

## Research Article

# Bearing Defect Detection with Unsupervised Neural Networks

Jianqiao Xu,<sup>1</sup> Zhaolu Zuo ,<sup>2</sup> Danchao Wu,<sup>3</sup> Bing Li,<sup>3</sup> Xiaoni Li,<sup>4</sup> and Deyi Kong<sup>2,5</sup>

<sup>1</sup>Department of Information Security, Naval University of Engineering, Wuhan 430033, China

<sup>2</sup>Hefei Institutes of Physical Science, Chinese Academy of Sciences, Hefei 230031, China

<sup>3</sup>Hefei Xiaobu Intelligent Technology Co., Ltd., Hefei 230011, China

<sup>4</sup>Shaanxi Aerospace Times Navigation Equipment Co., Ltd., Baoji 721000, China

<sup>5</sup>Innovation Academy for Seed Design, CAS, Beijing 10000, China

Correspondence should be addressed to Zhaolu Zuo; zuozl@iim.ac.cn

Received 16 June 2021; Revised 9 July 2021; Accepted 19 July 2021; Published 20 August 2021

Academic Editor: Jun Zhu

Copyright © 2021 Jianqiao Xu 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.

Bearings always suffer from surface defects, such as scratches, black spots, and pits. Those surface defects have great effects on the quality and service life of bearings. Therefore, the defect detection of the bearing has always been the focus of the bearing quality control. Deep learning has been successfully applied to the objection detection due to its excellent performance. However, it is difficult to realize automatic detection of bearing surface defects based on data-driven-based deep learning due to few samples data of bearing defects on the actual production line. Sample preprocessing algorithm based on normalized sample symmetry of bearing is adopted to greatly increase the number of samples. Two different convolutional neural networks, supervised networks and unsupervised networks, are tested separately for the bearing defect detection. The first experiment adopts the supervised networks, and ResNet neural networks are selected as the supervised networks in this experiment. The experiment result shows that the AUC of the model is 0.8567, which is low for the actual use. Also, the positive and negative samples should be labelled manually. To improve the AUC of the model and the flexibility of the samples labelling, a new unsupervised neural network based on autoencoder networks is proposed. Gradients of the unlabeled data are used as labels, and autoencoder networks are created with U-net to predict the output. In the second experiment, positive samples of the supervised experiment are used as the training set. The experiment of the unsupervised neural networks shows that the AUC of the model is 0.9721. In this experiment, the AUC is higher than the first experiment, but the positive samples must be selected. To overcome this shortage, the dataset of the third experiment is the same as the supervised experiment, where all the positive and negative samples are mixed together, which means that there is no need to label the samples. This experiment shows that the AUC of the model is 0.9623. Although the AUC is slightly lower than that of the second experiment, the AUC is high enough for actual use. The experiment results demonstrate the feasibility and superiority of the proposed unsupervised networks.

## 1. Introduction

With the continuous development and progress of manufacturing industry, the demand for bearings is increasing as a basic component widely used. Performance and life of the machine itself often have a great relationship with the quality of the bearings [1], so the requirements for the quality of the bearings in industrial production continue to increase. In the process of manufacturing and assembly of bearings, defects on the bearing surface are often caused by various reasons. Common defects include pull marks, dark spots, pits, scratches, rust, and yellow spots. These surface

defects will cause the corrosion resistance, elasticity, wear resistance, and lubricity of the bearing to decrease, resulting in a greatly reduced service life of the machine, and even serious safety accidents. Therefore, it is essential to detect the defects of the bearing.

For the detection of bearing surface defects, there are methods such as manual inspection, physical inspection, and machine vision inspection [2]. At this stage, the most important method is manual detection. However, manual inspection is very subjective, and it is often determined by the experience of the inspection operators based on their practice, which is time-consuming and labor-intensive. In

addition, when the operation is performed under continuous light, the inspecting staff are prone to misdetection or missed inspection due to visual fatigue, and it will cause serious harm to the health of the inspector. The common methods of testing in physics are eddy current testing, ultrasonic testing, magnetic particle testing, and so on. These physics-oriented inspection methods are widely used to detect the defects of bearing rollers, but this type of inspection method also has its own shortcomings; that is, it also requires operators to determine the defects of the bearing, but the inspection is not accurate. If the performance is still too low, it will cause missed detection or false detection.

With the continuous development and progress of modern science and technology, when we need to detect defects, machine vision begins to be more and more used. Ye and Hsu designed a new lighting system to collect images in a darkroom, avoiding the influence of external factors and light sources, and developed a rule-based local mask sensor algorithm to achieve high-precision detection of metal defects [3]. Shen et al. designed a new type of lighting and image acquisition system. By taking three photos of the bearing, the left and right photos are used to detect the deformation on the sealing ring, and other defects are detected by the central illumination image to correct the deformation on the sealing ring. Defects have high accuracy and efficiency [4]. Tao proposed a multi-threshold segmentation image based on OSTU to quickly detect defects on the bearing surface. After denoising the collected images, use OSTU to perform threshold segmentation to obtain two thresholds before detecting and locating defects [5].

Traditional surface detection algorithms obtain detected images through image preprocessing and then use statistical machine learning methods to extract image features to achieve the goal of defect detection. These algorithms have achieved good results in some specific applications, but there are still many shortcomings. For example, there are many image preprocessing steps and strong pertinence, with poor robustness; a variety of algorithms have an amazing amount of calculation and cannot accurately detect the size and shape of defects. Deep learning directly updates parameters through learning data, avoids manual design of complex algorithm processes, and has extremely high robustness and accuracy. Zhao et al. [6] proposed a new defect detection framework based on positive sample training, which combines GAN and autoencoder to reconstruct defect image, and LBP is used for image local contrast to detect defects. Wen et al. [7] proposed a multitask convolutional neural network to detect defects. Instead of using a large convolution kernel, a smaller convolution kernel is used to convolve the input data, and the shared neural network is used to classify and locate the defects after extracting the defect features of the sample data. Cha et al. [8] used a sliding window-based convolutional neural classification network to realize the location of crack surface defects, and the combination of two sliding window redundant paths to achieve full image coverage. Wang et al. [9] used a deep convolutional neural network to classify samples of defects

when detecting defects in cloth and then detect defects after classification. Chen et al. [10] use DCNNs combined with SSD, Yolo, and other network methods to build a cascaded detection network from coarse to fine, including firmware positioning, defect detection, and classification. DCNNs have good robustness and adaptability, which means that this method has a good application prospect in the defect detection and classification of fasteners. Mei et al. [11, 12] adopt the idea of image pyramid hierarchy and convolutional denoising autoencoder network to realize defect detection of cloth texture images. The results show that full use of unsupervised learning and multimodal result fusion strategy can improve the robustness and accuracy of defect detection. Bergmann et al. [13] propose an improving unsupervised defect segmentation by applying structural similarity to autoencoders, and the proposed method achieves significant performance gains on a challenging real-world dataset of nanofibrous materials. Yang et al. [14] propose an end-to-end surface quality detection method based on deep convolutional neural networks (CNNs) to improve the accuracy and efficiency of VDR surface quality detection. Essid et al. [15] develop a new machine vision framework for efficient detection and classification of manufacturing defects in metal boxes. The results show that the proposed autoencoder deep neural network (DNN) architecture can not only classify manufacturing defects, but also localize them with high accuracy. Wu et al. [16] propose a high-sensitivity magnetic flux leakage method based on magnetic induction head for the detection of tiny cracks in bearing rings. Xu et al. [17] propose a new multidefect detection method based on a combination of an improved visual attention model and image partitioning-weighted eigenvalue for surface defects of explosive cartridge in the automatic sorting process that are of small area, irregular shape, and random distribution. Kong et al. [18] propose a unified framework for detecting defects in planar industrial products or planar surfaces of nonplanar products based on a template-matching strategy. Tao et al. [19] propose an algorithm for pixel-level segmentation and classification of defects. The entire network can be divided into two stages: defect detection stage and defect classification stage. Fang et al. [20] propose an SLIC head of object instance segmentation in proposal regions (Mask R-CNN) containing a network block to learn the quality of the predict masks. Park et al. [21] propose a convolutional neural network (CNN) based method that inspects nonpatterned welding defects (craters, pores, foreign substances, and fissures) on the surface of the engine transmission using a single RGB camera. Ming et al. [22] propose a combined classifier with dynamic weights (CCDW) to classify the LPG samples considering both feature extraction diversity and base classifiers diversity after image segmentation and enhancement. Martínez et al. [23] propose a machine vision system, performing the detection of flaws on textured surfaces, and multiple images under different lighting conditions are processed and merged into one, which is used to extract features with a supervised classifier. Peng et al. [24] propose a precision measurement and inspection of O-rings with good accuracy and efficiency.

This research is to use the deep neural network to realize the defect detection of the bearing. The main content of this work focuses on the following topics: (1) how to increase the number of samples, (2) how to improve the AUC of the model, and (3) how to enhance the feasibility of the method. The organization of this paper is as follows. Section 2 describes the defect representation and data acquisition system, and Section 3 introduces the methodology. Experiment and results are illustrated in Section 4, and Section 5 gives some discussion. Finally, Section 6 summarizes this paper.

## 2. Defect Representation and Data Acquisition

**2.1. Data Acquisition System.** The data acquisition system is composed of cameras, lighting systems, and computers, as shown in Figure 1. The image capture device can capture images of the inner end surface, outer diameter, inner diameter, and lower end surface separately. Basler industrial camera as A1300-60gm with resolution of  $1282 \times 1026$  pixels is selected, and the lens is PCHI012. Different field of view sizes can be obtained by adjusting the focal length, so as to match the inner diameter, outer diameter, upper end surface, and lower end surface size. By adjusting the exposure time to obtain the largest signal-to-noise ratio, the light source is uniformly illuminated by the ring LED (the light source model is HZN DRL-70-60-W). The final images obtained are shown in Figure 2.

**2.2. Defect Representation.** Bearing defects mainly include the following types: outer diameter defects (stretch marks, dark spots, pits, scratches, rust, and yellow spots); lower end surface defects (dents, convex deformation, scratches, and embroidery); inner diameter defects (dimples, scratches, and embroidery); inner end surface defects (dents, convex deformation, rust, and yellow spots). There are many types of defects, and the characteristics of defects are not obvious, as shown in Figure 3.

## 3. Methodology

Carefully observe the samples obtained by the above-mentioned devices, and you can find that, in addition to useful information, there is some useless redundant information in the samples. In order to ensure the accuracy of detection, a series of pretreatments are required on the samples. Although the defects of the inner end surface, inner diameter, outer diameter, and lower end surface are different, their distributions are similar. They are all distributed along the circumference of the bearing, but the position is different. Therefore, this article selects the inner diameter sample with more complicated appearance and more interference factors. Processing: samples from other parts can be processed in the same way.

**3.1. Normalized Sample Method.** Since the bearing is taken on the liner, in addition to the bearing, images of other parts are also taken. To solve this problem, we first find the contours of the outer and inner edges and perform ellipse

fitting on the contours. Then, based on the center position of the fitted ellipse, move the bearing to the center of the image, and use perspective transformation to transform the ellipse into a circle based on the parameters of the ellipse. Finally, remove all the parts outside the outer edge and inside the inner edge after the transformation. The captured bearing image and the processing algorithm schematic are shown in Figure 4.

**3.2. Sample Split Based on Normalized Sample Symmetry.** After the sample is normalized, the inner diameter part of the bearing is converted into a standard ring, which satisfies the characteristics of stacking based on the center of the image. Since the defect part is generally very small and only occupies a small part of the ring, the symmetry can be used to split the sample into a large number of fan-shaped rings, as shown in Figure 5. The 12 samples obtained will be labelled, and the classifier will be trained based on the divided samples.

**3.3. Supervised Neural Networks Using ResNet Neural Networks.** Deep convolutional neural networks have already shined in image classification problems. Recent studies have also shown that the depth of the network plays a crucial role in accuracy. However, as the network deepens, there is a problem worth noting. As the network continues to stack and deepen, will the effect of the network always get better and better? Obviously, you will encounter the problem of gradient disappearance or gradient explosion, and this problem can already be solved by normalizing the input during initialization, but when the network finally converges, there will be a “degradation” problem, resulting in a decrease in accuracy (not overfitting), so although the number of network layers can be continuously stacked to allow it to train and converge, there is still no way to encounter degradation problems [25].

He et al. [25, 26] build a new network structure (ResNet) to solve the above problem that when the number of network layers is too high, the effect of the deeper network is not as good as the shallower network, and a proper explanation is made. ResNet uses the input of one layer and the output of another layer as the output of a block. Assuming that  $x$  is the input of a block, and one block is composed of two layers, then he first passes through a convolutional layer and activates relu to obtain  $F(x)$ , and then the result of  $F(x)$  after the convolutional layer is added to the previous input  $x$  to obtain a result, and the result is activated by relu as the output of the block. For ordinary convolutional networks, we output  $F(x)$ , but in ResNet, we output  $H(x) = F(x) + x$ , but we still use  $F(x) = H(x) - x$ . This changed the learning goal, changing the original learning to make the objective function equal to a known constant value to make the residual between the output and the input 0, which is the identity mapping. The result is that after the residual is introduced, the output is mapped to the output. The changes are more sensitive.

Based on the samples obtained with Sections 3.1 and 3.2, supervised neural networks can be trained with ResNet neural networks as the following process, as shown in

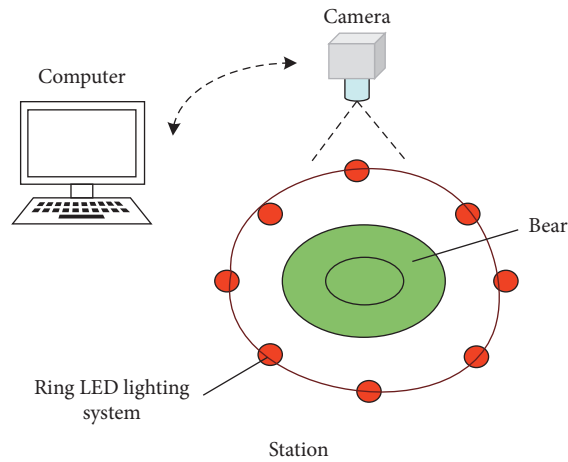


FIGURE 1: Schematic of measurement device.

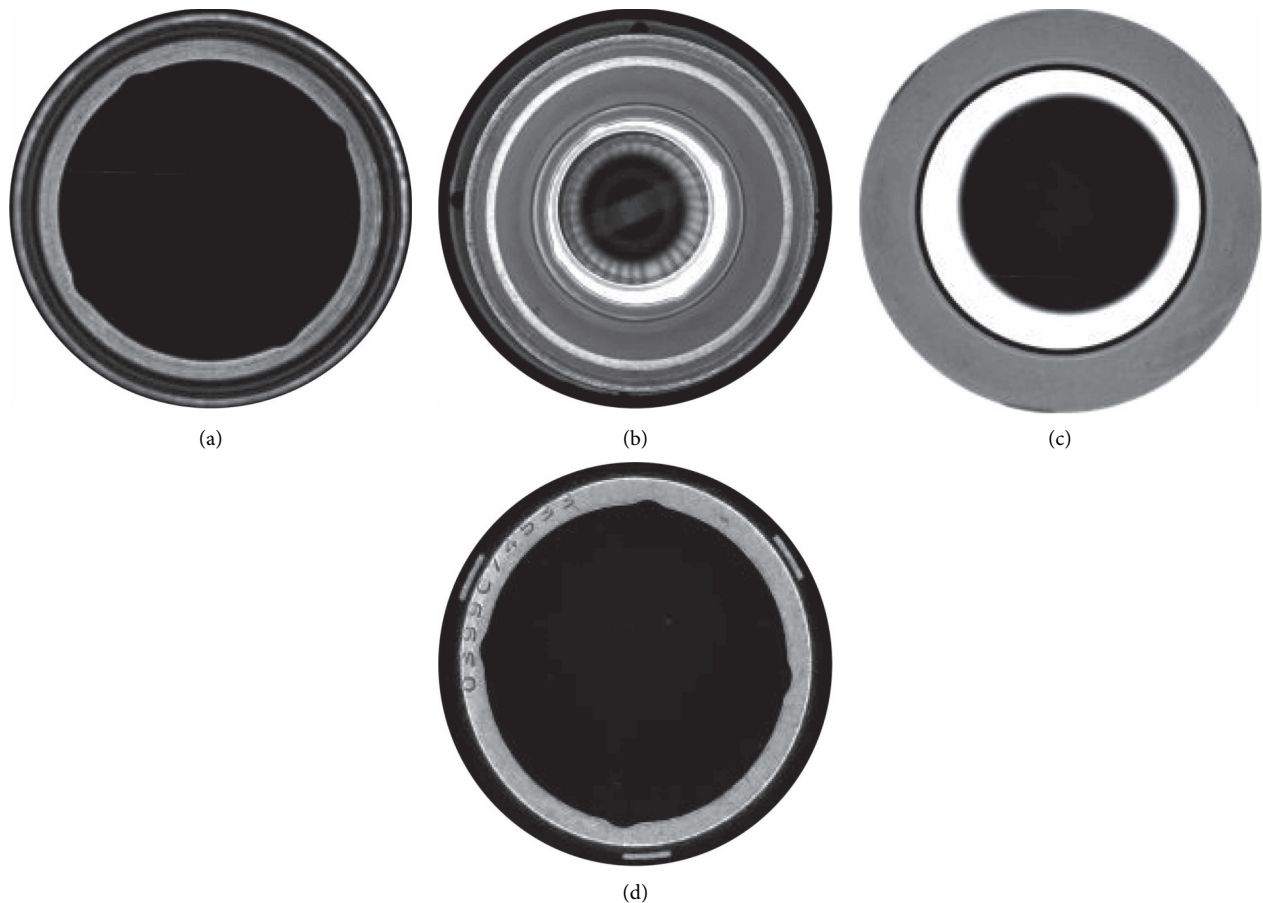


FIGURE 2: Bearing images of different parts. (a) Inner end surface. (b) Outer diameter. (c) Inner diameter. (d) Lower end surface.

Figure 6. Also, more details can be found in our previous work [27].

**3.4. Autoencoder Neural Networks Implemented with U-Net.** In the field of image generation, there is a very important network structure called Autoencoder [28]. An autoencoder

neural network architecture is a feedforward network composed of one or multiple connected hidden layers. It uses a nonlinear mapping function between the original data as input and output specific learned features. The feature of autoencoder is that the first half is the downsampling part, which is generally implemented by CNN; the second half is the upsampling part, which is generally implemented by

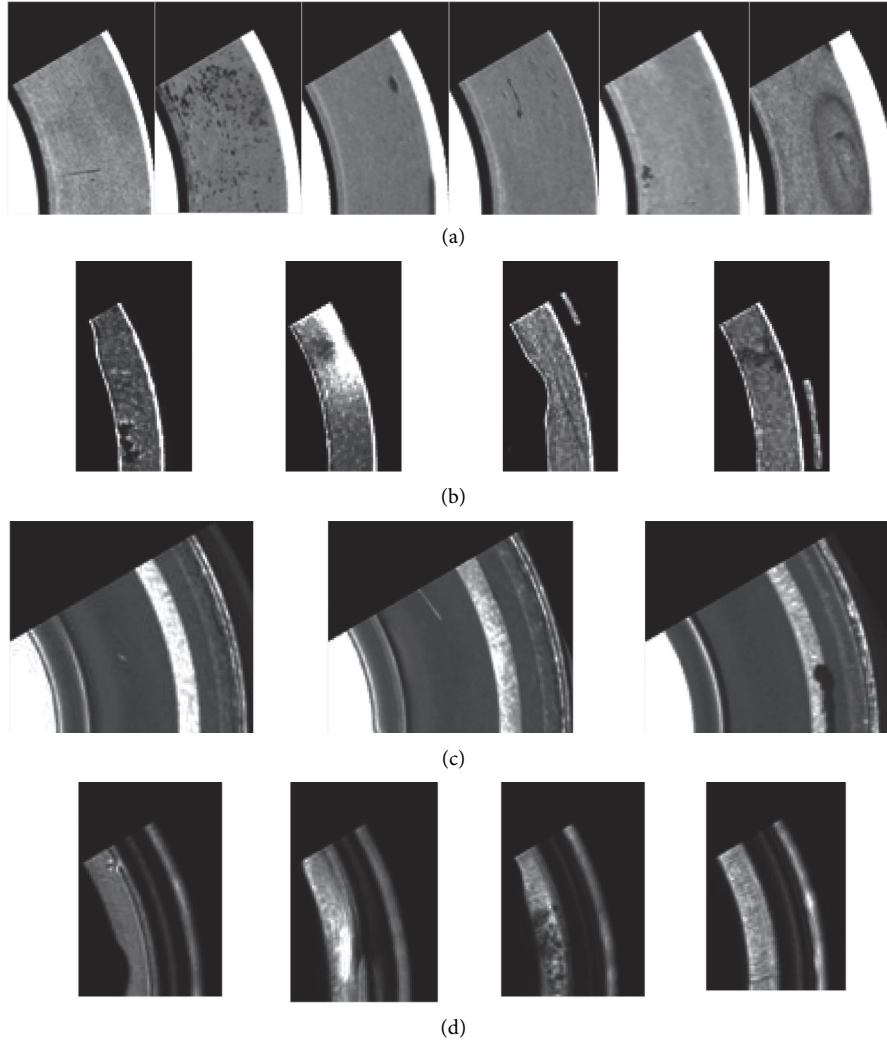


FIGURE 3: Defect classification. (a) Defects of outer diameter. (b) Defects of lower end surface. (c) Defects of inner diameter. (d) Defects of inner end surface.

inverse convolution. The most amazing thing about the entire autoencoder is that even if we only have the features of the middle layer, we can recover a picture that is very close to the original picture through the second half. Therefore, the entire autoencoder has at least two attractive applications: (1) use the first half for feature extraction; (2) use the second half for image generation.

U-Net itself is not used for autoencoder; it first appeared in the segmentation of medical images [29]. On the one hand, its structure is very similar to the traditional structure of autoencoder. On the other hand, its unique feedforward structure allows the network to capture a lot of spatial information. So recently, a lot of image synthesis and generation work are based on U-Net. In this paper, U-Net is used to extract feature map from the original image firstly, and then feature map is used to generate gradient image.

**3.5. The Proposed Unsupervised Neural Network.** Lighting attenuation or batches will affect the classification effect of the supervised network; therefore, an unsupervised neural

network is proposed to solve the disturbing factors, as shown in Figure 7. Based on the samples obtained with Sections 3.1 and 3.2, the proposed unsupervised neural networks can be trained with AE neural networks implemented with U-Net as the following process.

Step 1: raw bearing samples are normalized using Algorithm 1 in Section 3.1

Step 2: normalized samples are split based on normalized sample symmetry using Algorithm 2 in Section 3.2

Step 3: the gradient of the samples is extracted as label data, and Sobel operator is selected to calculate the gradient of the samples

Step 4: AE neural networks implemented with U-Net are used to predict the gradient of the samples

Step 5: the loss function is defined with the argmax of the difference between the label data and the predict data

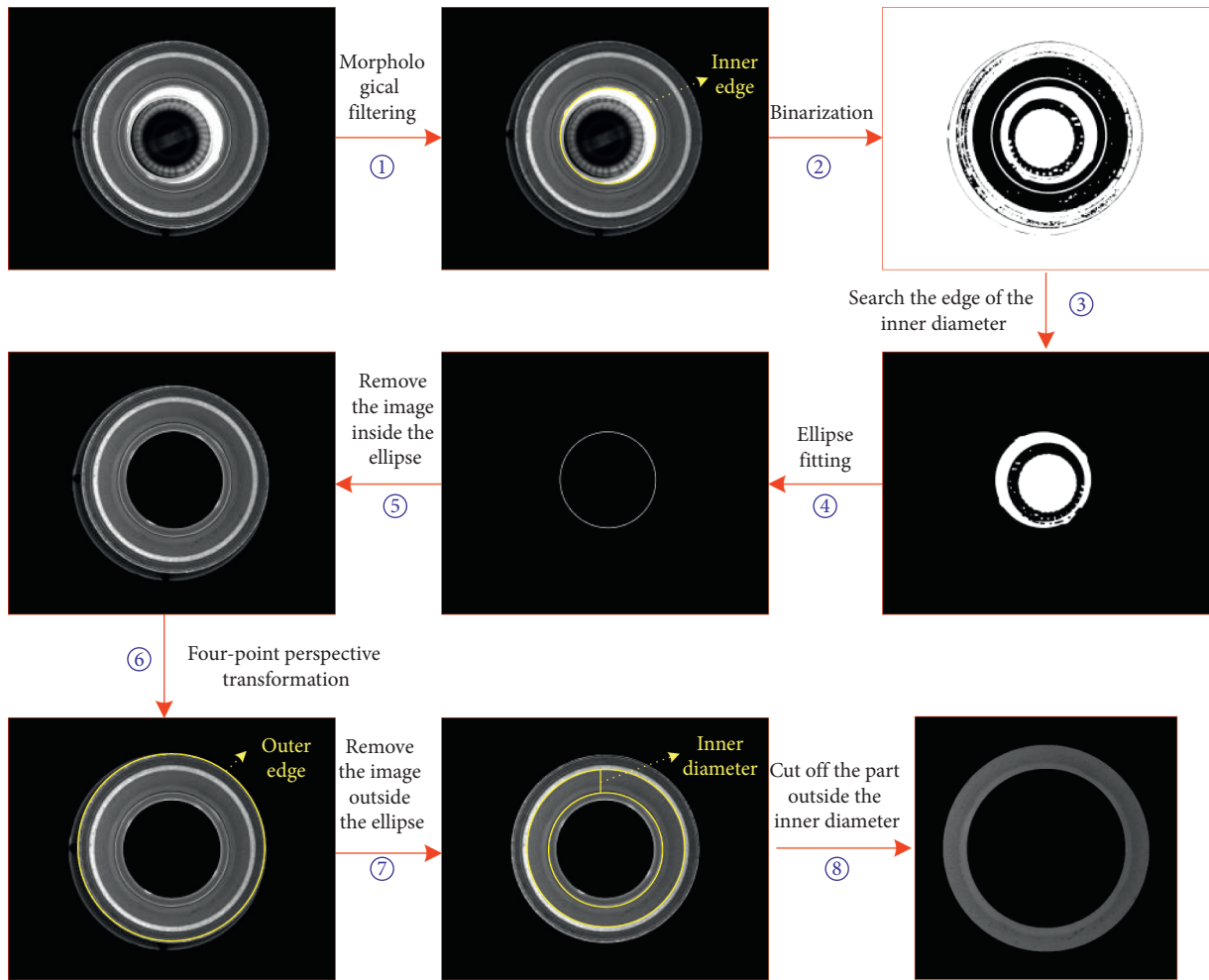


FIGURE 4: Normalization algorithm for bearing inner diameter samples.

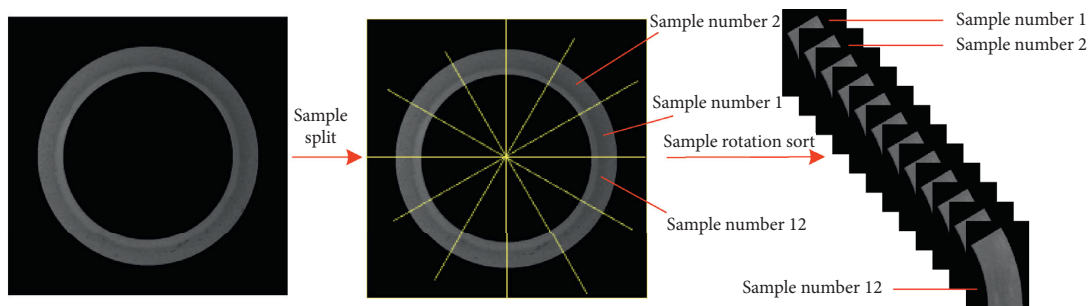


FIGURE 5: Sample splitting algorithm based on symmetry.

Step 6: new data can be updated to online train and online modify the model

## 4. Experiment and Results

The image processing algorithm in this article is trained and tested on the server. The server's processor is Intel(R) Xeon(R) CPU E5-2678v3@2.5 GHz, the graphics card is 2 GeForce GTX 1080 Ti from NVIDIA, and the deep learning architecture uses TensorFlow.

**4.1. Model Training Method.** Different datasets are made for different workstation training, and different defect classifiers are trained through the datasets of different workstations. This article selects the inner diameter sample as an example of the algorithm display. Three experiments are conducted.

The first experiment is the supervised neural networks using ResNet neural networks. The bearing inner diameter samples are divided into training set, validation set, and test set with numbers 16760, 2490, and 2076 separately. The numbers of positive samples and negative samples of the



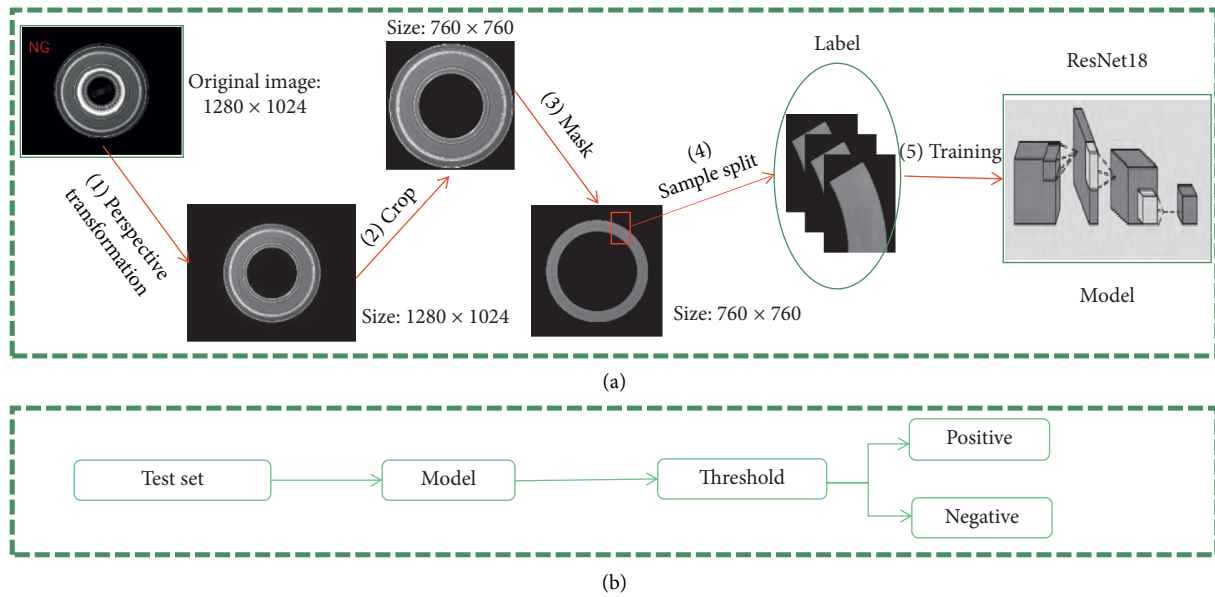


FIGURE 6: Classifier training with supervised neural networks. (a) Train stage. (b) Test stage.

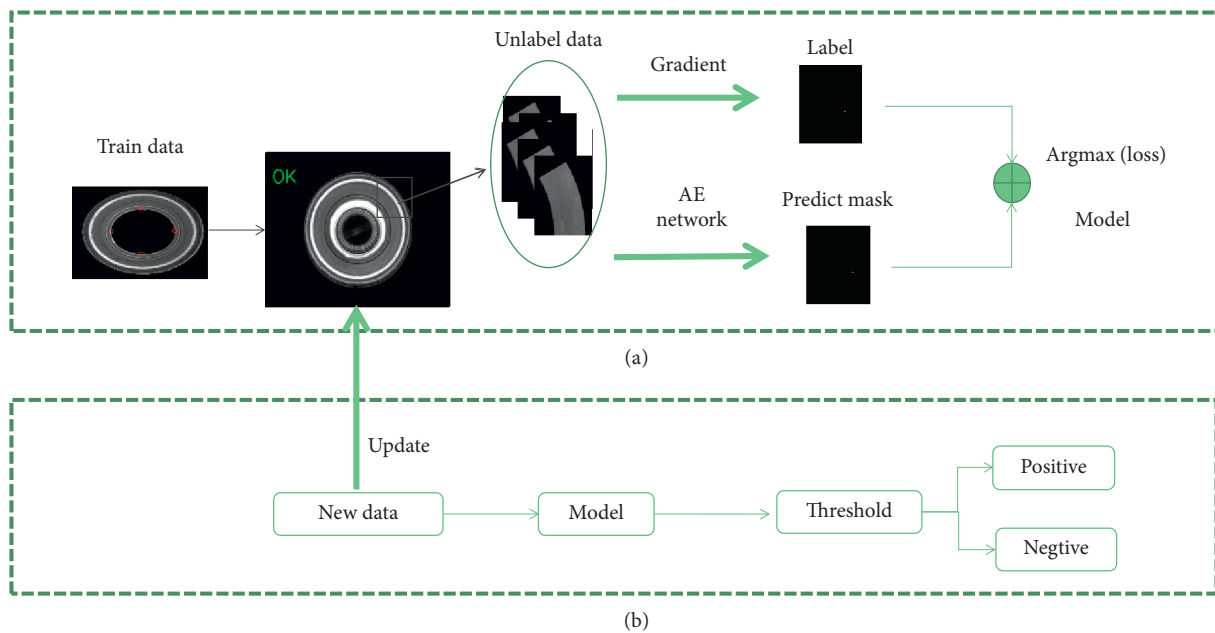


FIGURE 7: Classifier training with the proposed unsupervised neural network. (a) Train stage. (b) Test stage.

training set are 13440 and 3320, respectively. The numbers of positive samples and negative samples of the validation set are 2076 and 414, respectively.

The second experiment is the proposed unsupervised neural networks with AE neural networks. In this experiment, all the negative samples in training set and validation set are discarded. The bearing inner diameter samples are divided into training set, validation set, and test set with numbers 13440, 2076, and 2076 separately. The numbers of training set and validation set in this experiment are less than those of the first experiment. No negative samples are contained in this experiment.

The third experiment is also the proposed unsupervised neural networks with AE neural networks. The difference between this experiment and the second experiment is the training set and validation set. The samples sets are the same with the first experiment, but without any labels. All the positive samples and negative samples in training set and validation set are integrated together, respectively. The numbers of training set and validation set are also as 16760 and 2490 separately.

In order to evaluate the model trained with supervised neural network and the proposed unsupervised neural networks effectively, all the experiments share the same test

**Input:** inner diameter sample with  $1280 \times 1024$  pixels.

**Output:** normalized samples of inner diameter sample with  $760 \times 760$  pixels.

- (1) Morphological denoising: the original image is corroded and expanded, and the template is a  $5 * 5$  rectangular morphological structural element;
- (2) Binarize the original image, take the maximum gray value and minimum gray value of the inner diameter area as the threshold, set the image greater than the maximum threshold and less than the minimum threshold to 255, and the inner diameter area becomes 0;
- (3) Search the inner edge contour, and then fit the inner edge with an ellipse;
- (4) Use the ellipse fitted in step 3 to remove the extra part of the image;
- (5) Search the four points at the top, bottom, left, and right of the inner edge;
- (6) Map the above ellipse to a circle. Take the four points of the top, bottom, left, and right of the circle with the center of the image as the center and the radius of 290 as the target points to establish a projection transformation mapping matrix, and then use this transformation matrix to transform the image in step 4;
- (7) Search for the outer edge contour, fit the ellipse, and cut off the outside of the ellipse;
- (8) Search the area of the inner diameter, and cut off the outer part of the inner diameter area.

ALGORITHM 1: A normalized samples method to obtain the effective part of inner diameter sample from the raw bearing samples.

**Input:** normalized samples of inner diameter sample with  $760 \times 760$  pixels.

**Output:** 12 shares samples along the center of the circle with labels from 1 to 12.

- (1) Divide the inner diameter sample into 12 shares evenly along the center of the circle;
- (2) Label the 12 shares with numbers 1–12
- (3) Rotate the samples 2–12 by a certain angle to the position of sample number 1.

ALGORITHM 2: Sample split based on normalized sample symmetry.

TABLE 1: Classification confusion matrix.

Ground truth	Predictive value	
	Positive	Negative
True	TP	FN
False	FP	TN

set. The numbers of positive samples and negative samples of the test set are 1980 and 96, respectively.

**4.2. Model Evaluation Method.** Generally, the parameters of the classification confusion matrix in the following table are used for statistical calculation. Table 1 shows the classification confusion matrix.

In this paper, the accuracy rate ACC, accuracy rate  $P$ , and recall rate  $R$  of the training model on the black box set are used to evaluate the pros and cons of the model. The accuracy rate ACC is defined as follows: the proportion of the correct result of the classification model to the total observation sample, that is, the proportion of all the predicted results that is correctly predicted. The accuracy rate  $P$  is defined as follows: among the samples that are identified as positive samples, the model predicts the correct proportion. From the perspective of prediction, one type of prediction result is taken out to evaluate the prediction accuracy rate. The recall rate  $R$  is defined as the ratio of correctly identified samples in all positive categories, reflecting the sensitivity of the model.

The accurate rate, accuracy rate, and recall rate are defined as

$$\begin{aligned}
 \text{ACC} &= \frac{\text{TP} + \text{TN}}{\text{TP} + \text{FP} + \text{FN} + \text{TN}}, \\
 P &= \frac{\text{TP}}{\text{TP} + \text{FP}}, \\
 R &= \frac{\text{TP}}{\text{TP} + \text{FN}}.
 \end{aligned} \tag{1}$$

The accuracy rate can better represent the accuracy of the model. Accuracy and recall rate are better performance evaluation indicators than correct rate, which is an evaluation of a certain category. Accuracy and recall are a pair of contradictory measures. Generally speaking, when the accuracy is high, the recall is often low; when the recall is high, the accuracy is often low.

Another more comprehensive evaluation index is receiver operating characteristic (ROC) curve. The ROC curve is used to describe the performance of the two classification systems (the threshold of the classifier is



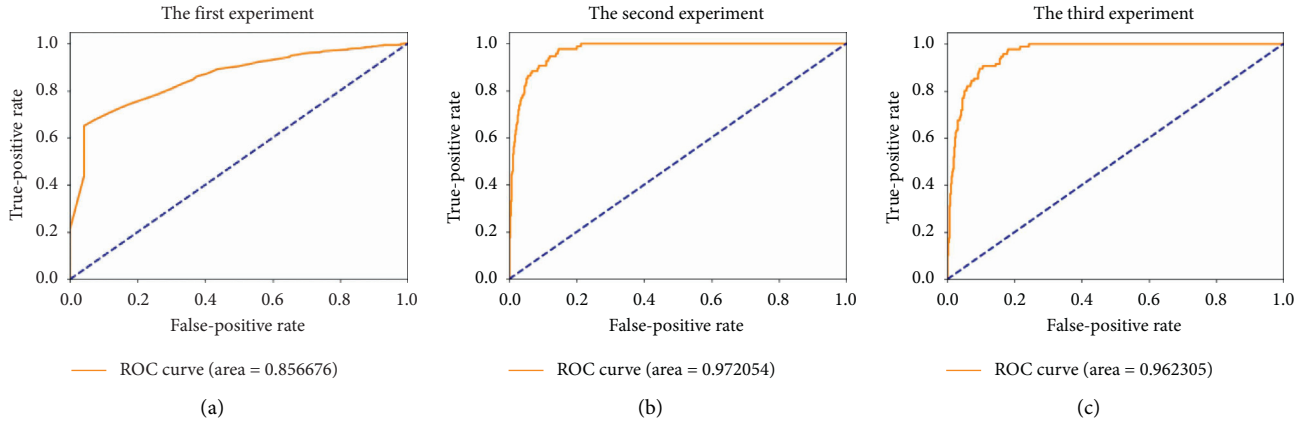


FIGURE 8: ROC curves of different models. The ROC of (a) the supervised neural networks (the first experiment), (b) the unsupervised neural networks (the second experiment), and (c) the unsupervised neural networks (the third experiment).

TABLE 2: Statistics of the above indicators for the supervised neural networks and the unsupervised neural networks.

	TP	FN	FP	TN	ACC	$P$	$R$	AUC
The first experiment	1980	0	96	0	0.9538	0.9538	1	0.8567
The second experiment	1980	0	25	71	0.9879	0.9875	1	0.9721
The third experiment	1980	0	47	49	0.9774	0.9768	1	0.9623

variable), a comprehensive index of continuous changes in response sensitivity and specificity, and the points on the ROC curve reflect the susceptibility of the same signal stimulus. ROC curve and AUC are indicators to evaluate the pros and cons of the two-class model as a whole, where AUC is the area between the ROC curve and its horizontal axis. The ROC curve is generally above  $y = x$ . The larger the AUC value, the better the model. The ROC curve is drawn by two indicators, the true-positive rate (TPR) and the false-positive rate (FPR). The true-positive rate (TPR) is defined as follows: the true label is the proportion of the positive sample, in which the prediction is also the positive sample. The false-positive rate (FPR) is defined as the proportion of positive samples, whose true labels are negative.

$$\begin{aligned} \text{TPR} &= \frac{\text{TP}}{\text{TP} + \text{FP}}, \\ \text{FPR} &= \frac{\text{FP}}{\text{TN} + \text{FP}}. \end{aligned} \quad (2)$$

**4.3. Results.** Train the three experimental models and test them on the same test set, draw the ROC curve, and calculate the AUC, as shown in Figures 8(a)~8(c). It is easy to find that the model of Figure 8(b) has the best performance, while Figure 8(a) has the worst.

Statistics of the above indicators are shown in Table 2. The  $R$  indicator of all the three networks is 100%. From the ACC,  $P$  and AUC indicators, the unsupervised networks have better performance than supervised network. The AUC of the three models is 0.8567, 0.9721, and 0.9623 separately. Though the indicators of the third model are slightly less

than those of the second model, the third model is still good enough for actual use. What is more, the third model is totally an unsupervised model, which is very convenient in actual use and can update the model online.

## 5. Discussion

Some experiments about the supervised neural networks with ResNet networks and unsupervised neural networks with AE networks for bearing defect detection have been carried out in Section 4. According to the results, some points should be discussed further:

- (1) Why does the unsupervised network have better performance than the supervised network? We think the supervised network can have good performance if the defect characteristics are obvious. However, the defects of the bearing are very small and very inconspicuous. The unsupervised networks are good at identifying small defects. Thus, the unsupervised network has better performance.
- (2) Training process: in experiment 2, the unsupervised networks are trained with positive samples, which have the best performance; however, the samples have to be selected manually. In experiment 3, the unsupervised networks are trained with positive samples and negative samples; that is to say, the process of selecting samples is not necessary, which will be of great convenience for industrial site processing.
- (3) Automatic networks update process: the environment of the industrial site may change over time; in this condition, the networks should be updated

automatically for good performance of the networks. The proposed networks can update the networks with the update samples.

## 6. Conclusions

This paper proposes new unsupervised neural networks based on AE networks for bearing defect detection. Sample preprocessing algorithm based on normalized sample symmetry of bearing is adopted to greatly increase the number of samples. Gradients of the unlabeled data are used as labels, and AE networks are created with U-net to predict the output. Three experiments, one with supervised network and the other two with the unsupervised network, are conducted. The AUC of the three models is 0.8567, 0.9721, and 0.9623 separately. Though the indicators of the third model are slightly less than those of the second model, the third model is still good enough for actual use. What is more, the third model is totally an unsupervised model, which is very convenient in actual use and can update the model online. The experiment results demonstrate the feasibility and superiority of the proposed unsupervised networks. It can be expected that, with the widespread application of visual inspection systems in bearing automation production lines, the proposed method can greatly improve production efficiency and make a certain contribution to the improvement of bearing production quality.

## Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

## Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this study.

## References

- [1] J. Xingxing, S. Changqing, S. Juanjuan, and Z. Zhu, "Initial center frequency-guided VMD for fault diagnosis of rotating machines," *Journal of Sound and Vibration*, vol. 435, no. 24, pp. 36–55, 2018.
- [2] K. Wang, *Application of Image Fusion in Detection of Micro Bearing Surface Defect*, Jilin University, Changchun, China, 2011.
- [3] J.-H. Ye and Q.-C. Hsu, "Automatic optical apparatus for inspecting bearing assembly defects," *Sensors and Materials*, vol. 30, no. 11, pp. 2637–2652, 2018.
- [4] H. Shen, S. Li, D. Gu, and H. Chang, "Bearing defect inspection based on machine vision," *Measurement*, vol. 45, no. 4, pp. 719–733, 2012.
- [5] Q. Tao and X. Wu, "Rapid detection method of the bearing surface defects," *Microelectronics & Computer*, vol. 28, no. 10, pp. 98–100, 2011.
- [6] Z. Zhao, L. Bo, D. Rong, and P. Zhao, "A surface defect detection method based on positive samples," in *Proceedings of the Pacific Rim International Conference on Artificial Intelligence*, Springer, Nanjing, China, July 2018.
- [7] S. Wen, Z. Chen, and C. Li, "Vision-based surface inspection system for bearing rollers using convolutional neural networks," *Applied Sciences*, vol. 8, no. 12, Article ID 2565, 2018.
- [8] Y. J. Cha, W. Choi, and O. Büyüköztürk, "Deep learning-based crack damage detection using convolutional neural networks," *Computer-Aided Civil and Infrastructure Engineering*, vol. 32, 2017.
- [9] T. Wang, Y. Chen, M. Qiao, and H. Snoussi, "A fast and robust convolutional neural network-based defect detection model in product quality control," *International Journal of Advanced Manufacturing Technology*, vol. 94, 2017.
- [10] J. Chen, Z. Liu, H. Wang, A. Nunez, and Z. Han, "Automatic defect detection of fasteners on the catenary support device using deep convolutional neural network," *IEEE Transactions on Instrumentation and Measurement*, vol. 67, no. 2, pp. 257–269, 2018.
- [11] S. Mei, H. Yang, and Z. Yin, "An unsupervised-learning-based approach for automated defect inspection on textured surfaces," *IEEE Transactions on Instrumentation and Measurement*, vol. 67, no. 6, pp. 1266–1277, 2018.
- [12] S. Mei, Y. Wang, and G. Wen, "Automatic fabric defect detection with a multi-scale convolutional denoising autoencoder network model," *Sensors*, vol. 18, no. 4, Article ID 1064, 2018.
- [13] P. Bergmann, S. Lwe, M. Fa User et al., "Improving unsupervised defect segmentation by applying structural similarity to autoencoders," in *Proceedings of the 14th International Conference on Computer Vision Theory and Applications*, January 2018.
- [14] T. Yang, S. Peng, and L. Huang, "Surface defect detection of voltage-dependent resistors using convolutional neural networks," *Multimedia Tools and Applications*, vol. 79, no. 9–10, pp. 1–16, 2020.
- [15] O. Essid, H. Laga, and C. Samir, "Automatic detection and classification of manufacturing defects in metal boxes using deep neural networks," *PLoS One*, vol. 13, 2018.
- [16] J. Wu, Y. Yang, E. Li et al., "A high-sensitivity MFL method for tiny cracks in bearing rings," *IEEE Transactions on Magnetics*, vol. 54, 2018.
- [17] L. Xu, H. Xu, X. Li et al., "A defect inspection for explosive cartridge using an improved visual attention and image weighted eigenvalue," *IEEE Transactions on Instrumentation and Measurement*, vol. 69, no. 99, p. 1, 2019.
- [18] H. Kong, J. Yang, and Z. Chen, "Accurate and efficient inspection of speckle and scratch defects on surfaces of planar products," *IEEE Transactions on Industrial Informatics*, vol. 13, no. 4, pp. 1855–1865, 2017.
- [19] X. Tao, D. Zhang, W. Ma, X. Liu, and D. Xu, "Automatic metallic surface defect detection and recognition with convolutional neural networks," *Applied Sciences*, vol. 8, no. 9, Article ID 1575, 2018.
- [20] X. Fang, W. Jie, and T. Feng, "An industrial micro-defect diagnosis system via intelligent segmentation region," *Sensors*, vol. 19, no. 11, Article ID 2636, 2019.
- [21] J.-K. Park, W.-H. An, and D.-J. Kang, "Convolutional neural network based surface inspection system for non-patterned welding defects," *International Journal of Precision Engineering and Manufacturing*, vol. 20, no. 3, pp. 363–374, 2019.
- [22] W. Ming, F. Shen, H. Zhang et al., "Defect detection of LGP based on combined classifier with dynamic weights," *Measurement*, vol. 143, pp. 211–225, 2019.
- [23] S. S. Martínez, C. O. Vázquez, J. G. García, and J. G. Ortega, "Quality inspection of machined metal parts using an image fusions technique," *Measurement*, vol. 111, pp. 374–383, 2017.

- [24] G. Peng, Z. Zhang, and W. Li, "Computer vision algorithm for measurement and inspection of O-rings," *Measurement*, vol. 94, pp. 828–836, 2016.
- [25] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 770–778, Las Vegas, NV, USA, June 2016.
- [26] K. He, X. Zhang, S. Ren, and J. Sun, "Identity mappings in deep residual networks," in *Proceedings of the European Conference on Computer Vision*, pp. 630–645, Amsterdam, The Netherlands, September 2016.
- [27] J. Xu, J. Wu, X. Chen, D. Wu, and B. Li, "Bearing defects detection based on standardized sample split," *Journal of Applied Optics*, vol. 42, no. 2, pp. 327–333, 2021.
- [28] C. Shen, Y. Qi, J. Wang, G. Cai, and Z. Zhu, "An automatic and robust features learning method for rotating machinery fault diagnosis based on contractive autoencoder," *Engineering Applications of Artificial Intelligence*, vol. 76, pp. 170–184, 2018.
- [29] O. Ronneberger, P. Fischer, and T. Brox, "U-net: convolutional networks for biomedical image segmentation," in *Proceedings of the International Conference on Medical Image Computing and Computer-Assisted Intervention*, pp. 234–241, Springer, Munich, Germany, November 2015.