Research Article

Size Detection Method of Embedded Ball Based on Aviation Bearing Image

Ping Xue (iD), **Yihui Wang** (iD)**, and Hongmin Wang**

School of Automation, Harbin University of Science and Technology, Harbin 150080, China

Correspondence should be addressed to Ping Xue; xueping@hrbust.edu.cn

In the aerospace industry, bearing is widely used in various rotating machinery. The performance of bearing affects the operation of the whole machinery and even aviation equipment. The wrongly assembled ball due to size is an important reason for unqualified bearing. To solve this problem, an accurate ball detection method based on the bearing image is proposed. Firstly, according to the imaging characteristics of bearing and light propagation characteristics, an image collection system based on the coaxial light source is designed. Then, aiming at the problem that the embedded ball is occluded by the bearing ring and the cage, only partial ball in the narrow gap can be used to predict the full ball and the high-precision requirement of ball detection, a ball segmentation model based on DeepLab v3+ network is used to segment the local ball, and CBAM is added in the Xception network of the original network. According to the characteristics of the segmentation result, a circle detection algorithm based on circle fitting evaluation designed for incomplete short arc is proposed. Finally, according to the detection results, judge whether the bearing is qualified or not and evaluate the feasibility of this method. Experimental results show that the ball detection accuracy is about 27 microns, and the wrongly assembled ball with a size difference of only 198 microns can be distinguished. The false detection rate of unqualified bearing is 1%. As the last line of defense of bearing quality inspection, this method can achieve zero false detection rate of unqualified bearing in the industry.

1. Introduction

As one of the most commonly used components of various rotating machinery, high-precision aviation bearing is widely used in aviation equipment, such as steam turbines, generators, and gas turbines. Aviation bearing has very high requirements on precision, and even if there is a small bearing quality problem, a chain reaction may occur, affecting the operation of the entire aviation equipment. Therefore, it is necessary to ensure that its quality is 100% qualified. The assembly of the bearing ball will affect the load-bearing performance of bearings. If one or more balls are unqualified, it will have a great impact on the performance of bearings, and the factor that most easily causes the unqualified bearing ball is the size error of the ball. Although the size of the ball is measured strictly before assembly, there are a large number of bearings in the assembly line, and there is only a micron difference in the sizes between balls of different specifications, which may lead to the wrong assembly of the ball bearing. So it is necessary to carry out the last strict inspection on the size of the aviation bearing ball before delivery. At present, some research studies on bearing assembly inspection use signal processing to judge the performance of bearing when it is used [1–3]. However, for the new bearing that has not been delivered from the factory, the trial one by one will make it wear. Therefore, some manufacturers will manually carry out the last assembly inspection, but the embedded ball is occluded by the bearing ring and cage, the human eye can only observe the micron-level part of the gap, so the efficiency and accuracy of manual detection cannot be guaranteed. For the question, a nondestructive, high-efficiency, and high-precision detection method is urgently needed. Based on the development of machine vision detection technology and its successful application in the accurate detection of various parts [4, 5], this paper intends to adopt machine vision technology to solve the problem.

Some researchers use machine vision to detect the balls of bearing, but the detection targets are mostly slightly

blocked balls with visible complete outline; these methods are not applicable to the embedded balls in bearing which are seriously shielded. After the aviation bearing assembly is completed, the embedded ball is occluded by the cage and inner and outer ring, only the local image of the ball bearing can be obtained from the micron-level gap, and then the complete bearing shape can be fitted. Therefore, the automatic locating and segmentation of the bearing ball are the premise of its recognition and size detection. Traditional segmentation methods, such as threshold method and region method [6], are based on low-level semantic segmentation of images, and they are not effective in target segmentation of complex scenes. How to use the basic information in the image and combine the high-level semantics to improve the segmentation effect is a research hotspot in recent years. Convolutional neural network has made great progress in the field of image segmentation [7]. By constructing loss functions and training the network, the model has the ability to extract image features efficiently and to automatically segment the target of interest. And then, various semantic segmentation algorithms based on the convolutional neural network have been proposed [8–10], among which Chen proposed a relatively mature model—DeepLab [11]. Then, a series of DeepLab semantic segmentation improvement models have been proposed, and in 2018, DeepLab v3+ was proposed [12–14]. This model adds a decoder based on DeepLab v3. It not only uses dilated convolution to increase the receptive field without increasing the size of the convolution kernel to obtain more information and uses the pyramid structure of dilated convolution to obtain multiscale information to integrate the global context information, but also uses decoder to restore the details of the image step by step and obtain finer object boundaries. In addition, the fully connected conditional random fields are creatively connected to the last layer of the convolutional neural network to the structural prediction, thus enhancing the locating ability of objects and improving the accuracy of semantic segmentation. Since then, some new algorithms have also provided ideas for the improvement of DeepLab v3+. In [15], Yang embeds the attention mechanism module in the encoder of the DeepLab v3+ model, and the high-level features of the image are extracted by dense depth-wise separable convolution. The surgical instruments are segmented by the improved network, thus realizing the accurate segmentation of surgical instruments. In [16], Wang designs a power line detection method combining Haar-like features and DeepLab v3+ network. On the basis of DeepLab v3+ network, a more complex decoder structure is designed, and this network is applied to the segmentation of power line, which overcomes the problems of complex background and small pixel width of power lines and realizes the accurate segmentation of power lines. Based on the high segmentation accuracy in the above segmentation experiment, a ball segmentation model based on DeepLab v3+ network is used to locate and segment the ball bearings automatically.

In view of the fact that the previous detection methods are not suitable for the bearing ball detection in this paper, this paper proposes a detection method for the ball target which is mostly occluded, to pick up the unqualified bearings. The innovations of this paper are as follows: Firstly, to solve the problem that the traditional segmentation method cannot accurately locate the partially occluded ball target, the semantic segmentation method based on the convolutional neural network is creatively used to extract the ball feature. Secondly, according to the characteristics of small-scale ball target and edge thinning, CBAM is added in the residual branch of the residual block in the Xception network of the original network to make it more suitable for the characteristics of ball segmentation. Finally, according to the short edge of the segmented ball, a circle detection algorithm based on circle fitting degree is proposed, which is suitable for fitting a small number of pixels. After multiple sampling, the fitting circle can be obtained by looking for the optimal value.

2. Analysis of Imaging Characteristics and Design of Lighting Source

2.1. Problem Description. Figure 1 shows the bearing image, which can be used to analyze the imaging characteristics and image characteristics of the embedded ball. The ball is occluded by the bearing ring and cage and dynamically rotatable, and the local information of the bearing ball can only be obtained through the micron-level dynamic gap. And the surface of the bearing is very smooth; the reflection problem is serious. It is easy for the image details to be unclear due to reflection. What is more, the bearing and the ball surface are covered with mechanical oil, easy to stick burr, and cutting residue and cause boundary blurring; all these make it difficult to collect high-quality images. However, the size difference between balls of different specifications which may be wrongly assembled is only 198 um, we need to control the detection error within 99 microns, and the premise of accurate measurement is the clear outline and details of bearings. Therefore, how to overcome these problems and get high-quality images of bearing balls is the problem that should be solved, and designing a suitable lighting source is the key to solve this problem.

2.2. Design of Lighting Source. The bearing surface is very smooth and has serious light reflection, and it is close to specular reflection, which may lead to halo phenomenon, which is a manifestation of energy concentration. As the light emitted by the light source directly enters the CCD after mirror reflection, the photosensitive element in the CCD camera receives excessive photons, resulting in halo phenomenon, which adds a lot of interference information to the bearing image and brings great difficulties to subsequent processing. Therefore, the selection of light source plays an important role in the quality of the bearing ball image.

Coaxial lighting is composed of high-density LED lamps, which have the advantages of high brightness, good heat dissipation, and high stability. Figure 2 shows the principle of the coaxial light source. The principle is that the light source hits the spectroscope (semipermeable membrane) through the diffuse reflection plate, the spectroscope can
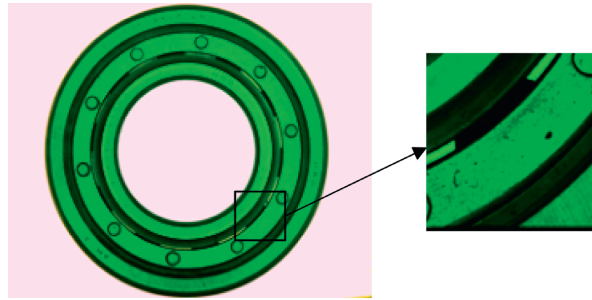
FIGURE 1: The bearing image. Complete aviation bearing and partially occluded ball image.
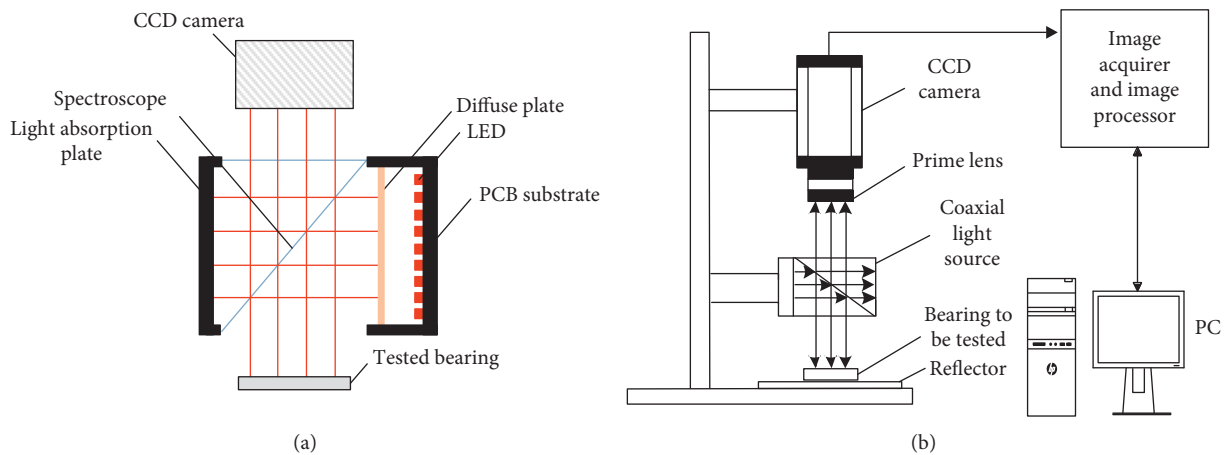

(a)

(b)

FIGURE 2: Lighting source and its application in the image acquisition system. (a) Schematic diagram of the coaxial light source. (b) Structure diagram of the image acquisition system.

transmit and reflect two parts of light respectively, and the reflected light can irradiate the tested bearing vertically. Because the object and the camera are on the same axis, the light reflected from the surface of the measured object will be collected vertically by the CCD camera. Coaxial light source has clear imaging and uniform brightness because its parallel light source irradiates the object without stray light, which can highlight the surface of the object and strengthen the defect features.

What is more, the backlight light source makes the light passing through the gap stronger, which can better reflect the details of the balls in the gap. Therefore, the coaxial light source is combined with the backlight light source, and a reflector is arranged below the bearing, so that the light emitted by the coaxial light source can be reflected by the reflector and can penetrate the gap more, so the details of the balls in the gap are clearer.

3. Partially Occluded Ball Segmentation Model Based on DeepLab v3+

3.1. DeepLab v3+ Network. DeepLab v3+ is a new typical network for semantic segmentation developed by Google in 2018. Its structure is divided into encoder and decoder, that is, a decoder structure is added on the basis of DeepLab v3. The encoder extracts features with Xception and ASPP modules. Xception is a DCNN feature extraction network

with entry flow, middle flow, and exit flow [17]. ASPP is the atrous spatial pyramid pooling structure, which is a feature extraction module with multiple dilated convolutions with different rates [18]. The decoder is used to restore the details of the image step by step and obtain finer object boundaries. In this model, the dilated convolution is used instead of the traditional convolution method to expand the receptive field and obtain more context information. What is more, depth separable convolution replaces max pooling, and depth separable convolution divides the general form of convolution into a depth convolution and a convolution with a convolution kernel size of $1*1$. Depth convolution convolutes different channels of input data, and $1*1$ convolution integrates the results of depth convolution. The depthwise separable convolution is used so that the amount of learning parameters is reduced and the running speed of the network has also increased. As shown in Figure 3, it is proved that it is feasible to use DeepLab v3+ network to segment ball features.

3.2. Improved DeepLab v3+ Network. Although ball features are segmented by the original DeepLab v3+ network, the segmentation results show that the distinction between ball and background is not clear, and the ball boundary is rough. One of the reasons is that the gray distribution of the ball boundary area is uneven, which makes segmentation
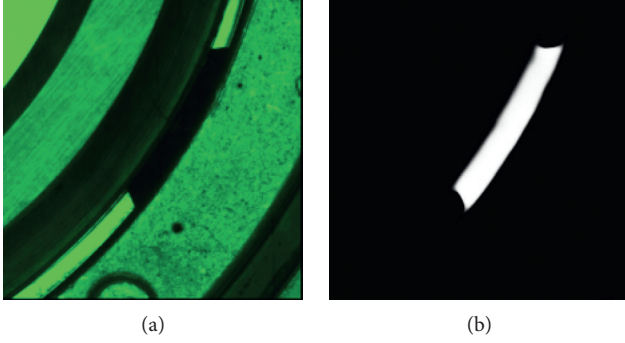
FIGURE 3: Segmentation result of DeepLab v3+ network. (a) The original ball image. (b) The segmented local ball.

difficult. The other is that because of the lack of attention mechanism, the model is not effective in segmenting ball target with small size and inconspicuous features. Therefore, this paper adds attention mechanism on the basis of the original network, combines the decoding method of the full convolution network, connects the information of the encoder to the decoder by jumping, and uses UP for upsampling to restore the edge details layer by layer, which makes it more suitable for the characteristics of small pixel width occupied by local balls.

3.2.1. CBAM. CBAM (Convolutional Block Attention Module) is the attention mechanism module of convolution module. Attention is another important quota which is different from the depth, width, and cardinality of convolutional neural networks. It points out which spatial positions and channels are more important in the multichannel feature map output by a convolution layer [19]. CBAM proposes to combine spatial dimension information on the basis of channel, considering both channel feature correlation and spatial pixel feature correlation. The specific structure of CBAM is shown in Figure 4. CBAM module is divided into two parts: channel attention module and spatial attention module.

The channel attention mechanism module mainly focuses on useful features and uses parallel global maximum pooling and global average pooling to compress the feature map in spatial dimension. The 2 outputs are calculated by the shared network composed of MLP (multilayer perceptron), and then the channel attention map M_c is finally obtained by sigmoid function. The calculation process of M_c is as follows:

$$M_c(F) = \sigma(\text{MLP}(\text{Avg Pool}(F)) + \text{MLP}(\text{Max Pool}(F))), \tag{1}$$

where F is the input image and σ represents the sigmoid operation.

Spatial attention module mainly focuses on pixel location information. The obtained channel feature map is used as the input feature map of this module, after the global maximum pooling and average pooling of channel dimensions, the two feature maps are merged by concatenation and reduced to one channel by a convolution operation, and finally, the spatial

attention map M_s is obtained by sigmoid function. The calculation process is as follows:

$$M_s(F) = \sigma(f([\text{Avg Pool}(F); \text{Max Pool}(F)])), \tag{2}$$

where f is the convolution layer of 3^*3.

3.2.2. Network Structure. Figure 5 shows the improved network structure, adding CBAM in the residual branch of the residual block in Xception of the encoder and behind the convolutional layer in the 3^*3 convolution of the decoder module. Adding CBAM has little effect on the original structure of the network and hardly increases the training cost but enables the network to learn more important features and spatial positions.

In the encoder part, the feature map from exit flow in Xception was input into the ASPP. In this paper, the kernel sizes of the four dilated convolutions of ASPP are 1^*1, 3^*3, 3^*3, and 3^*3, respectively, and the dilated rates of the three dilated convolutions are 6, 12, and 18, respectively. Then, the channel number of the feature map which is the output of the ASPP was transformed by the convolutional layer of 1^*1. It should be noted that the convolution operation of 1^*1 is required for each extracted feature map in order to adjust the number of channels. It is prevented that the proportion of channels originally belonging to the low-level features in the stacked feature map is too large, resulting in the prediction result inclining to the low-level features. This is different from the general encoder-decoder which directly uses skip structure to combine the low-level features with the high-level semantic.

In the decoder part, first take out the lower-layer feature map in the Xception network, then transform its channels with 1^*1 convolution operation, finally, it is unified with the feature map output by the encoder, and the two are merged. After merging the features, the features are reextracted through a convolution layer of 3^*3, where 3^*3 convolution is to learn and increase nonlinearity, so as to effectively combine low-level features with high-level semantic. Then, the feature map with the same size as the input map is obtained through bilinear interpolation upsampling. The upsampling operation is to increase the size of the current feature map so as to be combined with the next extracted low-level feature map with a larger size.

4. Ball Recognition and Measurement Based on DeepLab v3+ and Circle Detection

4.1. Circle Detection Based on Multiple Random Circle Detection and Fitting Degree Evaluation. Aiming at the incomplete and short ball edges after segmentation, and in order to make the locating method have strong anti-interference ability and fast execution speed, a circle detection method based on random circle detection and random circle sampling consistency fitting is proposed, including three-point random circle detection, circle fitting degree evaluation, and determination of fitting circle [20, 21]. Compared with traditional circle detection methods such as least square method and Hough transform, this method is especially suitable for incomplete circle fitting [22, 23]. The idea of three-point random circle detection is to randomly sample
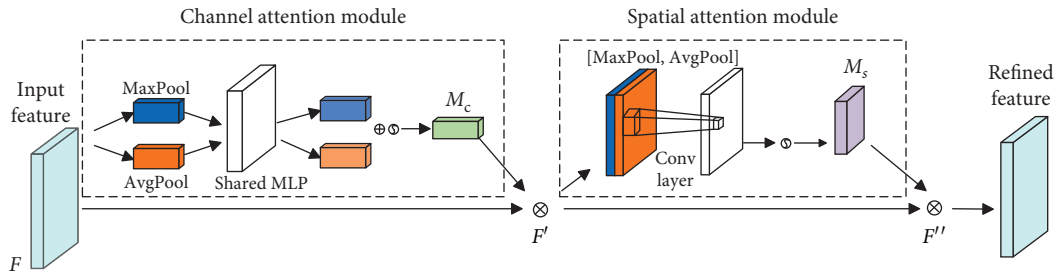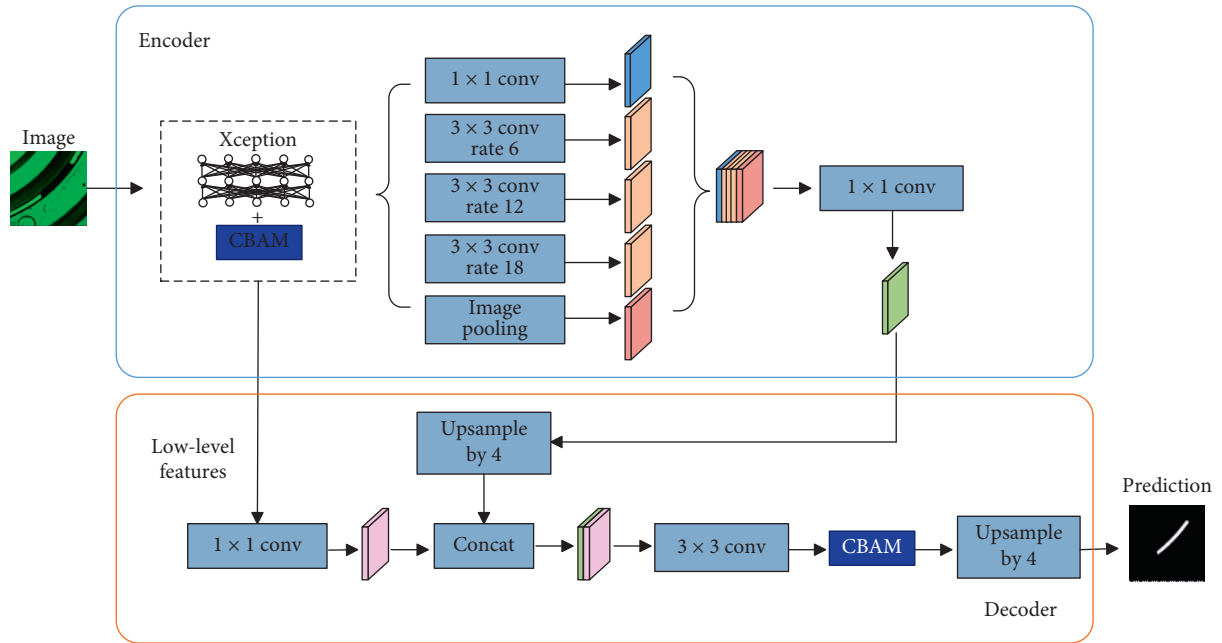
Figure 4: CBAM module structure.

Figure 5: Improved DeepLab v3+ network structure.

three points from the segmented ball edge points for circle detection. Randomly sampling three points from N edge points, assume that the coordinates of the three points are $(x_{i1}, y_{i1}), (x_{i2}, y_{i2})$, and (x_{i3}, y_{i3}), respectively, where i is the number of samples. According to the three points, a random

circle is determined, the center coordinate of the circle is (C_X_i, C_Y_i), and the radius is C_R_i. Save them in a two-dimensional array $C_{XYR}(L, 3)$ with L rows and three columns, where L is the total number of samples. The calculation method of circle parameters is as follows:

$$C\_X_i = \frac{\begin{vmatrix} x_{i2}^2 + y_{i2}^2 - x_{i1}^2 - y_{i1}^2 & 2(y_{i2} - y_{i1}) \\ x_{i3}^2 + y_{i3}^2 - x_{i1}^2 - y_{i1}^2 & 2(y_{i3} - y_{i1}) \end{vmatrix}}{4\left[(x_{i2} - x_{i1})(y_{i3} - y_{i1}) - (x_{i3} - x_{i1})(y_{i2} - y_{i1})\right]},$$

$$C\_Y_i = \frac{\begin{vmatrix} 2(x_{i2} - x_{i1}) & x_{i2}^2 + y_{i2}^2 - x_{i1}^2 - y_{i1}^2 \\ 2(x_{i3} - x_{i1}) & x_{i3}^2 + y_{i3}^2 - x_{i1}^2 - y_{i1}^2 \end{vmatrix}}{4\left[(x_{i2} - x_{i1})(y_{i3} - y_{i1}) - (x_{i3} - x_{i1})(y_{i2} - y_{i1})\right]},$$

$$C\_R_i = \sqrt{(x_{i1} - C\_X_i)^2 + (y_{i1} - C\_Y_i)^2}.$$

$$(3)$$

Then, the circle fitting degree is evaluated for each random circle, and the random circle satisfying the requirements is selected as the candidate circle by taking the circle fitting degree as the evaluation standard. In order to evaluate the fitting degree, the edge points (including randomly selected three points) whose distance to the random circle is less than the threshold T_d are defined as interior points; otherwise, they are exterior points. Defining the ratio of the number of interior points to the total number of edge points as a circle fitting degree and determining a fitting circle according to the circle fitting degree comprise the following steps.

Calculate the distance from each edge point to the random circle according to the following formula:

$$d_{ij} = \sqrt{(X_j - C\_X_i)^2 + (Y_j - C\_Y_i)^2} - C\_R_i, \qquad (4)$$

(X_j, Y_j) is the coordinate of the edge point $(0 < j < N)$ and d_{ij} is the distance from the jth edge point to the random circle in the ith sampling.

For each sampled random circle, count the number n of interior points in N edge points, and the characteristic of interior points is $d_{ij} < T_d$. The circle fitting degree $gf = (n/N)$ is calculated and stored in row i of an array $C_{gf}(L)$ with L rows and one column, respectively.

After L times of random circle detection, the circle fitting degree corresponding to the obtained L random circles is evaluated, from which the maximum circle fitting degree $gf_{\max}$ is selected, and its corresponding random circle is taken as the final fitting circle to obtain its circle parameters.

4.2. Prediction and Fitting of Complete Ball. The aviation bearing ball segmented by the improved DeepLab v3+ network has higher accuracy, but after a step of edge detection, the accuracy of its edge is reduced again. In this paper, the process as shown in Figure 6 is adopted. Firstly, the edge detection is used to detect the edge of the bearing image, then the ball segmentation network is used to segment the bearing ball, and edge detection is carried out on the segmented balls; the pixels and positions of the 2 edge images are compared to further determine the local edge of the bearing ball. Finally, the circle is fitted and the size is measured. The method can not only improve the accuracy of ball edge locating, but also avoid errors caused by direct edge detection.

5. Experimental Result and Analysis

5.1. Comparison of Images Collected by Various Light Sources as Illumination Sources. The STC15W204S is used as the main control chip, and the driver of WS2812 LED is written, so as to control the intensity and color of the light source. The upper computer communication module drives the serial port according to the universal asynchronous serial transceiver data transmission protocol, and the CH340 is used to build a circuit from USB to serial port to achieve communication between the upper computer and the lower computer. The real experimental equipment is shown in Figure 7.

As shown in Figure 8, serious halo phenomenon appears under normal light source illumination, and the ball details in the gap and bearing surface information are seriously disturbed. Although the outline of the bearing is clear under backlight illumination, the light is blocked by the bearing, and only a few rays enter the camera, so the surface information and details of bearing cannot be reflected. Under the illumination of the coaxial light source, the interference of reflection is basically overcome, and the details of the bearing surface are clear, but the light in the gap is insufficient, so it is difficult to observe the details in the gap. Therefore, the splitting principle of the spectroscope is analyzed systematically, an illumination system combining coaxial light source and backlight light source is designed, and the bearing image collected under the light source of this system has clear outline and complete surface details.

5.2. Result and Analysis of Ball Segmentation

5.2.1. Dataset and Network Training. The model of bearing tested in this experiment is 3D180506KJ8T2R3, with an inner diameter of 61.78 mm, an outer diameter of 29.92 mm, and a ball diameter of 9.79 mm. First of all, we collected a large number of bearing ball images with similar angles and positions, and then fixed-size region of 256×256 containing complete ball features is easy to cut out from the original image. The conditions and scenes of image acquisition used in this experiment are similar, and the similarity of samples is relatively high, so we may not need too many samples. However, in order to get a model with stronger generalization ability and avoid the problem of underfitting caused by fewer data samples, this paper expands the data through data enhancement. Finally, 750 ball images were obtained as training set and 200 as test set. An example of a sample is shown in Figure 9.

In this paper, Tensorflow, the mainstream development framework of deep learning, is selected to build the DeepLab V3+ network model, and the specific configuration of the experimental environment is shown in Table 1.

In this model, the adjusted Xception network is used as the basic network, the batch size is set to 4, and the initial learning rate is 0.01. The momentum optimizer is used to adjust the learning rate and update the model parameters in the iterative process, and the momentum coefficient is set to 0.9. Based on the fact that the ball pixels only occupy a small part of the whole image, the weight value of the loss function
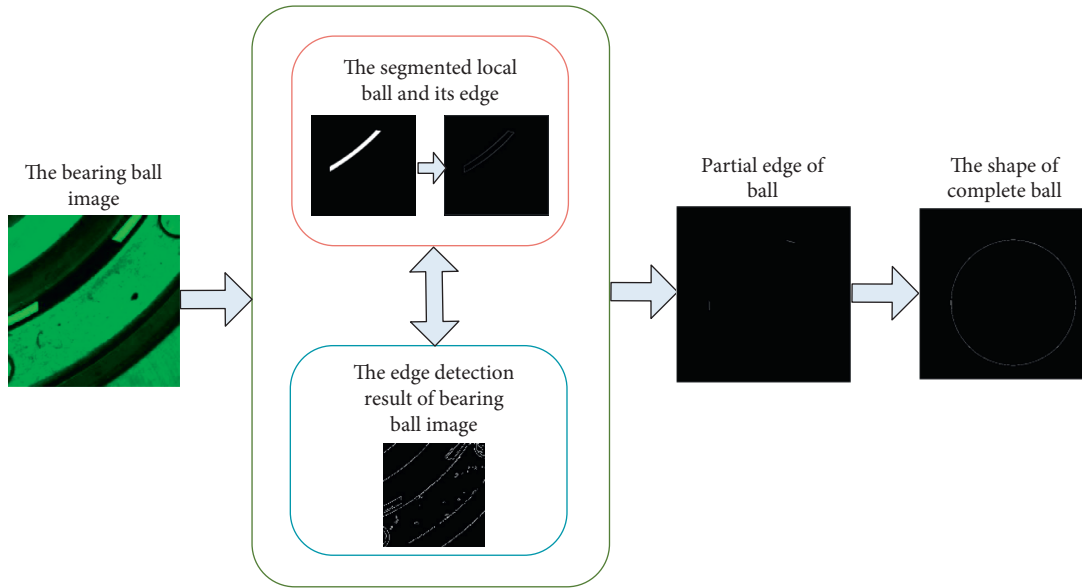
FIGURE 6: Algorithm flow of ball recognition and measurement.

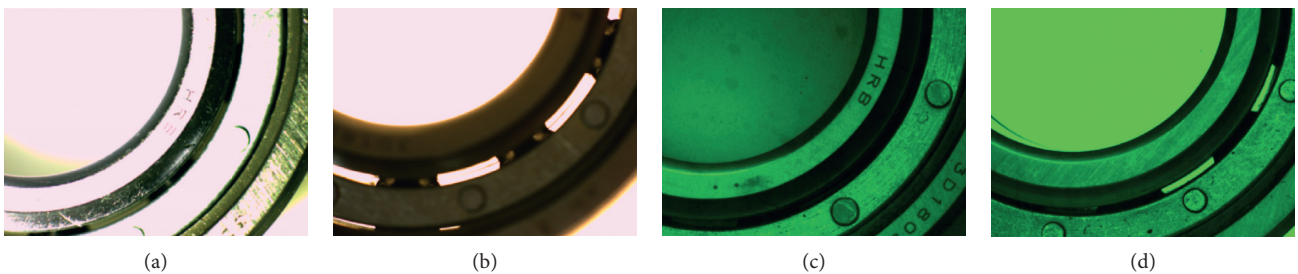FIGURE 7: The real scene of image collection in this experiment.

FIGURE 8: Image collection effect under various light sources. (a) Normal light source. (b) Backlight source. (c) Coaxial light source. (d) The light source designed in this paper.

is much larger than the background, so the loss function is set as a weighted cross-entropy function.

5.2.2. Evaluation Criteria and Comparison of Segmentation Results. In view of the performance evaluation of the improved aviation bearing ball segmentation network, this paper selects quotas such as mIOU (mean intersection over union), PA (pixel accuracy), recall, and *F* value to evaluate the network segmentation performance. mIOU is the most commonly used quota in semantic segmentation experiments. Its value is to first calculate the ratio of intersection and union of two sets of real and predicted values on each class and then calculate the average of intersection and union ratio of all classes, that is, the mean intersection over union. The formula is expressed as follows:
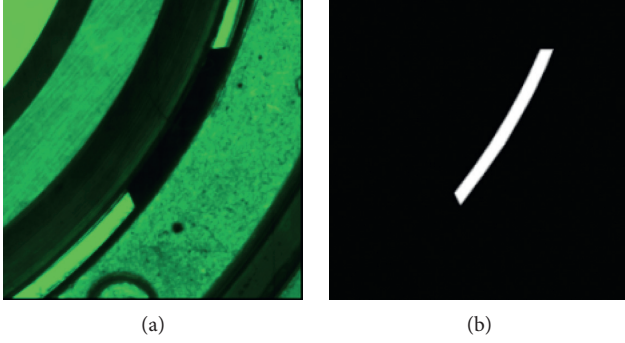
(a)   (b)

FIGURE 9: Example of sample labeling. (a) Original image. (b) Label.

TABLE 1: Specific configuration of experimental environment.

Name	Deploy
Operating system	Win10
Deep learning framework	Tensorflow1.11.0
CUDA	Cuda9.0
Cudnn	Cudnn7.0
CPU	Intel Core i7 7700HQ
Running memory	16 GB
Graphics chip	GTX 1080 max-Q

$$\text{mIOU} = \frac{1}{k+1} \sum_{a=0}^{k} \frac{p_{aa}}{\sum_{b=0}^{k} p_{ab} + \sum_{b=0}^{k} p_{ba} - p_{aa}}, \quad (5)$$

$$\text{PA} = \frac{\text{TP}}{\text{TP} + \text{FP}},$$

where k represents the number of foreground classes and an image can be divided into $k + 1$ classes; in this paper, k is 1, and there are only two types of ball and backgrounds. p_{aa} represents the number of pixels where class a is predicted to be class a. p_{ab} represents the number of pixels where class a is predicted to be class b. p_{ba} represents the number of pixels where class b is predicted to be class a. PA is the proportion of each correctly classified pixel to the total number of pixels marked as belonging to the classification, named pixel accuracy, as follows:

$$\text{recall} = \frac{\text{TP}}{\text{TP} + \text{FN}}. \quad (6)$$

However, due to the negative correlation between precision and recall rate in general, considering both precision and completeness, F value is introduced:

$$F = \frac{2\text{PA}^* \text{recall}}{\text{PA} + \text{recall}}. \quad (7)$$

Firstly, the traditional threshold segmentation method is used to segment the ball feature. From the gray histogram of Figure 10, it can be seen that although the gray distribution of the image is bimodal, and there is a big difference between the target and the background in the displayed image, the result after threshold segmentation is not optimistic, some inner rings and dust cover edges with similar gray values of the ball are mistaken for ball features, and there are many noise spots on the segmented edges. This segmentation result obviously cannot meet the requirements of accurate positioning and detection in this paper.

When using semantic segmentation to do experiments, in order to verify the feasibility of the proposed network, the U-Net and the original DeepLab v3+ network are compared. The same test set is segmented by trained networks. The segmentation performance comparison is shown in Table 2. From the differences of four performance quotas mentioned in this paper, some conclusions can be drawn. The segmentation results of DeepLab v3+ network are more outstanding in mIOU, PA, recall, and F values compared with U-Net, which are 0.732, 0.805, 0.913, and 0.857, respectively. And the performance of the improved network in this paper is further improved. A 1.6% increase of mIOU shows that the effect of distinguishing target and background is better, a 2.2% increase of PA shows that the segmentation accuracy is improved and the increase of F shows that the overall segmentation performance is better although the recall decreases slightly, which shows that the integrity of segmentation is maintained.

The 3 networks are used to carry out comparative segmentation experiments. Some predicted results are shown in Figure 11, and the segmentation effect can be judged subjectively from the figure. It can be seen from the results of the U-Net that there are some errors in the classification of ball features and background; especially, the segmentation of areas with unclear grayscale contrast of ball boundary is not accurate enough, and the extracted edge details are fuzzy. Although the DeepLab v3+ network uses dilated convolution to expand the receptive field on the basis of encoder and decoder structure, the segmentation results have certain progress, but there are still errors in the distinction between background and region of interest.

The improved network uses CBAM to combine spatial dimension information on the basis of channel, considering both channel feature correlation and spatial pixel feature correlation, and learns the parts that need to be paid attention to in the feature map, and the weighted summation operation makes the features of important areas in the image more obvious and improves the expression ability of the convolutional neural network. It has a higher degree of discrimination for the boundary area of the target where grayscale contrast is not obvious and has a better effect on boundary refinement. So the improved network is more suitable for ball segmentation which occupies a small pixel width with serious background interference.

5.3. Result of Bearing Ball Detection.

Firstly, aiming at the circle detection, the relationship between the accuracy of extracted edge points and the times of three-point sampling is analyzed. Assuming that only nt of the N edge points are really the edge points of the ball, the probability that the sampling points are real points is $p = (nt/N)$, and the probability of getting the real shape of the ball after three-point random circle detection is
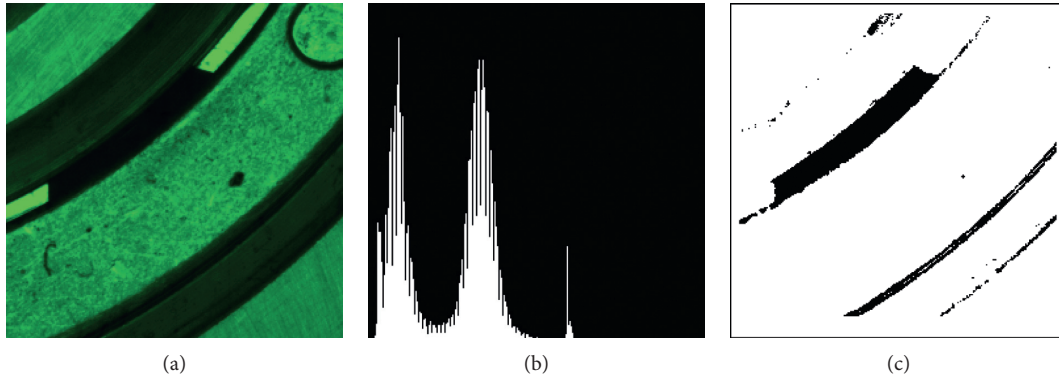
(a)   (b)   (c)

FIGURE 10: Image feature analysis of aviation bearing ball. (a) Original ball image. (b) Gray histogram. (c) Threshold segmented image.

TABLE 2: Comparison table of segmentation performance of different segmentation networks.

Segmentation network	Quota			
	mIOU	PA	Recall	F
U-Net	0.709	0.776	0.911	0.838
DeepLab v3+	0.732	0.805	0.913	0.857
The improved network	0.748	0.827	0.906	0.864

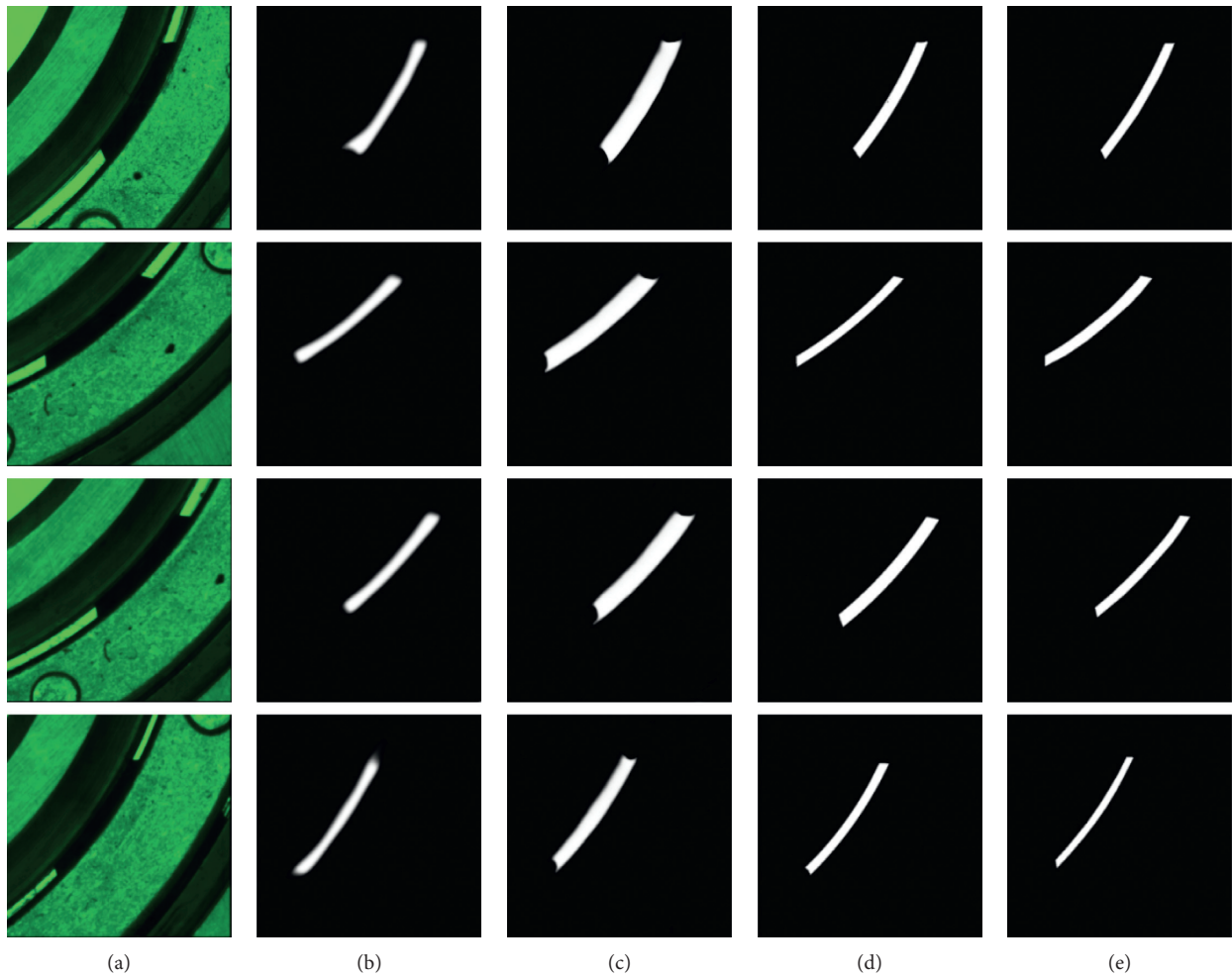(a)   (b)   (c)   (d)   (e)

FIGURE 11: Comparison of segmentation results. (a) Test images. (b) U-Net segmentation results. (c) Original DeepLab v3+ segmentation results. (d) Improved DeepLab v3+ segmentation results. (e) Labels.
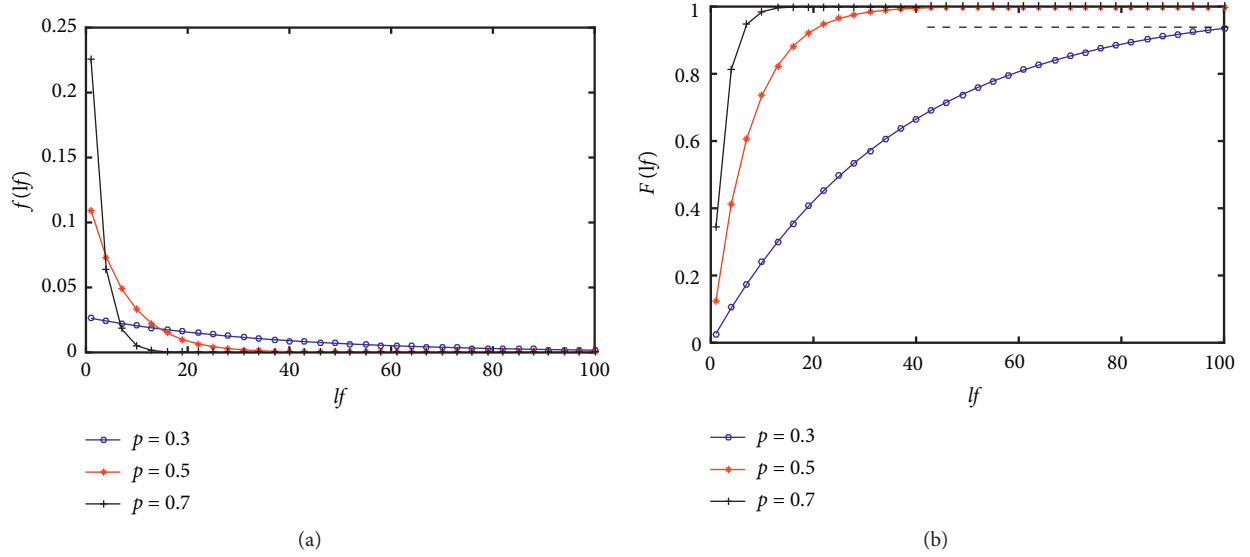
(a)

(b)

FIGURE 12: Relationship between the accuracy of extracted edge points and the times of three-point sampling. (a) Probability function curve. (b) Distribution function curve.

TABLE 3: Statistical table of experimental results.

Size measurement error (μm)	Original DeepLab v3+ network	Improved DeepLab v3+ network
0–20	27	99
20–40	86	73
40–60	41	17
60–80	35	9
Over 80	11	2
Average measurement error	41.0526	26.7317
Standard deviation	20.8490	17.1806
False detection rate	5.5%	1%

$$P = \frac{nt\,(nt-1)\,(nt-2)}{N\,(N-1)\,(N-2)}, \tag{8}$$

where P is approximately equal to p^3, and the probability and distribution function of the optimal circle obtained after L times of three-point circle detection are as follows:

$$f(lf) = \left(1 - p^3\right)^{lf}\left(p^3\right),$$
$$F(lf) = \sum_{j=1}^{lf} \left(1 - p^3\right)^{j}\left(p^3\right), \tag{9}$$

where the random circles obtained in the first lf times are not the final fitting circles, $lf = L - 1$.

When p is 0.3, 0.5, and 0.7, the corresponding probability function curve and distribution function curve are shown in the figure.

It can be seen from Figure 12 that when p is 0.7, the optimal circle can be obtained in less than 20 samples. In the three-point random circle detection experiment, the optimal circle can be obtained after 10–20 sampling, which also reflects that the accuracy of ball edge points obtained by the ball feature recognition method proposed in this paper is

over 70%. For the sake of conservatism, the sampling number is set to 20.

On the basis of DeepLab v3+ network segmentation and circle detection, the complete ball shape is fitted and the size of the ball is measured. The original DeepLab v3+ network and the improved DeepLab v3+ network are used to carry out segmentation experiments on bearing balls, respectively. Then circle detection and size measurement are carried out on the segmentation results and whether the ball is wrong assembled is judged. The statistic of experimental results is shown in Table 3.

Because of the large test set, the experimental results are counted in sections. From Table 3, it can be seen that the ball size measurement errors after segmentation by the original DeepLab v3+ network are mostly concentrated in the range of 30–80 microns, the average measurement error is about 41 microns, and the standard deviation is 20.85 microns. The error of ball size measurement after segmented by the improved DeepLab v3+ network in this paper is mostly concentrated in 0–40 microns, the average error is about 26.73 microns, and the standard deviation is 17.18 microns. The result shows that the improved network locating and segmentation effect is better so that the size measurement result

is more accurate, and the accuracy of size measurement is high enough to be used as a standard to judge whether the assembled ball is qualified or not.

The detection of ball after the assembly of the aviation bearing is the last threshold to determine the assembly qualification, so the correct rate of wrong assembly judgment is required to be extremely high in order to achieve zero miss rate of unqualified bearings in industrial production. Since the sizes of balls of different specifications which may be wrongly assembled differ only by 198 microns, the measurement error of ball size must be within 99 microns. In order to ensure the correct rate of judgment, when the measurement error is greater than 80 microns, it is determined that the bearing ball is wrongly assembled and the bearing is unqualified. It can be seen from Table 2 that the false detection rate is reduced from 5.5% when the original network is segmented to 1%, and when the improved network is segmented, the detection accuracy rate is obviously improved. In view of the large number of industrial inspections, the zero false detection rate of unqualified bearings can be realized. So the method proposed in this paper can be effectively applied to the last assembly detection of aviation bearing.

6. Conclusion

Aiming at the problem that the full shape of the ball in aviation bearing can only be predicted by small-scale local features, this paper studies the segmentation and recognition method of the embedded ball in aviation bearing. Firstly, an image acquisition system based on the coaxial light source and backlight light source is designed, which is used for solving the problem of light reflection on the surface of aviation bearing and the problem of insufficient imaging light in a gap. Then, a ball segmentation network model based on DeepLab v3+ semantic segmentation network is designed to accurately locate the ball features in complex background. What is more, the CBAM module is introduced into DeepLab v3+ network, which can make the features of important areas in the image more significant through weighted summation operation and improve the expression ability of the network, so that the network can distinguish between target and background better, and the edges can be thinned. Finally, according to the characteristic that the ball edges are 2 short arcs composed of a few pixels, a random circle detection method based on circle fitting degree is proposed, and the final fitted circle is obtained by finding the optimal value. Experimental verification is carried out in this paper, the experimental result shows that the size detection accuracy is about 26.73 microns, and the false detection rate of unqualified bearing is 1%, which meets the requirement of zero missed inspection in industry. At present, the method proposed in this paper is only for single ball detection. If we want to detect more than two balls, the semantic segmentation will not be able to classify them well because the ball features are too similar, which is what we will study later.

Data Availability

The codes used in this paper are available from the author upon request.

Conflicts of Interest

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

Acknowledgments

References

[1] Q. Wang, "Research on key technologies of surface defect classification of bearing outer ring based on deep learning," Master thesis, University of Electronic Science and Technology, Chengdu, China, 2020.

[2] X. Du, Z. Chen, Y. Wang et al., "Rolling bearing fault diagnosis based on IEWT and IFractalNet," *Vibration and Shock*, vol. 39, no. 24, pp. 134–142, 2020.

[3] Z. Xu, L. Hu, and Y. Chen, "Brake pad dimension detection based on dimension tolerance and image processing," *Chinese Journal of Scientific Instrument*, vol. 38, no. 9, pp. 2317–2325, 2017.

[4] H. Jiang, C. Xu, Y. Chen et al., "Detection and counting method of small-sized rice ears with partial occlusion based on field images," *Transactions of the Chinese Society for Agricultural Machinery*, vol. 51, no. 9, pp. 152–162, 2020.

[5] F. Jiang, Q. Gu, H. Hao et al., "Summary of content-based image segmentation methods," *Journal of Software*, vol. 28, no. 1, pp. 160–183, 2017.

[6] H. Zang, K. Wang, and F. Wang, "The application progress and prospect of deep learning in target vision detection," *Acta Automatica Sinica*, vol. 43, no. 8, pp. 1289–1305, 2017.

[7] J. Long, S. Evan, and D. Trevor, "Fully convolutional networks for semantic segmentation," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 3431–3440, Boston, MA, USA, June 2015.

[8] V. Badrinarayanan, A. Cipolla, and C. Roberto, "SegNet: a deep convolutional encoder-decoder architecture for image segmentation," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 39, no. 12, pp. 2481–2495, 2017.

[9] R. Olaf, F. Philippand, and B. Thomas, "U-net: convolutional networks for biomedical image segmentation," in *Proceedings of the International Conference on Medical Image Computing and Computer-Assisted Intervention (MICCAI)*, pp. 234–241, Munich, Germany, October 2015.

[10] L. Chen, G. Papandreou, and I. Kokkinos, "Semantic image segmentation with deep convolutional nets and fully connected CRFs," in *Proceedings of the International Conference on Learning Representations(ICLR)*, San Diego, CA, USA, May 2015.

[11] L.-C. Chen, G. Papandreou, I. Kokkinos, K. Murphy, and A. L. Yuille, "DeepLab: semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected CRFs," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 40, no. 4, pp. 834–848, 2018.

[12] L. Chen and G. Papandreou, "Rethinking atrous convolution for semantic image segmentation," 2017, http://arxiv.org/abs/1706.05587.

[13] L. Chen, Y. Zhu, and G. Papandreou, "Encoder-decoder with atrous separable convolution for semantic image segmentation," in *Proceedings of the European Conference on Computer Vision (ECCV)*, pp. 801–818, Munich, Germany, September 2018.

[14] K. He, X. Zhang, S. Ren et al., "Deep residual learning for image recognition," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 770–778, Las Vegas, NV, USA, July 2016.

[15] B. Yang, Q. Tao, and P. Dong, "Surgical instrument segmentation method based on improved deeplab v3+ network," 2020, http://kns.cnki.net/kcms/detail/11.2127.TP.20200221.0832.004.html.

[16] X. Wang, "Research on semantic segmentation technology of power line based on image," Master thesis, Zhejiang University, Hanzhou, China, 2019.

[17] F. Chollet, "Xception: deep learning with depthwise separable convolutions," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 1800–1807, Honolulu, HI, USA, July 2017.

[18] K. He, X. Zhang, S. Ren et al., "Spatial pyramid pooling in deep convolutional networks for visual recognition," *IEEE Transactions on Pattern Analysis & Machine Intelligence*, vol. 37, no. 9, pp. 1904–1916, 2014.

[19] S. Woo, J. Park, J. Lee et al., "CBAM: convolutional block attention module," in *Proceedings of the European Conference on Computer Vision (ECCV)*, pp. 3–19, Munich, Germany, September 2018.

[20] D. Yu and W. Wei, "Improved random circle detection algorithm," *Journal of Image and Graphics*, vol. 14, no. 8, pp. 1590–1593, 2009.

[21] X. Zhou, Y. Wang, K. Li et al., "New bottle mouth positioning method based on multiple randomized circle detection and fitting degree evaluation," *Chinese Journal of Scientific Instrument*, vol. 36, no. 9, pp. 2021–2029, 2015.

[22] J. Zhu, X. Li, W. Tan et al., "short arc measurement based on least squares circle fitting with center constraint," *Optical Precision Engineering*, vol. 17, no. 10, pp. 2486–2492, 2009.

[23] H. Qu, Y. Wang, and C. Wu, "Circle detection algorithm based on Hough transform of diameter accumulation," *Journal of Xi'an University of Posts and Telecommunications*, vol. 22, no. 5, pp. 89–93, 2017.