted through multiple convolutional layers for abstract feature extraction. Finally, the image is classified and predicted. Combining the types and characteristics of defects during actual gear manufacturing and the complex background of image acquisition in the factory production line, an improved online defect detection network for YOLOv3 gear is proposed. This network is called S-YOLO, which stands for smoothing-YOLOv3. The network structure is shown in Figure 3.

In the network structure of S-YOLO, the end-to-end Darknet-53 convolutional network formed in YOLOv3 is maintained. Moreover, an image-smoothing layer is added at the front end of the network to weaken the background noise of gear image collection during production.

In the smoothing layer, an average filter with pixel 8×8 is used to filter and smoothen the collected image for the first time. This process is aimed at weakening the influence of impurities, such as oil and fine dust particles, in the image. A Gaussian filter with a pixel of 3×3 is then used for the secondary smoothing of the image. This filter mainly reduces

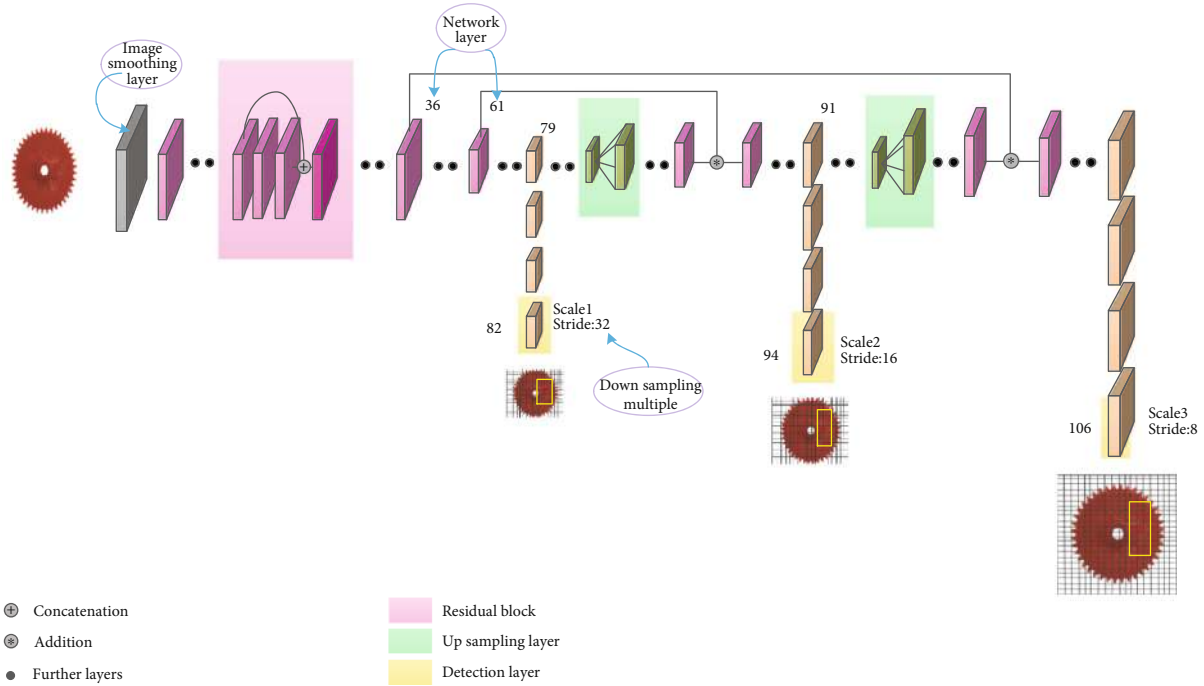


FIGURE 3: S-YOLO detection principle.

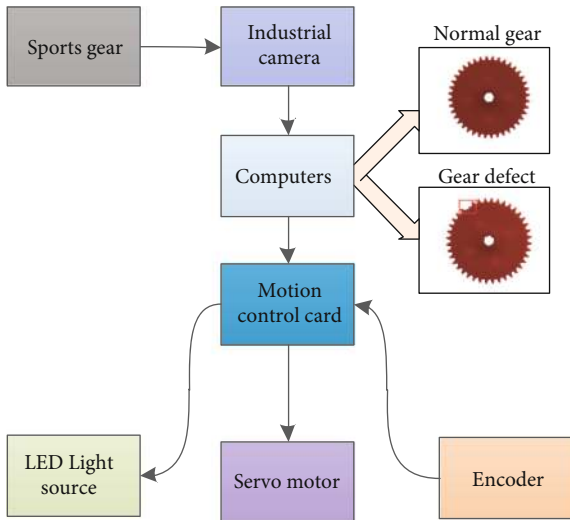


FIGURE 4: Flow chart of the experimental platform.

the high-frequency noise in the collected image and further reduces the influence of oil, dust particles, and other impurities in the gear production workshop.

After passing the smooth layer, the pixel size and gear defect characteristics remain unchanged. The following YOLOv3 network uses three different scale feature maps for defect detection. As shown in Figure 3, a scale detection result is obtained through several yellow convolution layers after the 79th convolutional layer. The input image size during the experiment is 416×416 pix. Hence, the feature image pixel size at this time is 13×13 pix. The receptive field of the feature map is relatively large at this time because the downsampling factor is high, which is suitable

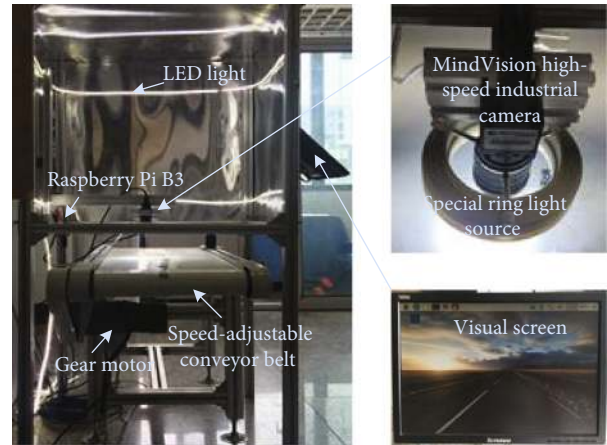


FIGURE 5: Experimental platform.

for detecting relatively large defect size in the image. The network starts upsampling from the feature map of the 79th layer. It then fuses with the 61st layer feature map to obtain the 91st layer of the finer-grained feature map. After several convolution layers, the feature map 16 times of the input image is obtained. It has a medium-scale receptive field and is suitable for detecting objects with medium defect size. Finally, the 91st layer feature map is again upsampled and merged with the 36th layer feature map to obtain a feature map that is downsampled for eight times from the input image. It has the smallest receptive field and is suitable for detecting small defect sizes.

4.3. *k*-Means Clustering-Based A Priori Box Acquisition. Although the YOLO network itself can improve the value

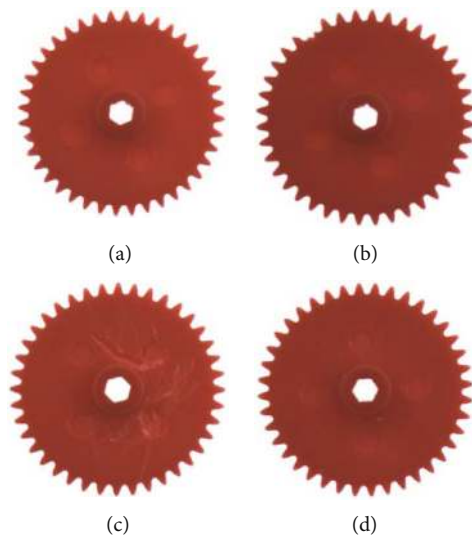


FIGURE 6: Common defects in the gear manufacturing process. (a) Break. (b) Lack. (c) Scratch. (d) Normal.

TABLE 1: Distribution of common manufacturing defect datasets for gears.

Defect type	Broken tooth		Missing tooth		Scratch		Normal	
	Test	Train	Test	Train	Test	Train	Test	Train
Number of images	100	900	100	900	100	900	100	900

of the IOU and constantly adjust the size of the bounding box via training, allowing the network to modify through a large amount of data will slow down the network training and prevent the value of the IOU from gaining substantial improvement. With the gear training dataset as a basis, the k -means method is used to find the anchors of the a priori box that best fits the size of the gear defect. The standard k -means method uses Euclidean distance, and this usage will result in large boxes that generate more errors than small boxes. Therefore, Formula (18) is used to represent the distance and obtain a large IOU value in network prediction:

$$d(\text{box}, \text{centroid}) = 1 - \text{IOU}(\text{box}, \text{centroid}). \quad (18)$$

5. Online Platform for Industrial Defect Detection

Figure 4 depicts the system flow chart of the online testing platform for gear manufacturing defects designed by the research group. Figure 5 is an online test platform for gear manufacturing defects built by the research team [10, 32]. This platform includes the conveyor belt, data processor, data acquisition sensor, light source, and other mechanical supports, wherein the touch display for inputting and displaying data is the 32-inch industrial touch screen. The vision sensor device uses the MindVision high-speed industrial camera with an electronic rolling shutter, which can collect high-speed moving samples for real-time testing. The data processor is the Raspberry Pi B3. To ensure sufficient light in the system box, a band-shaped ambient light source LED

with adjustable brightness is installed. A dedicated circular light source of Microscope LED Ring Light is installed outside the industrial camera to fill the test sample with light and to obtain a clear sample image. The device uses a variable speed motor to drive the conveyor belt. The outside of the box is equipped with a display for visualizing the test results. The Dell workstation of GPU1080 graphics card, which is mainly used for data analysis, is used to reduce the computational load of data processor. At the same time, Raspberry PI B3 has a wireless communication module, which can realize end-to-end communication between the test experimental platform and the workstation. The SQL SERVER 2008 R2 database is installed on the workstation to realize real-time local data capturing and automatic real-time data storage in the cloud.

The gear is transported to the field of view of the industrial camera's lens through the conveyor belt. After detecting the gear passing, the fiber optic sensor sends a trigger pulse to the image acquisition part. The image acquisition part then sends a start pulse to the industrial camera and the illumination system according to the pre-set program and delay. Industrial cameras begin to capture images, and the Microscope LED Ring Light's dedicated ring light source provides illumination that matches the exposure time of the industrial cameras. After capturing the image, the image acquisition of the camera receives the analog signal and digitizes it via an analog to digital conversion. The image acquisition part stores the digital image in the processor or computer memory. The processor then processes, analyzes, and recognizes the collected gear image. It then obtains and saves the detection result.

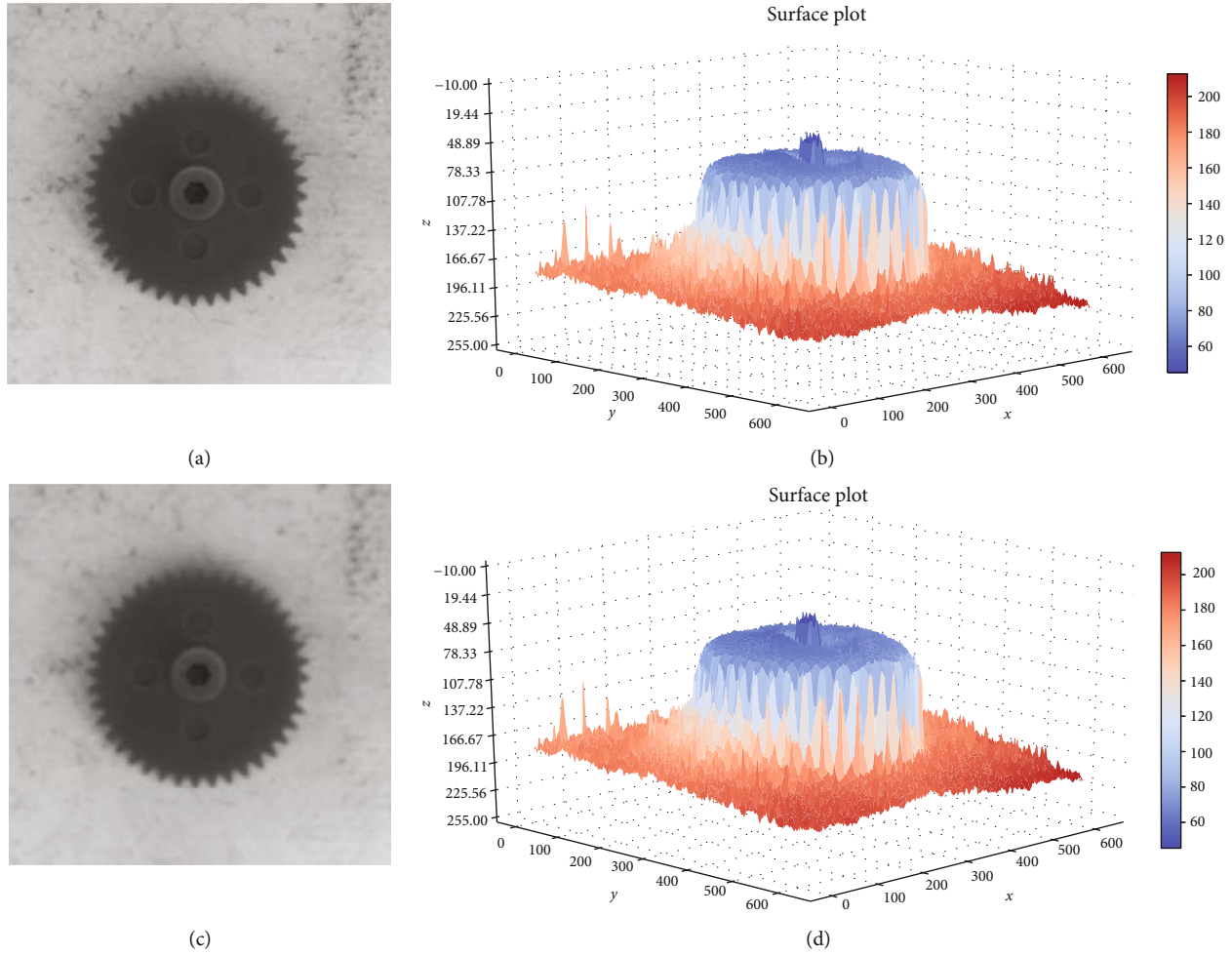


FIGURE 7: Comparison of image smoothness after mean filtering. (a) Original grayscale image. (b) The pixel gradient of the original grayscale image. (c) Image after mean filtering. (d) Image pixel gradient after mean filtering.

6. Experimental Results and Analysis

6.1. Production of Gear Datasets. During gear manufacturing, the bluntness of the turbine hob or the uneven material of the gear billet often causes gear tooth surface tear, tooth fracture, and gear surface scratches, among others, as shown in Figure 6.

Gear defect datasets $P = \{L, B, S, N\}$ are collected according to the types of defects commonly found in gear production. The four types of datasets are broken tooth image set $B = \{B_1, B_2, \dots, B_{300}\}$, missing tooth image set $L = \{L_1, L_2, \dots, L_{300}\}$, gear surface scratch image set $S = \{S_1, S_2, \dots, S_{300}\}$, and normal image set $N = \{N_1, N_2, \dots, N_{300}\}$.

Data enhancements can enrich small datasets or poorly diverse datasets. Common data enhancement methods include color jittering, PCA jittering, random scale, random crop, and horizontal/vertical flip. After collecting 300 pieces of gear data for each type of gear through industrial cameras, the images are rotated at random angles to achieve data enhancement. Finally, 1000 pieces of image data for each type are obtained, thereby collecting a total of 4000 pieces of gear image data. The specific data distribution is shown in Table 1.

6.2. Double Filtering Background Weakening. The effect of mean filtering on image noise removal in the complex background is considered. As shown in Figure 7, the original grayscale image has background noises, such as dust and oil stains, as illustrated in Figure 7(a). These noises have a certain influence on the later gear feature extraction. After the mean filtering operation, as shown in Figure 7(c), the background noise is partially weakened, and the degree of weakening depends on the convolution kernel size of the mean filter. After the mean filtering operation, the entire part of the gear still has all the features required for defect detection. As indicated in the comparison between Figures 7(b) and 7(d), the smoothness of the image increases after mean filtering. Moreover, the overall pixel gradient tends to be smooth, which is a good data condition for defect recognition and classification via the deep learning algorithm.

The comparison in Figure 8 shows that the high-frequency noise in the image is suppressed after the secondary filtering by the Gaussian filter, and the low-frequency part of the image is highlighted. Thus, the effect of the main part of the gear in the protruding image is achieved, which lays the foundation for the following feature extraction.

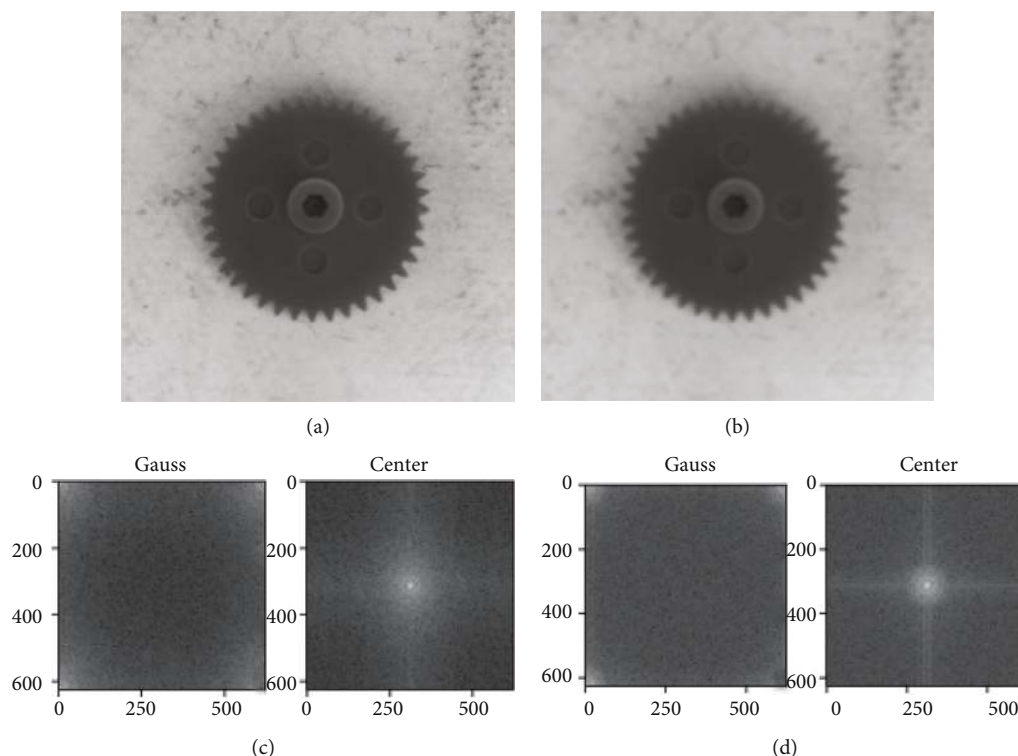


FIGURE 8: Comparison with the Fourier transform of the Gaussian filtered image. (a) Original. (b) Gaussian filtered. (c) The original Fourier transform. (d) The gaussian filtered Fourier transform.

TABLE 2: Training part of the main parameter settings.

Parameter	Numerical value	Parameter	Numerical value	Parameter	Numerical value	Parameter	Numerical value
Batch	64	Angle	0	Burn_in	2000	Scales	0.1, 0.1
Subdivisions	32	Saturation	1.5	Max_batches	50000	Learning_rate	0.001
Momentum	0.9	Exposure	1.5	Policy	Steps	Random	1
Decay	0.0008	Hue	0.1	Steps	4000, 4500	Jitter	3

TABLE 3: The effect of different k values on the clustering effect of datasets.

k value	5
Accuracy	82.32%
Boxes	(47, 42), (68, 45), (40, 60), (37, 25), (31, 43)
Ratios	[0.67, 0.72, 1.12, 1.48, 1.51]
k value	7
Accuracy	82.32%
Boxes	(72, 37), (37, 25), (46, 36), (39, 60), (66, 52), (42, 44), (50, 43)
Ratios	[0.65, 0.95, 1.16, 1.27, 1.28, 1.48, 1.95]
k value	9
Accuracy	86.06%
Boxes	(39, 61), (58, 53), (54, 38), (33, 25), (43, 45), (41, 26), (31, 41), (46, 40.5), (73, 38)
Ratios	[0.64, 0.76, 0.96, 1.09, 1.14, 1.32, 1.42, 1.58, 1.92]

6.3. *Experimental Results under Different k -Means.* To constantly adjust the size of the bounding box, the value of the IOU must be increased. Under the parameter settings in Table 2, the clustering effect of different k values on training

data in different k -means algorithms is tested. The experimental results are listed in Table 3.

The clustering effect is conducive to the gear defect situation. S-YOLO allocates three different sizes of a priori boxes

TABLE 4: A priori boxes for different receptive field assignments.

Feature map	13 × 13	26 × 26	52 × 52
Receptive field	Big	Medium	Small
Prior box	(58, 53) (54, 38) (73, 38)	(39, 61) (43, 45) (46, 40.5)	(33, 25) (31, 41) (41, 26)

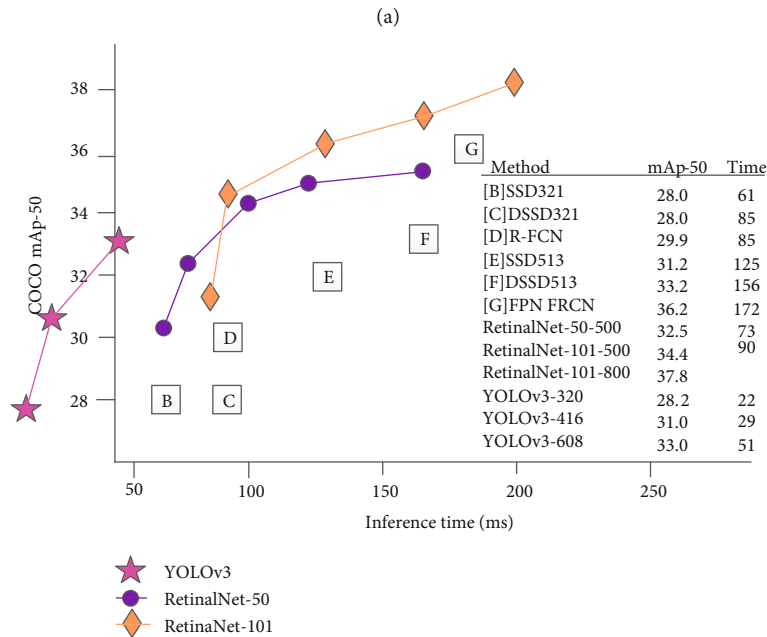
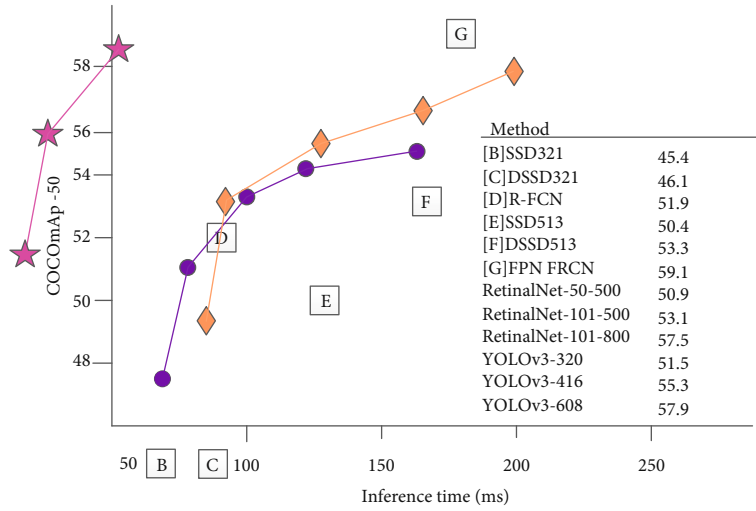


FIGURE 9: Comparison between the detection speed and accuracy of YOLOv3 and other algorithms [11].

for each scale when performing three-scale feature detection. When the k value is equal to 9, nine kinds of a priori boxes are available for allocation. Hence, when assigning, three a priori boxes may be assigned for each scale feature. Details are shown in Table 4.

At the smallest feature map 13×13 (larger receptive field), the larger priority box (58, 53) (54, 38) (73, 38) is applied to the feature map, which is suitable for detecting surface scratches with large defect sizes. Medium feature map 26×26 (medium receptive field) applies a medium

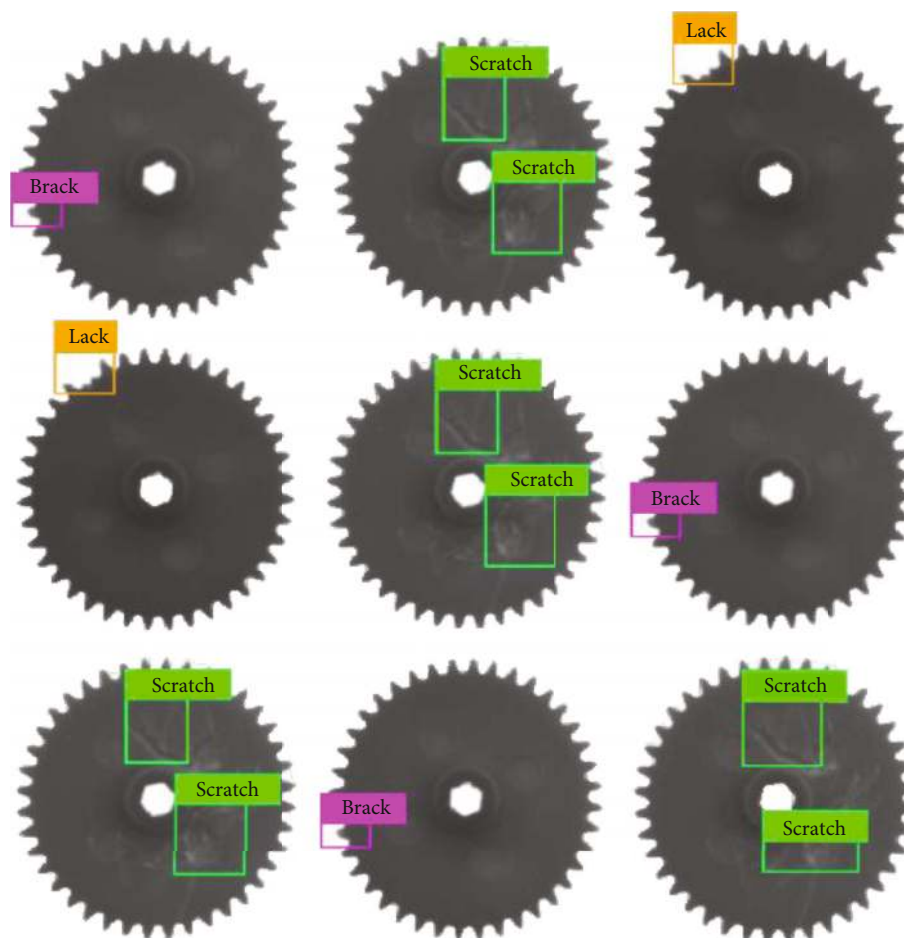


FIGURE 10: S-YOLO network test results without background interference.

priority box (39, 61) (43, 45) (46, 40.5), which is suitable for detecting objects of medium-size defects. A smaller priority box (33, 25) (31, 41) (41, 26) is applied on the larger feature map 52×52 (small receptive field), which is suitable for detecting objects with small defect sizes, such as broken and missing teeth. When training, the model training using the cluster generated by $k = 9$ can significantly shorten the model training time and improve the model IOU value.

6.4. Analysis of Gear Defect Detection Results. Figure 9 shows the combined performance of the YOLOv3 object detection network and other mainstream networks on the COCO datasets. After modifying the YOLOv3 model, the S-YOLO target detection model is trained. Through model training, the gear defect detection verification is finally performed on the detection platform. Figure 10 shows the detection of the S-YOLO model in the absence of complex background conditions, such as oil stains and dust particles. Figure 11 depicts the testing situation of the S-YOLO model when oil and dust particles are filled in the background in the simulation of the actual factory production on the platform for high-speed gear manufacturing defect testing. The experimental test results are provided in Table 5.

A comparison between Table 5 and Figure 11 shows that proposed network S-YOLO increases the complex back-

ground of gear manufacturing while retaining the advantages of traditional YOLOv3, which are detection speed and multiscale prediction. The image-smoothing layer and k -means clustering method are used to assign the most priority box to multiscale detection, which greatly inhibits the influence of the complex background on the detection effect of the model. It also makes the model lose stability and improves the average IOU value during training. S-YOLO is applied to the high-speed gear manufacturing defect detection experimental platform. Its classification effect reaches 100% accuracy, and the average confidence reaches 93.96%. The algorithm has good robustness.

7. Summary

The manufacturing defects in the gear manufacturing process were analyzed and studied. A dual-filtering background-weakening algorithm was proposed to address oil pollution, dust, and other complex backgrounds during production. Combined with the deep learning algorithm and target detection network model of YOLOv3, the network model of S-YOLO for gear manufacturing defect detection was proposed. Nine optimal anchor values were obtained via k -means clustering, which reduced the declining fluctuation of loss during model training and improved the

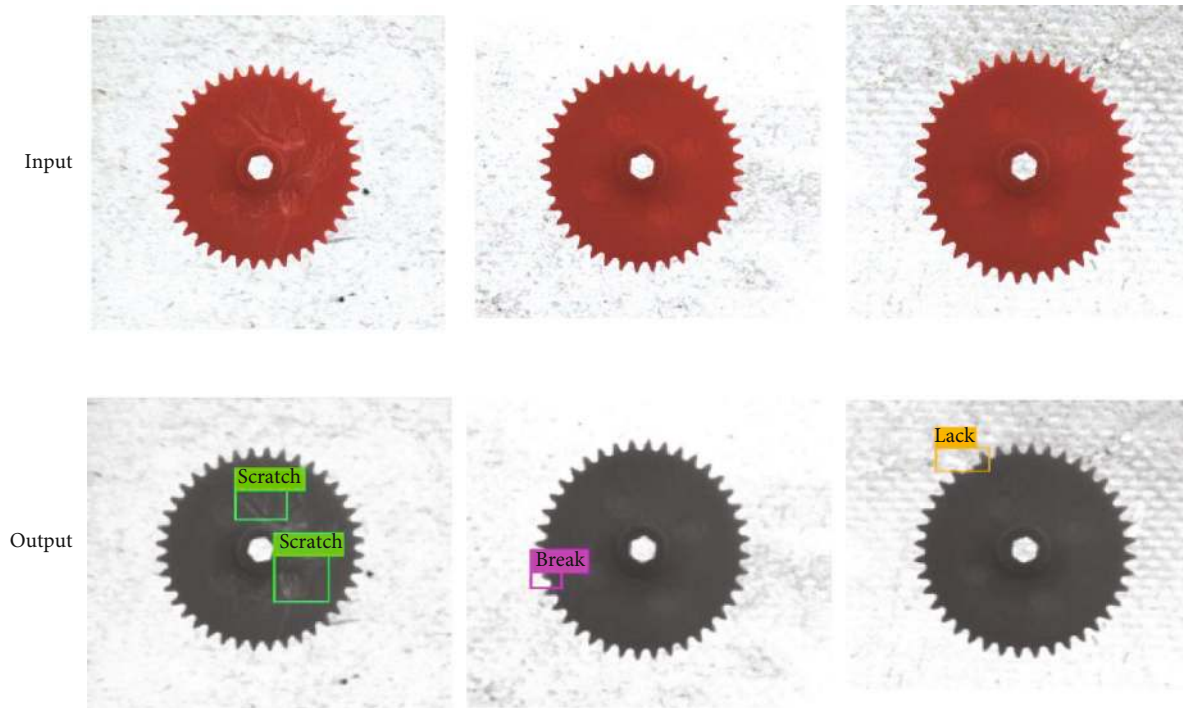


FIGURE 11: S-YOLO network defect detection when background interference exists on the detection platform.

TABLE 5: S-YOLO network detection gear defect situation.

Defect type	Number of defects in the test set	S-YOLO classification accuracy	YOLOv3 classification accuracy	Average confidence of each type of defect in S-YOLO	YOLOv3 average confidence of each type of defect
Broken tooth	300	100%	98.2%	98.30%	88.4%
Missing tooth	300	100%	98.7%	97.05%	84.9%
Surface scratch	300	100%	98.6%	86.54%	80.2%

average IOU value of the model. The gear manufacturing defect dataset was established using the data enhancement method. The application of the proposed algorithm and model was verified by building an online platform for industrial defect detection. The results showed that the proposed algorithm can meet actual production requirements.

Data Availability

Code and data have been made available at <https://github.com/Yuli-Ya/Detecting-Gear-Surface-Defects>.

Conflicts of Interest

The authors declare no conflict of interest.

Authors' Contributions

L.Y. and Z.W. worked on conceptualization and data curation. Z.W. performed the methodology. L.Y. worked with software and resources and did writing (original draft prepara-

tion and review and editing) and funding acquisition. Z.W, L.Y., and Z.D. carried out the validation.

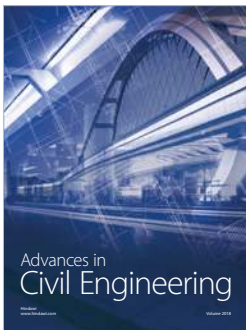
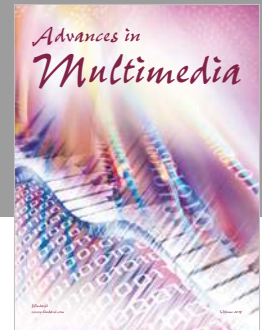
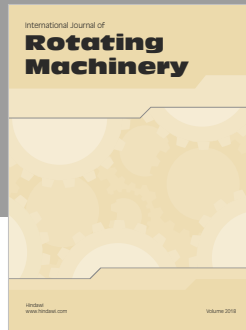
Acknowledgments

This work was supported in part by the National Natural Science Foundation of China under Grant Nos. 91746116 and 51741101, the Science and Technology Project of Guizhou Province under Grant Nos. [2017]2308, [2015]4011, and [2016]5013, and the Collaborative Innovation of Guizhou Province (YJSCXJH[2018]052).

References

- [1] M. J. Robinson, J. P. Oakley, and M. J. Cunningham, "The accuracy of image analysis methods in spur gear metrology," *Measurement Science and Technology*, vol. 6, no. 7, pp. 860–871, 1995.
- [2] M. A. Ayub, A. B. Mohamed, and A. H. Esa, "In-line inspection of roundness using machine vision," *Procedia Technology*, vol. 15, pp. 807–816, 2014.

- [3] Q. Guo, C. Zhang, H. Liu, and X. Zhang, "Defect detection in tire X-ray images using weighted texture dissimilarity," *Journal of Sensors*, vol. 2016, Article ID 4140175, 12 pages, 2016.
- [4] J. Yuan, Q. Wang, and B. Li, "A flexible and high precision calibration method for binocular structured light scanning system," *The Scientific World Journal*, vol. 2014, no. 8, Article ID 753932, 8 pages, 2014.
- [5] J. Yang and G. Yang, "Modified convolutional neural network based on dropout and the stochastic gradient descent optimizer," *Algorithms*, vol. 11, no. 3, pp. 28–43, 2018.
- [6] G. Yang, J. Yang, W. Sheng, F. Junior, and S. Li, "Convolutional neural network-based embarrassing situation detection under camera for social robot in smart homes," *Sensors*, vol. 18, no. 5, pp. 1530–1553, 2018.
- [7] J. Yang, W. Sheng, and G. Yang, "Dynamic gesture recognition algorithm based on ROI and CNN for social robots," in *2018 13th World Congress on Intelligent Control and Automation (WCICA)*, pp. 389–394, Changsha, China, July 2018.
- [8] Y. Xu, C. Yang, J. Zhong, N. Wang, and L. Zhao, "Robot teaching by teleoperation based on visual interaction and extreme learning machine," *Neurocomputing*, vol. 275, pp. 2093–2103, 2018.
- [9] M. Trobe and M. D. Burke, "The molecular industrial revolution: automated synthesis of small molecules," *Angewandte Chemie International Edition*, vol. 57, no. 16, pp. 4192–4214, 2018.
- [10] J. Yang, S. Li, Z. Gao, Z. Wang, and W. Liu, "Real-time recognition method for 0.8 cm darning needles and KR22 bearings based on convolution neural networks and data increase," *Applied Sciences*, vol. 8, no. 10, p. 1857, 2018.
- [11] J. Redmon and A. Farhadi, "Yolov3: an incremental improvement," 2018, <http://arxiv.org/abs/1804.02767>.
- [12] R. Girshick, "Fast r-cnn," in *2015 IEEE International Conference on Computer Vision (ICCV)*, pp. 1440–1448, Santiago, Chile, December 2015.
- [13] W. Liu, D. Anguelov, D. Erhan et al., "Ssd: single shot multibox detector," in *Computer Vision – ECCV 2016. ECCV 2016. Lecture Notes in Computer Science, volume 9905*, B. Leibe, J. Matas, N. Sebe, and M. Welling, Eds., pp. 21–37, Springer, Cham, 2016.
- [14] T. Lin, P. Dollár, R. Girshick, K. He, B. Hariharan, and S. Belongie, "Feature pyramid networks for object detection," in *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 2117–2125, Honolulu, HI, USA, July 2017.
- [15] J. Mączak and M. Jasiński, "Model-based detection of local defects in gears," *Archive of Applied Mechanics*, vol. 88, no. 1-2, pp. 215–231, 2018.
- [16] J. M. Gandarias, A. J. García-Cerezo, and J. M. Gómez-De-Gabriel, "CNN-based methods for object recognition with high-resolution tactile sensors," *IEEE Sensors Journal*, vol. 19, no. 16, pp. 6872–6882, 2019.
- [17] S. Lu, G. Qian, Q. He, F. Liu, Y. Liu, and Q. Wang, "Insitu motor fault diagnosis using enhanced convolutional neural network in an embedded system," *IEEE Sensors Journal*, p. 1, 2019.
- [18] Z. Cheng and N. Hu, "Quantitative damage detection for planetary gear sets based on physical models," *Chinese Journal of Mechanical Engineering*, vol. 25, no. 1, pp. 190–196, 2012.
- [19] F. Nabih, C. Jérôme, V. Fabrice, and V. Philippe, "Detection of gear tooth pitting based on transmission error measurements," in *Design and Modeling of Mechanical Systems. Lecture Notes in Mechanical Engineering*, Springer, Berlin, Heidelberg.
- [20] R. Younes, N. Ouelaa, N. Hamzaoui, and A. Djebala, "Experimental study of combined gear and bearing faults by sound perception," in *Advances in Acoustics and Vibration. Applied Condition Monitoring, vol 5*, T. Fakhfakh, F. Chaari, L. Walha, M. Abdennadher, M. Abbes, and M. Haddar, Eds., vol. 76, no. 5-8pp. 927–940, Springer, Cham, 2015.
- [21] M. Zhao, J. Lin, Y. Miao, and X. Xu, "Feature mining and health assessment for gearboxes using run-up/coast-down signals," *Sensors*, vol. 16, no. 11, p. 1837, 2016.
- [22] T. Kidar, M. Thomas, M. Elbadaoui, and R. Guilbault, "Phase monitoring by ESPRIT with sliding window and Hilbert transform for early detection of gear cracks," in *Advances in Condition Monitoring of Machinery in Non-Stationary Operations. Lecture Notes in Mechanical Engineering*, G. Dalpiaz, Ed., pp. 287–299, Springer Berlin Heidelberg, Berlin, Heidelberg, 2014.
- [23] K. Debray, F. Bogard, and Y. Q. Guo, "Numerical vibration analysis on defect detection in revolving machines using two bearing models," *Archive of Applied Mechanics*, vol. 74, no. 1-2, pp. 45–58, 2004.
- [24] R. Moreno, J. M. Chicharro, and P. Pintado, "Comparison of minimum detectable crack size in a geared system from three different vibration transducer types," *Experimental Techniques*, vol. 38, no. 1, pp. 76–87, 2014.
- [25] H. Ozturk, I. Yesilyurt, and M. Sabuncu, "Detection and advancement monitoring of distributed pitting failure in gears," *Journal of Nondestructive Evaluation*, vol. 29, no. 2, pp. 63–73, 2010.
- [26] R. Liu, D. Zhong, H. Lyu, and J. Han, "A bevel gear quality inspection system based on multi-camera vision technology," *Sensors*, vol. 16, no. 9, p. 1364, 2016.
- [27] S. Fedala, D. Rémond, R. Zegadi, and A. Felkaoui, "Gear fault diagnosis based on angular measurements and support vector machines in normal and nonstationary conditions," in *Advances in Condition Monitoring of Machinery in Non-Stationary Operations. CMMNO 2014. Applied Condition Monitoring, vol 4*, F. Chaari, R. Zimroz, W. Bartelmus, and M. Haddar, Eds., pp. 291–308, Cham: Springer International Publishing, 2014.
- [28] A. Djebala, N. Ouelaa, C. Benchaabane, and D. F. Laefer, "Application of the wavelet multi-resolution analysis and Hilbert transform for the prediction of gear tooth defects," *Meccanica*, vol. 47, no. 7, pp. 1601–1612, 2012.
- [29] X. Zhang and S. Fan, *Identification and modification of closed defect of the location function for N-lobed noncircular gears*, Springer Singapore, Singapore, 2016.
- [30] X. Tian, L. Jiao, Y. Duan, and X. Zhang, "Video denoising via spatially adaptive coefficient shrinkage and threshold adjustment in Surfacelet transform domain," *Signal, Image and Video Processing*, vol. 8, no. 5, pp. 901–912, 2014.
- [31] H. D. J. O. Domínguez and V. G. Jiménez, "Evaluation of denoising methods in the spatial domain for medical ultrasound imaging applications," in *Current Trends on Knowledge-Based Systems. Intelligent Systems Reference Library, vol 120*, G. Alor-Hernández and R. Valencia-García, Eds., pp. 263–281, Springer, Cham, 2017.
- [32] J. Yang, S. Li, Z. Wang, and G. Yang, "Real-time tiny part defect detection system in manufacturing using deep learning," *IEEE Access*, vol. 7, pp. 89278–89291, 2019.



Hindawi

Submit your manuscripts at
www.hindawi.com

