

Article

The Bearing Surface Defect Detection Method Combining Magnetic Particle Testing and Deep Learning

Long Li ¹, Zhiyuan Liu ^{2,*}, Hengyi Zhao ², Lin Xue ²  and Jianbo Wu ³ 

¹ School of Naval Architecture, Ocean and Civil Engineering, Shanghai Jiaotong University, Shanghai 200240, China; llsjtu@outlook.com

² School of Mechanical Engineering, Dalian University of Technology, Dalian 116081, China; zhaohengyi2022@mail.dlut.edu.cn (H.Z.); linxue@dlut.edu.cn (L.X.)

³ School of Mechanical Engineering, Sichuan University, Chengdu 610065, China; wujianbo@scu.edu.cn

* Correspondence: zylui@mail.dlut.edu.cn

Featured Application: The research results enable the rapid and accurate detection of surface defects in the bearing-manufacturing process.

Abstract: As a critical foundational component, bearings find widespread application in various mechanical equipment. In order to achieve automated defect detection in the bearing-manufacturing process, a defect detection algorithm combining magnetic particle inspection with deep learning is proposed. Dynamic thresholding and generative adversarial network (GAN) methods are employed to extract defect samples from bearing images and augment the dataset, thereby enhancing data diversity. To mitigate the impact of irrelevant displays in bearing images, a coordinated attention (CA) mechanism is introduced into the backbone network of the deep learning model to focus on key information. Additionally, an adaptive spatial feature fusion module (ASFF) is incorporated during the multiscale fusion stage to maintain consistency in features across different hierarchical levels. The weighted intersection over union (WIoU) bounding box loss function is utilized to replace the original generalized intersection over union (GIoU) in the network, directing the model's attention towards common-quality anchor boxes to reduce the adverse effects of inconsistent annotations. The experimental results demonstrate that the improved network achieves a mean average precision (mAP) of 98.4% on the bearing dataset, representing a 4.2% improvement over the original network.

Keywords: bearings; defect detection; magnetic particle testing; deep learning; data augmentation



Citation: Li, L.; Liu, Z.; Zhao, H.; Xue, L.; Wu, J. The Bearing Surface Defect Detection Method Combining Magnetic Particle Testing and Deep Learning. *Appl. Sci.* **2024**, *14*, 1747. <https://doi.org/10.3390/app14051747>

Academic Editors: Grzegorz Psuj and Barbara Grochowalska (Szymanik)

Received: 13 January 2024

Revised: 14 February 2024

Accepted: 16 February 2024

Published: 21 February 2024



Copyright: © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Bearings, as crucial components widely used in manufacturing, are acclaimed as the “joints of industry” [1]. Their primary functions include reducing wear on various components and lowering friction coefficients to ensure the longevity of various equipment [2]. However, during the manufacturing process, various defects may occur on the surface of bearings, leading to a decrease in mechanical efficiency, an increase in failure losses, and ultimately a reduction in the lifespan of mechanical equipment [3]. The quality of bearings directly influences the stability and safety of mechanical equipment [4].

Surface defects on bearings, such as scratches and cracks, significantly reduce the service life of the bearings, directly leading to equipment failures [5]. Therefore, it is essential to conduct defect detection on bearing surfaces during the production stage. By promptly identifying and eliminating defects in bearings, it is possible to effectively enhance the stability and reliability of equipment, reduce failure rates, and thereby extend the operational lifespan of mechanical equipment. Currently, non-destructive testing methods for bearings mainly include ultrasonic testing [6] and magnetic particle testing [7]. However, these methods still require human intervention in the final decision-making process, making them susceptible to subjective factors and potentially resulting in missed

or false detections of bearing defects, thereby reducing the accuracy and efficiency of the detection process.

With the development of machine vision, some enterprises are opting to use industrial cameras instead of human eyes, employing defect detection algorithms to replace manual decision making [8]. Existing defect detection algorithms can be broadly categorized into traditional image processing methods [9] and deep learning-based object detection algorithms [10]. Traditional image processing methods require algorithm engineers to manually extract defect features and set different thresholds, exhibiting poor robustness and high false-detection rates. In contrast, deep learning-based object detection algorithms autonomously learn defect features through neural networks, requiring no human intervention and demonstrating superior robustness and controllable false detection rates, especially in recognizing complex defect images. However, deep learning methods require substantial data support to avoid overfitting. The characteristics of the deep learning method and traditional image processing method are shown in Table 1.

Table 1. Contrast between deep learning and traditional image processing.

	Deep Learning	Traditional Image Processing
accuracy	high	low
durability	high	low
cost	high	low
test time	long	short

This paper innovatively combines magnetic particle inspection with deep learning for bearing defect detection. The magnetic traces at defect locations during the magnetic particle inspection process exhibit distinctive image features, providing excellent learning samples for deep learning models. This approach allows for deep learning to comprehensively learn bearing defect features, resulting in a trained model with precise defect detection capabilities.

2. Magnetic Particle Testing

Presently, non-destructive testing methods commonly used for surface defect detection in bearings include ultrasonic testing and magnetic particle testing. Ultrasonic testing primarily aims to identify internal defects within the bearing, while magnetic particle testing is specifically employed for detecting defects on the bearing's surface and in the near-surface region. The primary steps involved in magnetic particle testing generally encompass the six procedures illustrated in Figure 1 [11].

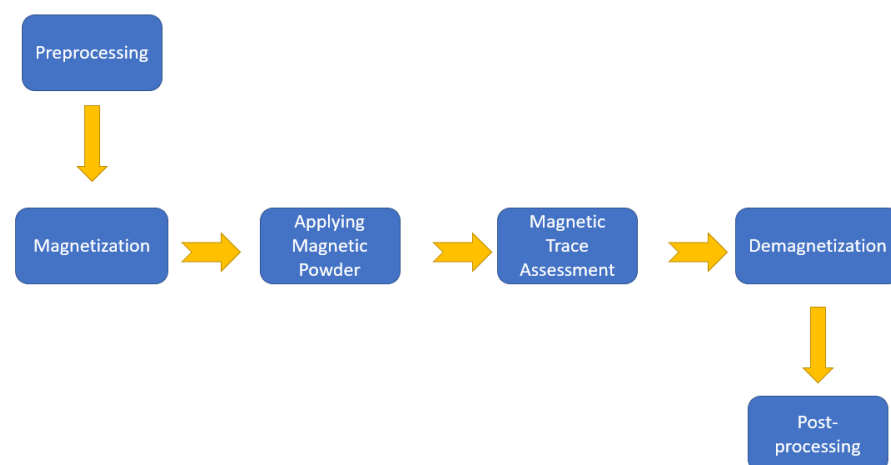


Figure 1. The main procedures of magnetic particle testing.

- (1) Preprocessing: Improving the surface condition of the workpiece before magnetization is essential to ensure an even distribution of magnetic particles on the workpiece's surface.
- (2) Magnetization: The application of a magnetic field to the workpiece is known as magnetization, inducing a leakage magnetic field at defect locations.
- (3) Applying Magnetic Powder: Spraying magnetic powder onto the workpiece's surface is the process of applying magnetic powder, attracting the powder to defect-induced leakage magnetic fields, resulting in the formation of magnetic traces. This application can be performed using either a dry method, employing solid magnetic powder, or a wet method, utilizing a magnetic suspension containing magnetic powder.
- (4) Magnetic Trace Assessment: Inspection personnel observe and assess the position, shape, and size of magnetic traces, determining the nature and dimensions of defects.
- (5) Demagnetization: Demagnetization involves subjecting the workpiece to an alternating magnetic field, utilizing the magnetic hysteresis loop's decreasing effect to reduce residual magnetism within the workpiece, ensuring that residual magnetic fields do not affect the workpiece's subsequent use.
- (6) Post-processing: Post-processing includes tasks such as cleaning the workpiece's surface and rust prevention, aiming to eliminate any impact on the workpiece's subsequent use from the applied magnetic powder, water-based magnetic suspension, contrast enhancement agents, etc.

In the current operation of magnetic particle testing for bearings, most of the procedural steps can be automated by inspection equipment. For instance, a bearing magnetic particle-testing machine can automatically handle tasks such as bearing loading, magnetization, spray application, rotation, and demagnetization. However, the observation, evaluation, and corresponding record-keeping of magnetic traces still require manual intervention. During the magnetic trace observation and evaluation process in bearing magnetic particle testing, inspection personnel must control the rotation of the bearing and carefully examine the magnetic traces on the surface. This process involves numerous observation points, considerable observation distances, and high work intensity. Particularly when using fluorescent magnetic powder for inspection (as shown in Figure 2), inspection personnel need to be exposed to ultraviolet light for an extended period, working in challenging conditions. Additionally, the fluorescence of the magnetic powder and the reflection from the ultraviolet lamp can be somewhat stimulating to the eyes, leading to eye fatigue. This may cause inspection personnel to lose concentration, resulting in potential defects being overlooked or misjudged.

Magnetic particle testing is a method for detecting defects by identifying the accumulation of magnetic particles forming magnetic traces at defect locations [12]. The principle is illustrated in Figure 3: after magnetization, a magnetic field is generated inside the ferromagnetic material of the workpiece. When defects are present on or near the surface of the workpiece, the discontinuity of material at the defect causes a local change in magnetic permeability, resulting in a distortion of the magnetic field and the leakage of magnetic flux from the workpiece surface, forming a leakage magnetic field. Under the influence of the leakage magnetic field, magnetic particles are attracted and gather at the defect location, forming magnetic traces associated with the defect's leakage magnetic field. These traces can highlight defects that are not easily visible, and by observing and analyzing the geometric shape of the magnetic traces, workers can identify surface defects on the workpiece. Fluorescent magnetic powder, containing fluorescent dyes, emits yellow-green fluorescence in the 510–550 nm wavelength range when exposed to ultraviolet light. This creates a high-contrast fluorescence against the workpiece's surface color, making the magnetic traces easier to observe. Especially when using fluorescent magnetic powder in a darkroom, the fluorescence of the magnetic traces exhibits extremely high contrast with the workpiece surface, significantly enhancing the sensitivity and inspection speed of magnetic particle testing.

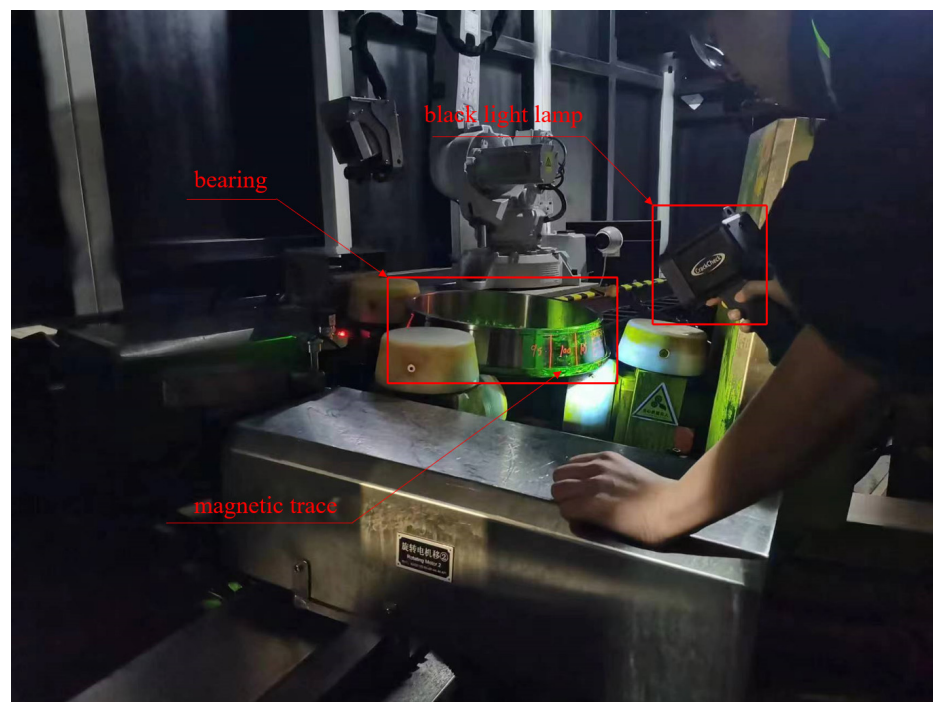


Figure 2. Manual observation of bearing defects.

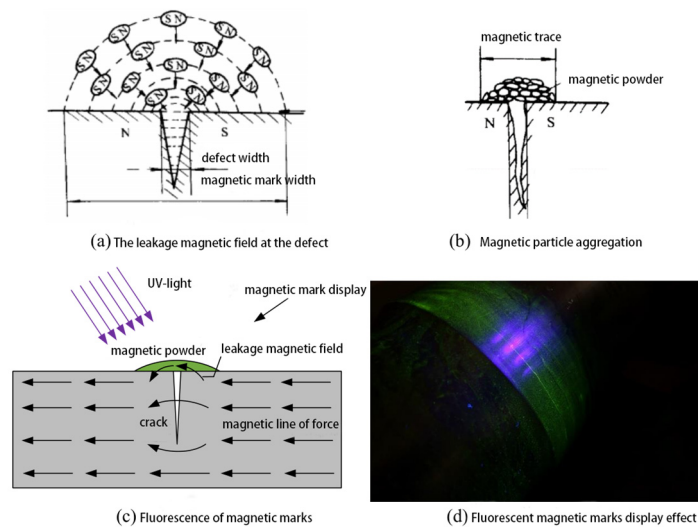


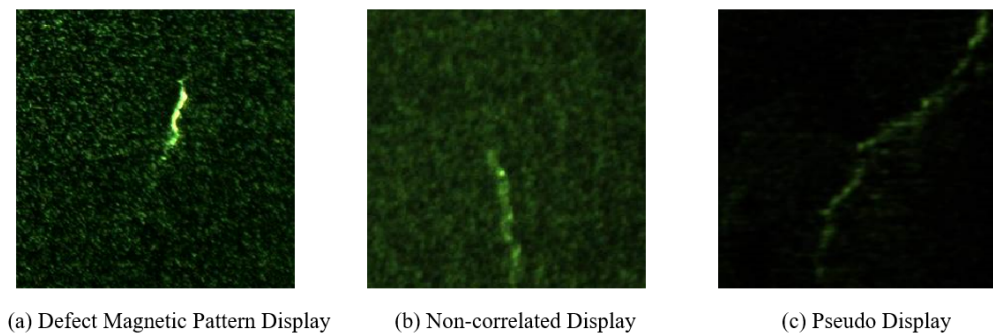
Figure 3. Magnetic particle testing principle.

The generation of magnetic traces involves various factors, including the leakage magnetic field at the defect, the leakage magnetic field in non-defective areas, and even the natural deposition and retention of magnetic particles on the workpiece surface [13]. Generally, magnetic traces formed by the leakage magnetic field generated by defects are referred to as defect magnetic trace displays, also known as correlated displays. Correspondingly, magnetic traces formed by the attraction of magnetic particles due to larger leakage magnetic fields caused by changes in the workpiece's cross-section and material property differences are referred to as non-correlated displays. Magnetic traces not formed by the leakage magnetic field on the workpiece surface are referred to as pseudo displays. The magnetic particle information used in this paper is shown in the following Table 2.

Table 2. Magnetic particle technical properties.

Technical Properties	Value
Particle size μm	approx. 3 μm
Density/15 °C	$1140 \pm 15 \text{ kg/m}^3$
weight percent	2.5%
pH value	8.75 ± 0.25
permeability	approx. 10^{-3}

Different causes result in magnetic traces with distinct geometric features, which are captured by the camera and form magnetic trace images, as shown in Figure 4.

**Figure 4.** Images of different types of magnetic marks.

3. Machine Vision

Machine vision is the automatic acceptance and processing of images of a real object through optical devices and non-contact sensors. It involves analyzing the images to obtain the required information or control machine movements. Currently, it is widely used in various non-destructive testing applications. In the context of bearing magnetic particle inspection, machine-vision methods are employed to detect defects, enabling automated inspection operations and defect recognition. This automation improves the working environment, reduces labor intensity, and lowers the rate of missed detections, showing promising applications.

In machine vision, the process of converting actual magnetic traces into magnetic trace images involves two main steps. The first step is optical imaging, where the three-dimensional magnetic traces on the surface of the object are projected onto the imaging target surface of the camera through the optical system, resulting in a two-dimensional light signal. The second step is image digitization, where the camera converts the light signals on the target surface into discrete digital signals for computer processing. These two steps together represent the transformation of the geometric features of magnetic traces into discrete brightness-distribution information on a two-dimensional plane in the image, as illustrated in Figure 5.

Image recognition refers to the technique of processing, analyzing, and understanding images using software programs. It transforms image data described in pixels into non-graphical image information, matching or classifying to identify the various patterns of targets and objects. When using machine-vision technology for detecting cracks in magnetic traces during bearing magnetic particle inspection, it is necessary to employ image recognition algorithms tailored for crack detection. These algorithms analyze the collected magnetic trace images to detect cracks and extract information about their shape, size, and position. Thus, image recognition technology is crucial for automated crack detection in the full-process magnetic particle inspection of bearings.

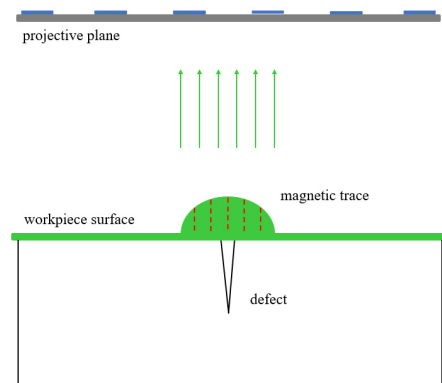


Figure 5. Collection of magnetic trace images.

Current image recognition technologies mainly include traditional image processing methods and deep learning-based object detection methods, each with their own characteristics in defect detection. Traditional image processing methods rely on algorithm engineers manually extracting defect features and setting thresholds, which may lead to issues such as insufficient robustness and high false-positive rates in practical applications. In contrast, deep learning object detection methods autonomously learn defect features through neural networks, requiring no human intervention and offering greater flexibility to adapt to various complex defect scenarios. This approach enhances algorithm robustness while effectively controlling false-positive rates. Deep learning methods possess powerful learning capabilities, handling large-scale data and alleviating the burden of manually designing features in traditional methods. In bearing defect detection, deep learning object detection methods can automatically learn the features of magnetic traces, resulting in more accurate defect identification and improved detection accuracy and efficiency.

Considering these factors, deep learning object detection methods demonstrate more advantages in bearing defect detection compared to traditional image processing methods, especially in handling complex scenes and large-scale data. Therefore, this paper adopts the deep learning object detection method combined with magnetic particle inspection for bearing defect detection to achieve superior detection results.

4. Principles and Optimization of the YOLOv5 Model

The YOLOv5 [14–16] network is widely used in the field of object detection, and is characterized by fast detection speed, high accuracy, and low computational complexity. It can meet the basic requirements for bearing image defect detection. Therefore, YOLOv5 is adopted as the detection algorithm in this paper. Figure 6 shows the architecture of the YOLOv5 network. The network is mainly composed of four parts: the input end (Input), the backbone network (Backbone), the neck network (Neck), and the detection network (Head).

This paper optimized the original YOLOv5 network architecture to improve the accuracy of bearing defects detection. This included incorporating a coordinate attention mechanism into the backbone network, adding an adaptive fusion module to the neck, and replacing the loss function. The improved network structure is illustrated in Figure 7.

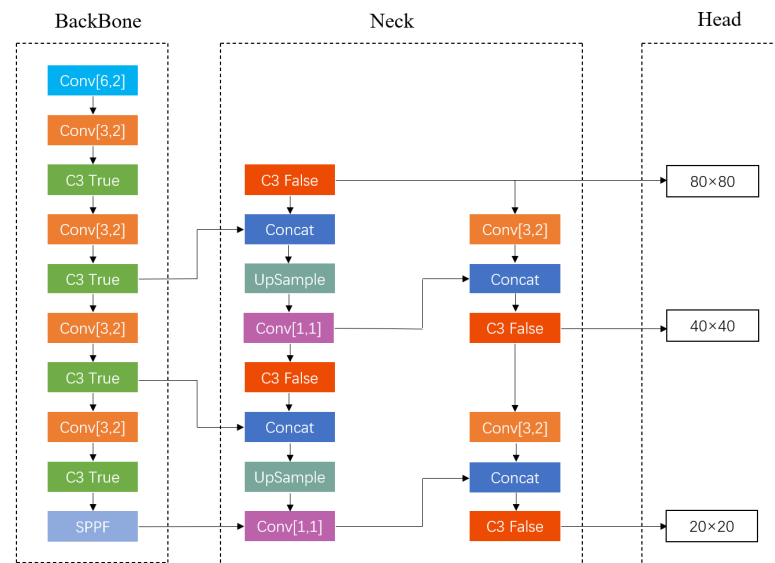


Figure 6. YOLOv5 network structure.

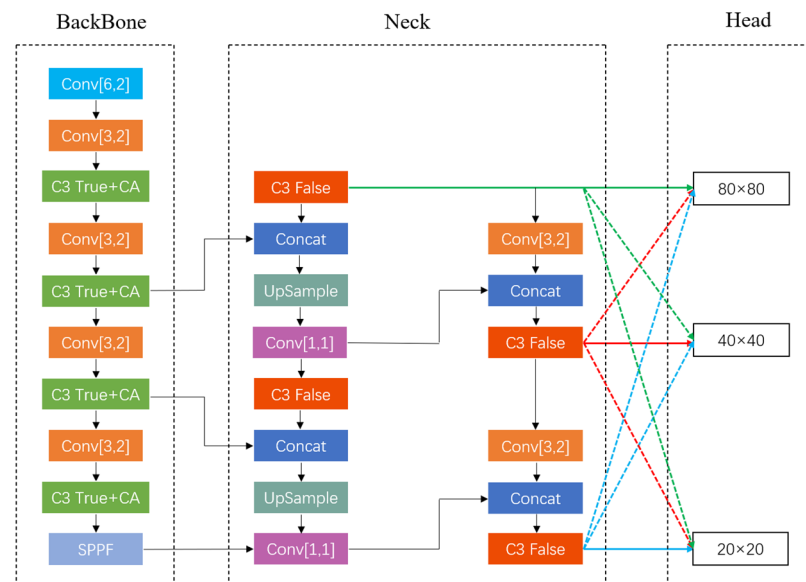


Figure 7. Improved YOLOv5 model.

4.1. CA Mechanism

The bearing images captured by industrial cameras have the characteristics of large size and numerous small defects. The pixel size of small defects is relatively small in the entire bearing image. Therefore, the feature information of these small defects is prone to being lost in multiple convolution and pooling operations. This can lead to the network inadequately learning the feature information of small defects and can result in the phenomenon of missed detection. Therefore, in the C3 module of the backbone network, this paper integrates the coordinate attention (CA) mechanism [17] to enhance the focus on the feature information of small defects, thereby improving the detection accuracy of small defects.

The channel attention mechanism was initially applied to improve deep learning models by adaptively learning the importance of each channel through compressing and exciting input features [18]. This mechanism can enhance the model's performance to some extent. However, the channel attention mechanism neglects the importance of positional information, which is a crucial aspect the network needs to focus on learning.

The CA mechanism embeds positional information into the channel attention, allowing the network to simultaneously consider encoding inter-channel information and positional information. The schematic diagram of the CA mechanism is shown in Figure 8.

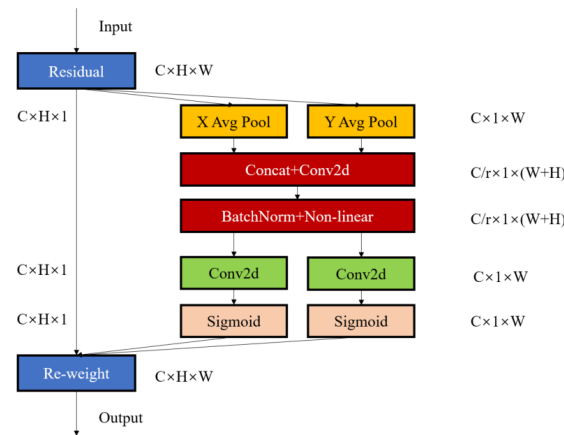


Figure 8. Schematic diagram of CA mechanism.

4.2. Adaptive Spatial Feature Fusion

In the neck network of YOLOv5, there is a solidified feature-matching mechanism; that is, the large target corresponds to the feature map of the high level, and the small target corresponds to the feature map of the low level. When the target of a feature layer is a positive sample, it means that the corresponding part of the other feature layer will be identified as a negative sample. The inconsistency between different levels of features will affect the gradient calculation in the training process and reduce the effectiveness of feature fusion.

In order to solve the inconsistency between different feature scales of the neck network, ASFF [19,20] is used in this paper to perform adaptive spatial feature fusion on the neck network (the principle of ASFF is shown in Figure 9) so that the neck network can filter the conflicting information in space to ensure the correct calculation of the gradient so as to improve the proportional invariance of the features. Since ASFF adopts an additive method to carry out adaptive spatial feature fusion, it is necessary to downsample or upsample the feature maps of different layers and adjust the number of channels to ensure that the feature maps of different feature layers have the same size and number of channels.

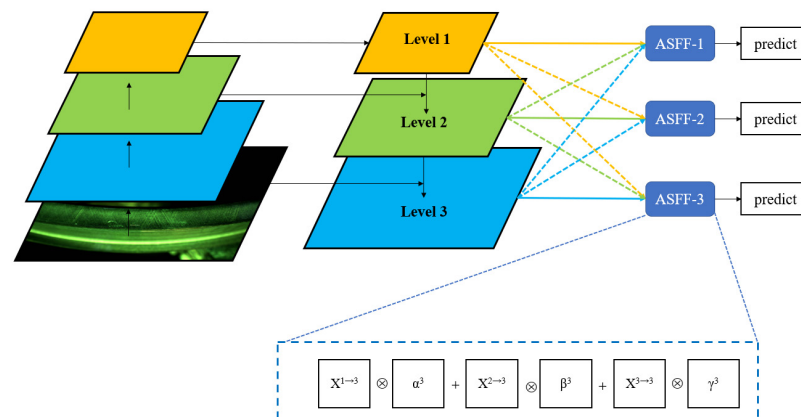


Figure 9. ASFF schematic diagram.

ASFF obtains new fusion features by assigning weights to different feature layers and adding them. The formula is shown in (1):

$$y_{ij}^1 = \alpha_{ij}^1 \cdot x_{ij}^{1 \rightarrow 1} + \beta_{ij}^1 \cdot x_{ij}^{2 \rightarrow 1} + \gamma_{ij}^1 \cdot x_{ij}^{3 \rightarrow 1} \quad (1)$$

where y is the new feature map after adaptive spatial feature fusion; x^1, x^2, x^3 are the three feature maps output by the neck network; and α, β, γ are the weights of the three feature maps in turn.

In the process of backpropagation, the weight of each feature layer will be updated, thus weakening the solidified feature-matching mechanism of the neck network and avoiding the interference of feature antagonism in the gradient calculation process so that the features can maintain consistency in different levels.

4.3. WIoU Loss Function

The performance of object detection models depends on the design of the loss functions. The bounding box loss function is a crucial component of the loss functions and is used to calculate the error between predicted bounding boxes and actual bounding boxes. In the YOLOv5 network, the bounding box loss function used is GIoU [20].

In the object detection dataset, a normal label refers to accurately and correctly bounding the position of the target by the annotator, while low-quality examples refer to cases where the annotator mistakenly selects the target's position or combines multiple targets into one bounding box. The basis of GIoU computation assumes that all examples in the dataset are of high quality, and strengthens the regression on anchor boxes accordingly, whereas the reality of datasets may not align with this assumption. Figure 10 shows normal labeling results, while Figure 11 illustrates low-quality examples. To address this issue, this paper uses WIoU [21] instead of the original GIoU in the network. WIoU focuses on regressing on regular-quality anchor boxes, which helps mitigate the impact of low-quality examples, thereby improving detection accuracy.

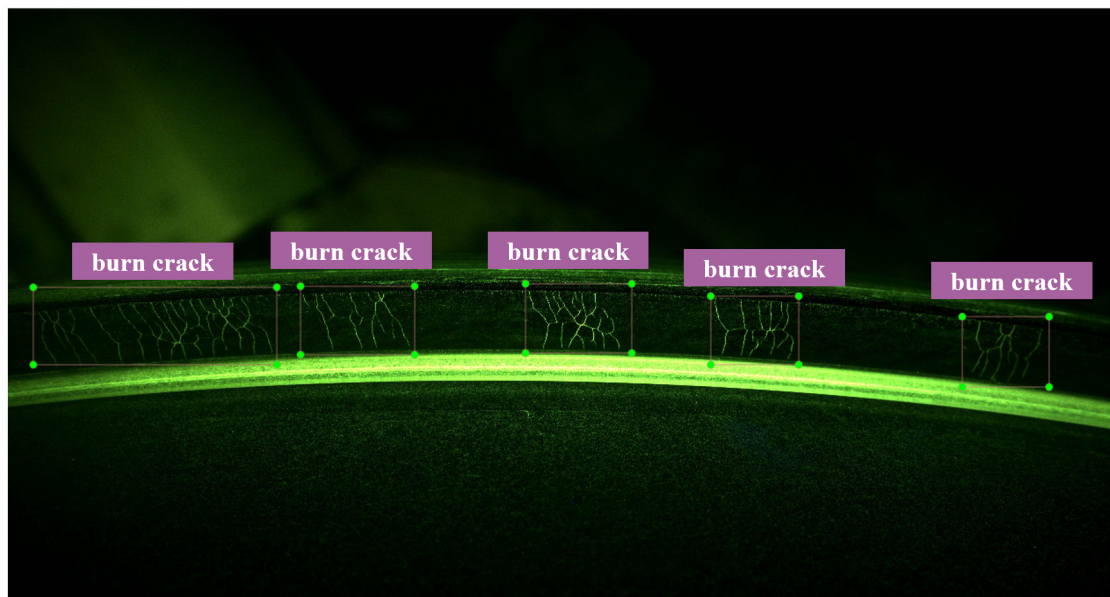


Figure 10. Normal labeling results.

This paper uses WIoU as the bounding box loss function, and the definition of the outlieriness β in WIoU is given by Equation (2):

$$\beta = \frac{L_{IoU}^*}{\overline{L_{IoU}}} \in [0, +\infty) \quad (2)$$

The definition of the outlier degree draws inspiration from the monotonic focusing mechanism of focal loss. L_{IoU}^* serves as a monotonic focusing factor, increasing monotonically with the rise in loss values. This effectively reduces the contribution of easy examples to the loss values, allowing the model to focus on challenging examples; $\overline{L_{IoU}}$ is the momentum with a sliding average of m .



Figure 11. Low-quality example.

Anchor boxes with lower outlieriness correspond to high-quality anchor boxes, while those with higher outlieriness correspond to low-quality anchor boxes. The gradient gain allocation strategy ensures that anchor boxes with both low and high outlieriness receive a small gradient gain, thereby emphasizing the regression for anchor boxes of ordinary quality. The expression for WIoU is shown in Equation (3).

$$L_{WIoUv3} = rR_{WIoU}L_{IoU} \quad (3)$$

Regarding the specific parameter expressions in WIoU, as shown in Equations (4) and (5):

$$r = \frac{\beta}{\delta\alpha^{\beta-\delta}} \quad (4)$$

$$R_{WIoU} = \exp\left(\frac{(x - x_{gt})^2 + (y - y_{gt})^2}{(W_g^2 + H_g^2)^*}\right) \quad (5)$$

In the equation, r is the non-monotonic focusing coefficient; R_{WIoU} is the enlargement coefficient for anchor boxes of ordinary quality; L_{IoU} is the loss value of the anchor box; W_g and H_g are the width and height of the minimum bounding box; (x, y) is the center point of the predicted box; (x_{gt}, y_{gt}) is the center point of the ground truth box; and $*$ denotes the separation of W_g and H_g from the computation graph.

According to the description of the outlier degree definition in the previous context, it is evident that anchor boxes with a small outlier degree correspond to high-quality anchor boxes, while anchor boxes with a large outlier degree correspond to low-quality anchor boxes. Because high-quality examples have a relatively large IoU, there is less need for further optimization. On the other hand, low-quality examples have a smaller IoU, making further optimization impractical. Therefore, we formulate a strategy for gradient gain allocation by controlling the hyperparameters α and δ , ensuring that both anchor boxes with a small outlier degree and those with a large outlier degree receive a small gradient gain. This directs the bounding box regression to focus on ordinary-quality anchor boxes, preventing significant harmful gradients from low-quality examples. The gradient gain allocation strategy is illustrated in Figure 12.

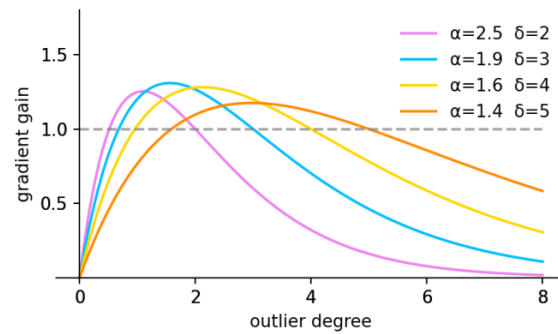


Figure 12. The mapping of outlier degree β and gradient gain r , which is controlled by the hyperparameters α , δ .

Because the component $\overline{L_{IoU}}$ within the outlier degree β is constantly changing, the criteria for classifying the quality of anchor boxes are also constantly changing. This enables the loss function to dynamically focus on ordinary-quality anchor boxes, continually strengthening the model's regression towards ordinary-quality anchor boxes, and thereby enhancing the detection performance of the network.

5. Data Augmentation

Due to the limited number of defective components in real production conditions, it is challenging to provide the required amount of data for deep learning models. A small dataset may lead to overfitting issues, resulting in missed detections [22]. Therefore, it is necessary to perform data augmentation on the collected bearing images. Traditional image augmentation methods typically involve simple affine transformations of the original images, which do not fundamentally generate new images [23]. Consequently, the challenge lies in how to generate entirely new defect samples based on the existing defect images.

To address these issues, this paper proposes a bearing image data augmentation method based on dynamic thresholds and generative adversarial networks (GANs) [24]. This approach leverages a small number of existing defect images for data augmentation, creating generated samples that closely resemble the original defect samples. This enhances the diversity of the original training data, addressing the problem of limited dataset size. The schematic diagram of this data augmentation method is shown in Figure 13.

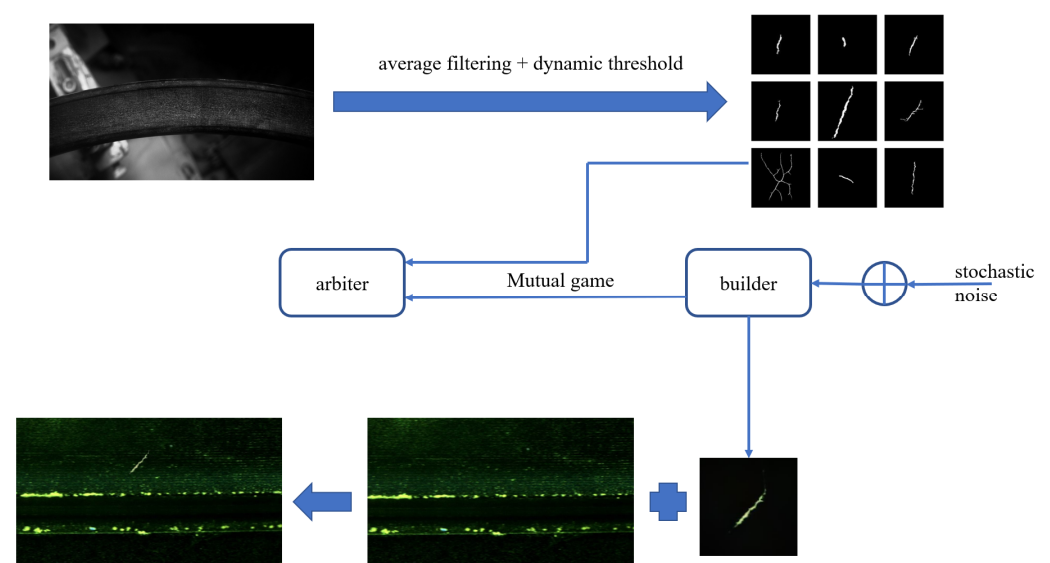


Figure 13. Data augmentation flowchart.

First, using the dynamic threshold algorithm and employing the mean filtering method, average local grayscale values of the original image are calculated. A thresh-

old for extracting defect grayscale values is then set. Defective regions are dynamically extracted based on the difference between the original grayscale image and the filtered image. The dynamic extraction formula is represented by Equation (6), separating the defective portions from the original image.

$$\begin{cases} g(o) \geq g(t) + \text{Offset} & (\text{if LightDark} = \text{light}) \\ g(o) \leq g(t) + \text{Offset} & (\text{if LightDark} = \text{dark}) \\ g(t) - \text{Offset} \leq g(o) \leq g(t) + \text{Offset} & (\text{if LightDark} = \text{equal}) \\ g(t) - \text{Offset} > g(o) \cup g(t) + \text{Offset} < g(o) & (\text{if LightDark} = \text{not_equal}) \end{cases} \quad (6)$$

In the equation, $g(o)$ represents the pixel values of the defect region to be extracted, $g(t)$ represents the pixel values of the grayscale image after mean filtering, Offset is the difference threshold between the original grayscale image and the grayscale image after mean filtering, and LightDark indicates the brightness or darkness of the defect area compared to its surroundings.

The generative adversarial network (GAN) is trained through the game between the generator and discriminator. This training enables the generator to produce generated samples that are extremely similar to real samples. Therefore, it is used to augment the defect samples extracted by the dynamic threshold, thereby increasing the diversity of defect samples. StyleGAN [25], based on ProGAN improvements, is a variant of generative adversarial network proposed by NVIDIA. The StyleGAN network structure consists of a generator and a discriminator. The generator network is composed of a mapping network and a generation network. The structure of StyleGAN is shown in Figure 14.

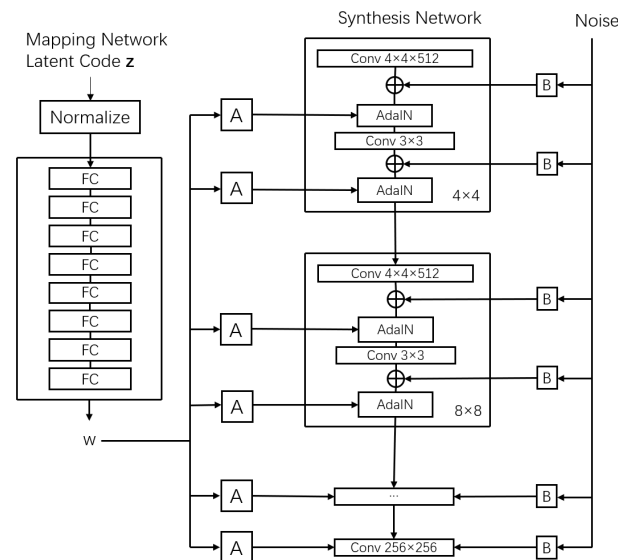


Figure 14. StyleGAN network architecture.

StyleGAN mainly decouples the latent space through a mapping network to explore the relationships between highly coupled features in the data. The mapping network transforms the latent feature z into an intermediate variable w through eight fully connected layers, thereby converting the distribution function of z into a non-uniform density distribution that better approximates the actual feature distribution. The mapping network model allows for the generated intermediate vector w to deviate from the training data distribution while reducing the coupling between features, enabling control over individual features without affecting their original interwoven characteristics.

As one of the state-of-the-art high-resolution image generation methods, StyleGAN has its shortcomings, including slow network-training speed and the presence of noticeable speckled artifacts in generated feature images. To address these issues, NVIDIA released StyleGAN2 [26]. Experimental findings revealed that the adaptive instance normalization

operation in the StyleGAN network structure causes feature images to contain speckled artifacts. StyleGAN2 redesigns the network architecture based on StyleGAN. Firstly, it decomposes the adaptive instance normalization operation into normalization and modulation steps. Secondly, it moves redundant biases and noise broadcast operations between different resolution layers. Finally, it uses demodulation instead of normalization. These adjustments in StyleGAN2 have been proven to effectively reduce the occurrence of speckled artifacts and accelerate network-training speed. The StyleGAN2 network structure is depicted in Figure 15.

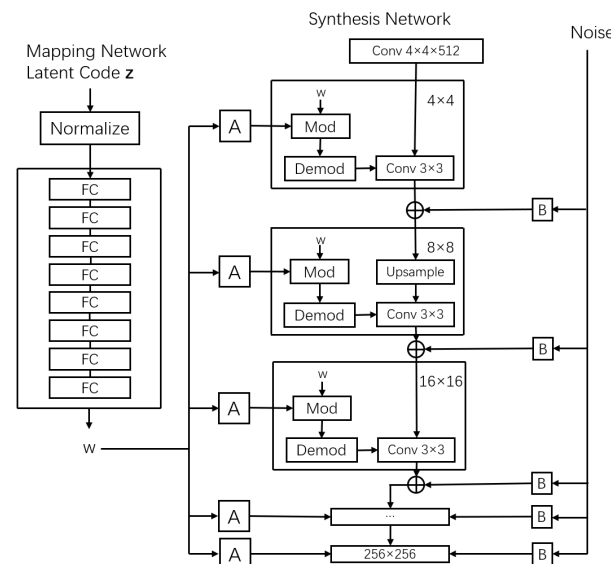


Figure 15. StyleGAN2 network architecture.

A comparison between Figures 14 and 15 shows that the StyleGAN2 generator incorporates improvements by using Mod and Demod modules to replace the AdaIN module in StyleGAN. Additionally, the overlay position of noise B has been adjusted from between the Conv module and AdaIN module to different resolution layers. These changes signify beneficial adjustments to the generator structure, making it more optimized and adaptable to the needs of image generation.

This study utilizes the generative adversarial network StyleGANv2 to generate entirely new defect images through the competition between the generator and discriminator. The generated defect images are then weighted and superimposed on defect-free bearing images, thereby constructing novel bearing defect images. This approach is employed to enhance the dataset with augmented samples of minor defects, enriching the training data and significantly improving the robustness of the detection model.

6. Experiment Results and Analysis

The experimental setup for this study was based on the Windows 11 operating system, with an Intel Core i7-11700 CPU, a 2 TB hard drive, NVIDIA RTX 3060 GPU with 12 GB of VRAM, and 16 GB of system memory.

6.1. Dataset Preprocessing

The data used in this study were collected from a bearing manufacturer in China, comprising a total of 1602 bearing images with a resolution of 4096×2160 . According to the morphology and origin of bearing surface defects, defects were categorized into three classes, denoted as a–c in Figure 16: inclusions, burn cracks, and scratches.

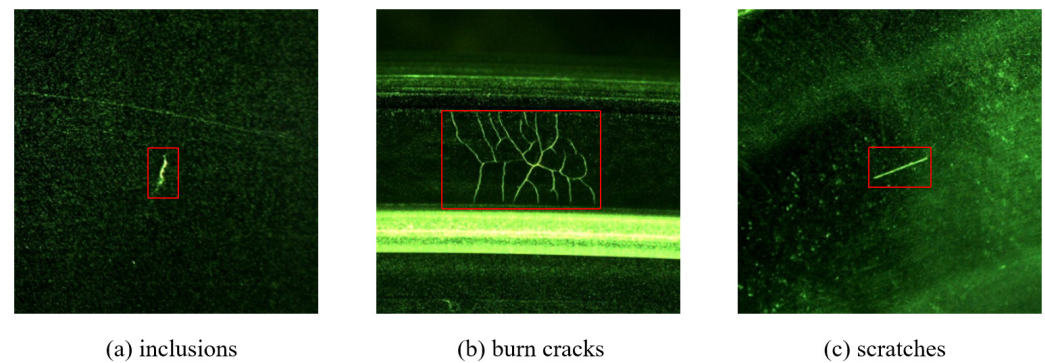


Figure 16. Categories of bearing defects.

To avoid losing small defect features during multiple convolutions and pooling operations, a block-wise training and sliding window detection approach was employed [27]. The images were cropped to a size of 2048×2160 , resulting in 3283 images remaining after removing defect-free images. The dataset was divided into training, validation, and test sets with a ratio of 8:1:1. Labeling was used to annotate the images, providing defect labels and completing the construction of the new dataset.

6.2. Evaluation Index

In order to assess the detection performance of the model, this paper employed precision (precision), recall (recall), average precision (AP), and mean average precision (mAP) as evaluation metrics [28].

6.3. Hyperparameter Setting

During the training process, the model was configured with a batch size of 32, a learning rate of 0.01, and the optimizer was set to SGD (Stochastic Gradient Descent). The number of training epochs was set to 100.

6.4. Compare the Improved Model with the Basic Model

The improved YOLOv5 model was compared with the baseline model in terms of mAP curve and training loss curve, as shown in Figures 17 and 18, respectively. Figure 17 illustrates the comparison of training loss curves, while Figure 18 presents the mAP curve comparison. It can be observed from Figures 17 and 18 that, in comparison to the baseline model, the improved model exhibited lower training loss values and higher detection accuracy.

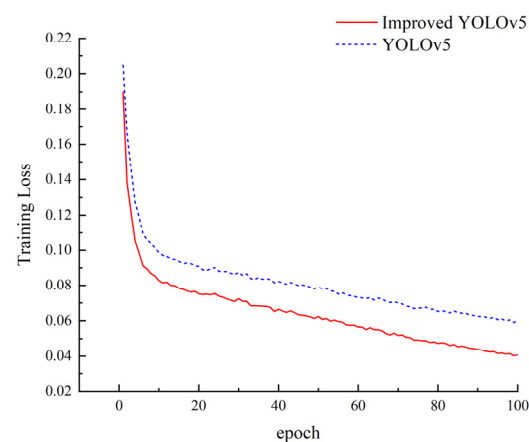


Figure 17. Training loss curve.

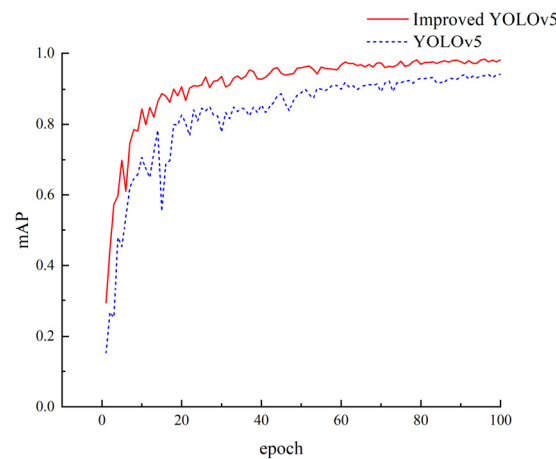


Figure 18. mAP comparison map.

6.5. Ablation Experiment

To investigate the impact of the aforementioned improvements in the model's detection accuracy, this paper conducted ablation experiments on the bearing image dataset using YOLOv5 as the base model. Firstly, the CA mechanism was added to the C3 module in the backbone network. Secondly, the ASFF adaptive spatial feature fusion structure was introduced to the neck network. Lastly, the original GIoU bounding box loss function in the network was replaced with WIoU, and the experimental results were recorded. The results are presented in Table 3:

Table 3. Experimental results.

Experiments	YOLOv5	CA	ASFF	WIoU	Precision/%	Recall/%	mAP/%
Experiments 1	✓				93.4	91.2	94.2
Experiments 2	✓	✓			95.4	93.5	96.1
Experiments 3	✓	✓	✓		96.9	95.8	97.3
Experiments 4	✓	✓	✓	✓	97.5	96.5	98.4

Analyzing the experimental results from Table 3, it can be observed that the original YOLOv5 network achieved a mAP of 94.2% on the bearing image dataset. After adding the CA mechanism to the C3 module of the backbone network, precision and recall increased by 2% and 2.3%, respectively, and mAP increased by 1.9%. The significant improvement in accuracy demonstrates that the CA mechanism can make the network focus more on the target regions in the images, thereby enhancing the model's detection performance.

Further, by adding the adaptive spatial feature fusion module (ASFF) to the end of the neck network, precision and recall increased by 1.5% and 2.3%, respectively, and mAP increased by 1.2%. The experimental results confirm that ASFF can eliminate feature inconsistencies at different levels, leading to improvements in all detection metrics.

Finally, replacing the bounding box loss function GIoU with WIoU resulted in precision and recall increasing by 0.6% and 0.7%, respectively, and mAP increased by 1.1%. The use of WIoU in this study, compared to GIoU, shows better fitting of the bounding boxes, thereby improving the model's detection performance.

In the end, the improved YOLOv5 achieved precision and recall of 97.5% and 96.5%, respectively, and a mAP of 98.4%. Compared to the original YOLOv5, these metrics improved by 4.1%, 5.3%, and 4.2%, respectively. Figure 19a shows the test results of the original YOLOv5, and Figure 19b shows the test results of the improved YOLOv5. The experiments demonstrate that, through the three improvements mentioned, the enhanced model exhibits effective improvement in detection accuracy, enhancing the precision of bearing defect detection and providing technical support for automated detection in bearing manufacturing.

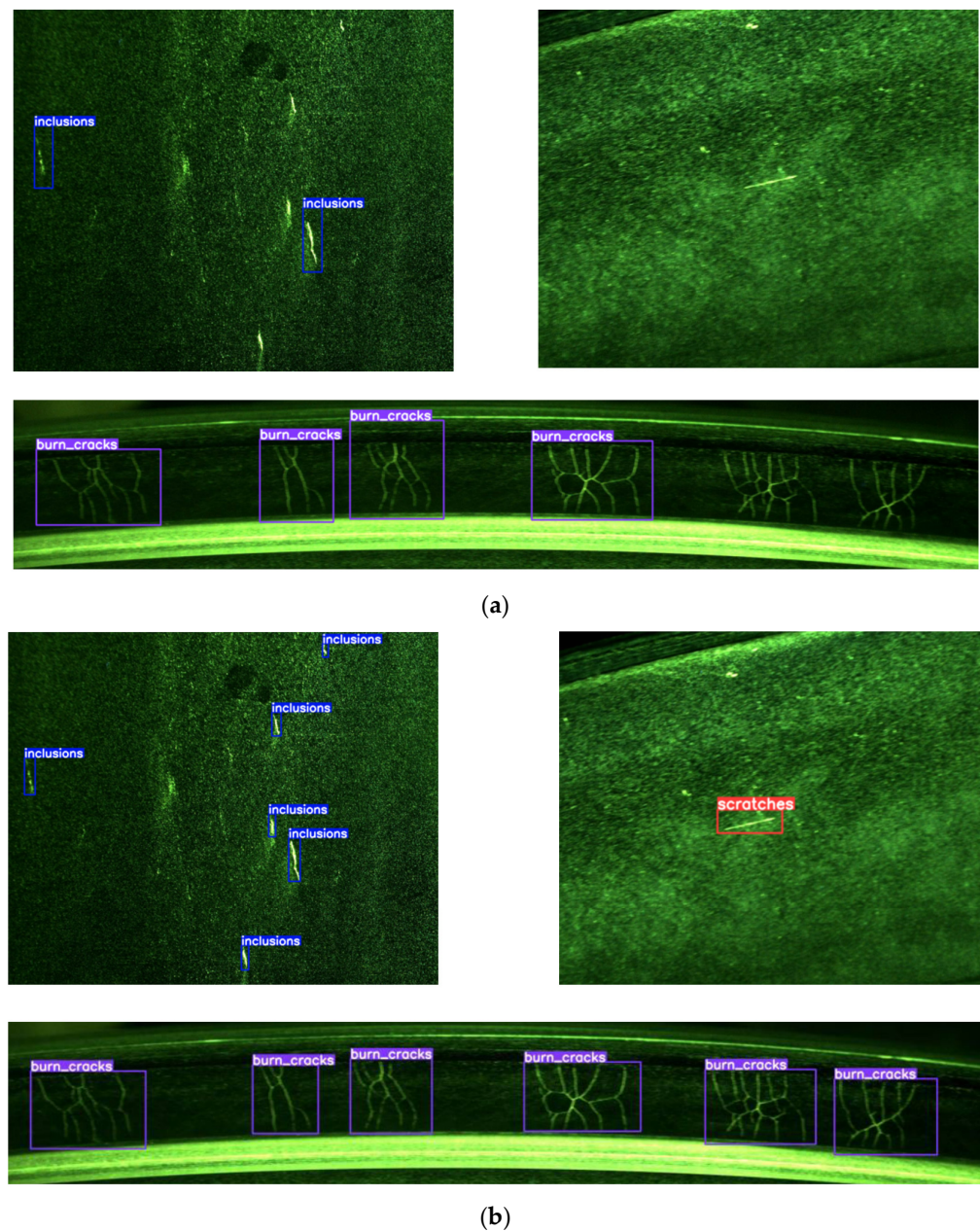


Figure 19. (a) Original YOLOv5 defect detection results. (b) Improved YOLOv5 defect detection results.

6.6. Comparative Experiment with Classical Algorithm

To validate the superiority of the improved YOLOv5 model in bearing defect detection, it was compared with three mainstream object detection models. The experimental results are presented in Table 4.

Table 4. Comparison experiment results of mainstream algorithms.

Model	Precision/%	Recall/%	mAP/%
Faster-RCNN	88.1	86.3	89.6
Cascade-RCNN	91.2	89.5	90.3
YOLOv3	86.7	83.1	85.3
ours	97.5	96.5	98.4

The results in the table indicate that the improved YOLOv5 model outperforms other mainstream algorithms in terms of accuracy, recall rate, and average precision. This demonstrates the superiority of this approach in surface defect detection on bearing components.

7. Conclusions

This paper proposes a method for bearing image defect detection that combines magnetic particle inspection with deep learning. Initially, the magnetic particle inspection method is employed to reveal defects on the bearing surface that are not easily observable through visual inspection, displaying them in the form of magnetic patterns. Subsequently, the diversity of defect samples is enhanced using a method based on dynamic thresholds and generative adversarial networks (GANs), thereby expanding the original defect sample dataset.

Targeted improvement of the original YOLOv5 network: firstly, integrating the CA mechanism into the original backbone network; secondly, adding the ASFF (adaptive spatial feature fusion) module to the neck network; and lastly, replacing the original GIoU loss function with the WIoU loss function. Through designed ablation experiments, this paper validates the effectiveness of the proposed data augmentation method and the three enhancements on the original network.

The proposed data augmentation method enhances the diversity of defect samples, with the improved network focusing more on the foreground regions in the images. This approach avoids feature inconsistencies across different levels and concentrates the bounding box loss function on regressing common-quality anchor boxes, thereby increasing the model's detection accuracy, reducing the likelihood of false negatives, and improving the production efficiency of bearings.

Author Contributions: Conceptualization, L.L. and Z.L.; methodology, L.L., Z.L. and H.Z.; software, Z.L.; validation, H.Z.; formal analysis, H.Z.; investigation, H.Z.; resources, L.L.; data curation, Z.L.; writing—original draft preparation, L.L.; writing—review and editing, L.L. and Z.L.; visualization, H.Z.; supervision, L.X.; project administration, L.X. and J.W. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The data presented in this study are available on request from the corresponding author.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Choudhury, A.; Tandon, N. Application of acoustic emission technique for the detection of defects in rolling element bearings. *Tribol. Int.* **2000**, *33*, 39–45. [\[CrossRef\]](#)
2. Betea, B.; Tomesc, L.; Dobra, P. Bearing Defects Diagnosis Based on Gaussian Filtering Shock Detection and Power Spectrum of Envelope. In Proceedings of the 17th International Conference System Theory, Control and Computing (ICSTCC), Sinaia, Romania, 11–13 October 2013; pp. 71–76.
3. Patil, M.S.; Mathew, J.; RajendraKumar, P.K. Bearing signature analysis as a medium for fault detection: A review. *J. Tribol. Trans. ASME* **2008**, *130*, 7. [\[CrossRef\]](#)
4. Holm-Hansen, B.T.; Gao, R.X.; Zhang, L. Customized wavelet for bearing defect detection. *J. Dyn. Syst. Meas. Control-Trans. ASME* **2004**, *126*, 740–745. [\[CrossRef\]](#)
5. Li, J.L.; Pan, H.P.; Li, J.F. ESD-YOLOv5: A Full-Surface Defect Detection Network for Bearing Collars. *Electronics* **2023**, *12*, 3446. [\[CrossRef\]](#)
6. Rimlyand, V.I.; Kazarbin, A.V.; Dobromyslov, M.B. Active ultrasonic nondestructive testing of rotating parts and bearings. *Res. Nondestruct. Eval.* **2004**, *15*, 19–29. [\[CrossRef\]](#)
7. Chen, Y.T.; Kang, Y.H.; Feng, B.; Li, Y.N.; Cai, X.; Wang, S.H. Automatic defect identification in magnetic particle testing using a digital model aided De-noising method. *Measurement* **2022**, *198*, 8. [\[CrossRef\]](#)

8. Liang, D.; Wang, D.C.; Chu, J.L.; Hu, K.; Xi, Y.L. Defect detection of bearing side face based on sample data augmentation and convolutional neural network. *J. Adv. Mech. Des. Syst. Manuf.* **2023**, *17*, 14. [\[CrossRef\]](#)
9. Liao, D.H.; Yin, M.S.; Luo, H.B.; Li, J.; Wu, N.X. Machine vision system based on a coupled image segmentation algorithm for surface-defect detection of a Si₃N₄ bearing roller. *J. Opt. Soc. Am. A Opt. Image Sci. Vis.* **2022**, *39*, 571–579. [\[CrossRef\]](#) [\[PubMed\]](#)
10. Wen, S.P.; Chen, Z.H.; Li, C.X. Vision-Based Surface Inspection System for Bearing Rollers Using Convolutional Neural Networks. *Appl. Sci.* **2018**, *8*, 19. [\[CrossRef\]](#)
11. Glazkov, Y.A. Analysis of Conditions of Detection of Indicator Patterns in Magnetic Powder Inspection. *Russ. J. Nondestr. Test.* **1993**, *29*, 880–887.
12. Shelikhov, G.S. Portable flaw detectors for magnetic-powder inspection of objects under operating conditions. *Russ. J. Nondestr. Test.* **2010**, *46*, 395–402. [\[CrossRef\]](#)
13. Shelikhov, G.S. Dependence of the detectability of flaws on the distribution of the magnetizing field in the vicinity of a flaw. *Russ. J. Nondestr. Test.* **2004**, *40*, 397–400. [\[CrossRef\]](#)
14. Redmon, J.; Divvala, S.; Girshick, R.; Farhadi, A. You Only Look Once: Unified, Real-Time Object Detection. In Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition, Las Vegas, NV, USA, 27–30 June 2016.
15. Redmon, J.; Farhadi, A. YOLO9000: Better, Faster, Stronger. In Proceedings of the 2017 IEEE Conference on Computer Vision and Pattern Recognition, Honolulu, HI, USA, 21–26 July 2017; pp. 6517–6525.
16. Redmon, J.; Farhadi, A. YOLOv3: An Incremental Improvement. *arXiv* **2018**, arXiv:1804.02767.
17. Hou, Q.; Zhou, D.; Feng, J. Coordinate Attention for Efficient Mobile Network Design. In Proceedings of the 2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition, Nashville, TN, USA, 20–25 June 2021.
18. Hu, J.; Shen, L.; Sun, G.; Albanie, S. Squeeze-and-Excitation Networks. In Proceedings of the 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition, Salt Lake City, UT, USA, 18–23 June 2018.
19. Liu, S.; Huang, D.; Wang, Y. Learning Spatial Fusion for Single-Shot Object Detection. *arXiv* **2019**, arXiv:1911.09516.
20. Rezaatofghi, H.; Tsoi, N.; Gwak, J.Y.; Sadeghian, A.; Savarese, S. Generalized Intersection Over Union: A Metric and a Loss for Bounding Box Regression. In Proceedings of the 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), Long Beach, CA, USA, 15–20 June 2019.
21. Tong, Z.; Chen, Y.; Xu, Z.; Yu, R. Wise-IoU: Bounding Box Regression Loss with Dynamic Focusing Mechanism. *arXiv* **2023**, arXiv:2301.10051.
22. Srivastava, N.; Hinton, G.; Krizhevsky, A.; Sutskever, I.; Salakhutdinov, R. Dropout: A Simple Way to Prevent Neural Networks from Overfitting. *J. Mach. Learn. Res.* **2014**, *15*, 1929–1958.
23. Chen, C. Insulator defect detection data enhancement method, involves extracting target region of interest characteristic, overlapping affine transformation to picture, and subjecting picture to contrast change, cutting and conventional data enhancement processes. Chinese Patent G06T-007/00 202215, 31 December 2021.
24. Goodfellow, I.; Pouget-Abadie, J.; Mirza, M.; Xu, B.; Warde-Farley, D.; Ozair, S.; Courville, A.; Bengio, Y. Generative adversarial networks. *Commun. ACM* **2020**, *63*, 139–144. [\[CrossRef\]](#)
25. Karras, T.; Laine, S.; Aila, T. A Style-Based Generator Architecture for Generative Adversarial Networks. In Proceedings of the 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), Long Beach, CA, USA, 15–20 June 2019.
26. Karras, T.; Laine, S.; Aittala, M.; Hellsten, J.; Aila, T. Analyzing and Improving the Image Quality of StyleGAN. In Proceedings of 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), Seattle, WA, USA, 13–19 June 2020.
27. Li, L.; Li, Q.; Liu, Z.Y.; Xue, L. Effective Fabric Defect Detection Model for High-Resolution Images. *Appl. Sci.* **2023**, *13*, 10500. [\[CrossRef\]](#)
28. Zheng, L.; Shen, L.; Tian, L.; Wang, S.; Tian, Q. Scalable Person Re-identification: A Benchmark. In Proceedings of the 2015 IEEE International Conference on Computer Vision (ICCV), Santiago, Chile, 7–13 December 2015.

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.