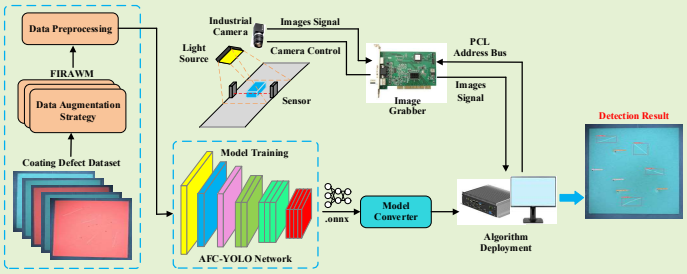
Citation: K. Tang, B. Zi, F. Xu, W. Zhu and K. Feng, "Coating Defect Detection Method Based on Data Augmentation and Network Optimization Design," in IEEE Sensors Journal, doi: 10.1109/JSEN.2023.3277979.

DOI: https://doi.org/10.1109/JSEN.2023.3277979

Coating Defect Detection Method Based on Data Augmentation and Network Optimization Design

Kai Tang, Bin Zi, Feng Xu, Weidong Zhu, and Kai Feng

***Abstract*—Coating defect detection is a critical aspect of ensuring product quality in the manufacturing process. However, due to the variety of coating defects and the complex detection background in actual production, detecting these defects can be challenging. To improve the accuracy and robustness of coating defect detection, a coating defect detection method based on data augmentation and network optimization design is proposed. First, a feature image random adaptive weighted mapping (FIRAWM) strategy is proposed, considering the prior accuracy, quantity and context information of each category. Then, several improvements are made to the YOLOv5 network. Specifically, to mitigate the aliasing effects and enhance feature richness during the feature fusion process, an additional detection layer is added, and the coordinate attention module and the adaptively spatial feature fusion (ASFF) module are introduced. Finally, ablation and comparison experiments are performed to demonstrate the effectiveness of the proposed method. The results show that the method achieves a 96.7 mAP_{50} with a processing speed of 61 FPS on the coating defect dataset, outperforming other popular detectors. Furthermore, the method is versatile and can be applied to detection tasks in various scenarios.**



***Index Terms*—Coating defect detection, data augmentation, object detection, network optimization design.**

I. Introduction

SPRAYING refers to covering the coating on the surface of an object. The main functions of these coatings have three aspects: 1) Protecting objects from being eroded by light, water and other media, and prolonging their service life; 2) Enhancing the color and luster of the object surface; 3) The surface of the object is sprayed with special patterns, which can achieve the effect of camouflage and so on. However, due to the factors such as the environment of the spraying workshop, the proportion of coating, the coating process and the improper operation of workers, inevitable defects may occur in the spraying process, so it is necessary to detect the coating defects.

Due to the randomness of coating defects, and the characteristics of coating defects such as multiple types, different shapes and sizes, limited total quantity, and uneven distribution and similar characteristics of some defects, it becomes difficult in the actual detection process. For coating defect detection, the traditional manual detection has some shortcomings, such as low efficiency, poor real-time performance, easy to be affected by subjective factors, and inability to meet the requirements of the modern mass production line for product quality and efficiency. Machine vision offers many advantages such as high efficiency, good stability, and superior detection accuracy. As a result, manual detection is gradually replaced by coating defect detection based on machine vision, which is becoming the mainstream method in the field of defect detection [1], [2]. According to different image processing and defect detection methods, it can be divided into traditional machine vision algorithms and machine vision algorithms based on deep learning. Rapid progress has been made in developing a variety of methods for detecting coating defects. Refs. [3], [4] and other scholars have studied coating defect detection methods based on traditional machine vision, mainly by preprocessing the image to alleviate the interference of background noise, and manually designing Gabor filters, Scale-invariant feature transform (SIFT) and other feature extraction operators to extract defect features according to the characteristics of defects. Finally, Support Vector Machine (SVM), K-Nearest-Neighbor (KNN), Random Forest (RF) and other classifiers are used to classify the defect features. Although traditional machine vision can detect coating defects to a certain extent, its feature extraction operator design is complex and susceptible to interference from external factors. The robustness of traditional machine vision is poor in dynamically changing complex

This work was supported by the National Natural Science Foundation of China (NSFC) under Grant 51925502. *(Corresponding author: Bin Zi.)*

Kai Tang, Bin Zi, and Feng Xu are with the School of Mechanical Engineering, Hefei University of Technology, Hefei 230009, China (e-mail: kaitang941025@163.com; binzi.cumt@163.com; xufeng_hfut@163.com).

Weidong Zhu is with the Department of Mechanical Engineering, University of Maryland, Baltimore, MD 21250, USA (e-mail: wzhu@umbc.edu).

Kai Feng is with the Department of Electromechanical Engineering, Faculty of Science and Technology, University of Macau, Macau 999078, China (e-mail: yc07491@um.edu.mo).

scenes, making it suitable only for simple coating detection in some specific environments. The method based on deep learning can use massive data to extract rich features of the same target to complete network training, and the accuracy and robustness of the algorithm are better. Therefore, the algorithms based on deep learning have been gradually applied to coating defect detection.

For a deep learning algorithm, the performance of a network depends not only on its structure but also on the size, feature diversity and distribution of its training dataset [5], [6]. However, it is expensive and difficult to collect a large number of defect samples in natural scenes, and due to the inherent characteristics of coating defects, some advanced data augmentation algorithms [7]–[9] are not suitable. Therefore, some traditional data augmentation methods, such as flipping, rotating, scaling, and so on, are often used in coating defect detection. Some researchers have paid more attention to optimizing and designing the network architecture. In [10], Chen *et al*. proposed a method for detecting surface coating defects of aluminum alloy based on deep learning, and improved Faster R-CNN as follows: 1) The Feature Pyramid Network (FPN) is used to realize the fusion of low-level location information and high-level semantic information to improve the detection ability of the network for small objects; 2) Add Deformable Convolutional Networks (DCN); 3) Optimize the training process with data augmentation techniques. Experimental results show that the method has excellent performance compared with other methods. Wen *et al*. proposed a new method for semiconductor surface coating defect detection based on a deep Convolution Neural Network in [11]. First, a novel FPN with Dilated Convolution is used to extract features, and then the generated feature map is fed into the Region Proposal Network (RPN) to generate candidate regions. Finally, the candidate regions are used as input to the Deep Multibranches Neural Network (DMBNN) to accurately classify and segment coating defects. Aiming at the problem that existing deep learning algorithms are insufficient in detecting some small and complex defect objects, Zhao *et al*. carried out a series of improvement measures on Faster R-CNN, and reconstructed the structure of Faster R-CNN network by using multi-scale feature fusion and Deformable Convolutional Networks in [12]. These improved the accuracy of detecting small defects on the steel surface and provide a reference for automatic defect detection of the coating. To balance the speed and accuracy of the coating defect detection algorithm, some researchers also studied the one-stage algorithm [13]. Many studies have proven that the coating defect detection method based on deep learning can achieve better results than the traditional machine vision. With the development of the deep learning algorithm and the improvement of hardware performance, the deep learning algorithm can be deployed on edge devices for real-time detection.

A coating defect detection method based on data augmentation and network optimization design is proposed in this paper. First, a coating defect dataset is constructed according to the causes of some coating defects. Then, a novel data augmentation strategy is proposed to improve the quality of the coating defect dataset and to compensate for the inadequacy of existing data augmentation methods in the coating defect dataset. Finally, a deep learning network for coating defect detection is designed by improving the YOLOv5 network. To verify the effectiveness of the proposed method, ablation and comparison experiments are performed on the coating defect dataset and other publicly available datasets. The results demonstrated that the proposed method can not only improve the accuracy of coating defect detection, but also achieve good performance in other detection tasks. In summary, the contributions of this paper are as follows.

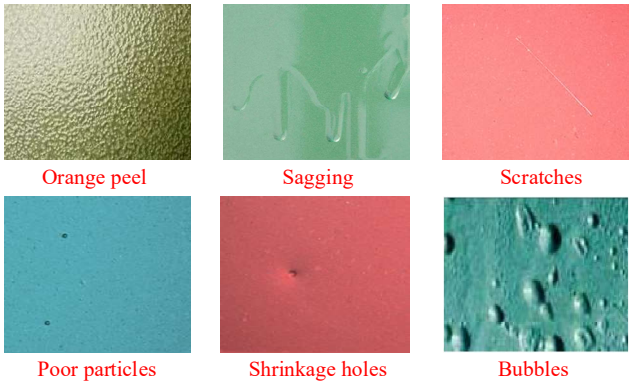


Fig. 1. Common coating defect categories.

1) A coating defect dataset with several typical defects is collected by controlling the proportion of the coating and spraying process.
2) For the coating defect dataset, a FIRAWM strategy based on the prior accuracy, quantity, and context information of each category is proposed, which improves the quality of the coating defect dataset and the detection performance of the network. This strategy is also effective for other detection tasks.
3) By optimizing the YOLOv5 network, a novel network for coating defect detection is designed. The network improves the accuracy and robustness of coating defect detection and is superior to other popular detectors.

The remainder of this paper is organized as follows. In Section II, the details and implementation of the proposed coating defect detection method are described, including the design of a novel data augmentation strategy and the improvement of the network architecture. In Section III, the experimental results and comparison with various advanced methods are presented. Finally, this study is concluded in Section IV.

II. METHOD

For the task of coating defect detection, the main difficulties include: 1) The amount of actual coating defect sample data is small and difficult to collect; 2) Exposed to the natural environment, light is easily reflected, and the background is complex and varied; 3) The morphological differences between the samples are large, and the number of each category is seriously unbalanced; 4) There are a large number of small objects, and they are densely clustered. These difficulties seriously hinder the further improvement of network detection accuracy and generalization. Therefore, this paper presents a unified solution for coating defect detection tasks based on the deep learning method. Since the accuracy and efficiency of coating defect detection based on deep learning mainly depend on the quality of datasets and the performance of networks, the study is mainly carried out from three aspects: coating defect dataset,

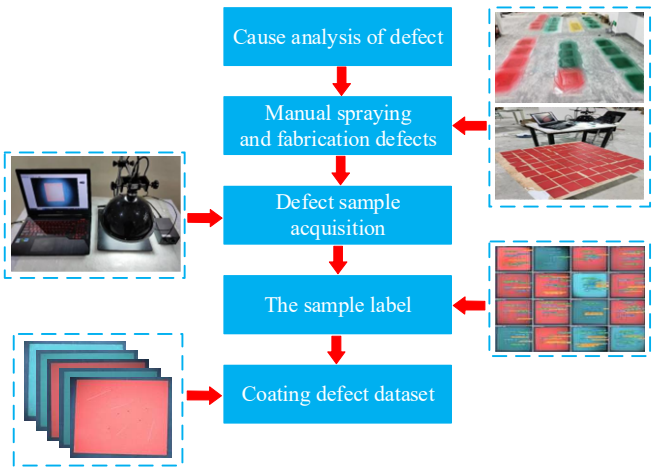


Fig. 2. Manufacturing process of coating defect dataset.

data augmentation strategy, and network optimization design. The specific contents are as follows.

A. Construction of the Coating Defect Dataset

Coating defect is one of the common appearance defects in industrial production, which plays an important role in the appearance and performance of products. Coating defects may be caused by equipment wear, the proportion of coating, oil pollution, the production environment, and improper operation of workers. Common coating defects include orange peel, sagging, scratches, poor particles, shrinkage holes, and bubbles, as shown in Fig. 1.

Some coating defects have obvious characteristics and are relatively easy to detect for the deep learning algorithm, this paper selects scratches, poor particles, and shrinkage holes as the detection objects, which are common but difficult to detect. Fig. 2 shows the production process of the entire coating defect dataset. First, a large amount of relevant coating literature is reviewed, such as spraying technology, and visited some spraying factories for investigation. Then, the causes of coating defects are analyzed, and produced coating defect samples by controlling the industrial environment, the spraying process, and other factors. Finally, the defect samples are collected using the image acquisition equipment, and the coating defect dataset is completed by manually labeling the samples. 1115 defect images are collected from the actual scene, and each image contains a large number of different defect types. According to the statistics, there are 3257 scratches, 3567 poor particles, and 2231 shrinkage holes.

B. Design of the Data Augmentation Strategy

Data augmentation is one of the core tasks of deep learning. It mainly expands and enriches the original dataset through various data transformation methods. The purpose is to provide more feature diversity in the network training process to reduce over-fitting and improve network generalization. Existing data augmentation methods can be divided into three categories [14], [15]: model-free, model-based, and optimizing policy-based. Model-free methods use image processing and mixed images [16]–[18], etc. Model-based methods need to generate new images according to the training model [19]–[21], etc. Optimizing policy-based methods aim to find the optimal operation or its combination in the parameter search space [22], etc. For practical tasks, the model-based and optimizing policy-based methods will introduce a lot of computational overhead and has high hardware requirements, so it is not suitable for coating defect detection tasks. In model-free data augmentation methods, CutMix [7], GridMask [8], and MixUp [9] are currently advanced data augmentation methods. However, due to the inherent characteristics (some of the defect features are tiny and slender) of coating defects, when using these data augmentation methods, some tiny coating defects may be occluded, and may also result in partial scratch features and bounding boxes regression precision loss. Therefore, they are not suitable for coating defect detection. To improve the quality and diversity of the dataset, and to compensate for the shortcomings of existing data augmentation methods in the coating defect dataset, operations combining various image transformations are considered to achieve the following goals: 1) Reduce the correlation between images before and after data augmentation to provide more feature diversity and reduce over-fitting during network training; 2) The number of different defect types can be dynamically adjusted to solve the problem of small sample size and uneven distribution of the coating defect dataset; 3) The merged image does not lose features and bounding boxes regression accuracy. Inspired by [23]–[25], a novel data augmentation strategy is proposed, referred to as FIRAWM. Fig. 3 shows the working flowchart of the algorithm, and its specific content is as follows.

1) Make Feature Image Database: First the prepared coating defect dataset is divided into pre-training set and test set by 7:3, then the categories of the pre-training set are counted and folders corresponding to each category are created. Finally, the feature images in the pre-training set are clipped according to the annotation information, and stored in the corresponding folders to build the coating defect feature image database.

2) Adaptive Random Sampling: In practical tasks, it is common for some datasets to have class-imbalance, and how to alleviate the problem of class-imbalance is a challenge. The existing data augmentation usually adopts the expansion of the number of classes to achieve the purpose of balancing the number of classes. These methods usually have certain randomness and uncertainty, and the recognition accuracy of some categories with a small amount of data is not necessarily low. The recognition accuracy of the network to the samples is related to the number of samples, feature complexity, background environment, and other factors, while the single consideration of balancing the number of classes in data augmentation is suboptimal. Compared with other data augmentation methods, we take the prior accuracy, quantity, and context information of each category in the image as one of the guiding principles for data augmentation.

The current pre-training set for each category is characterized by the small amount of data, the imbalanced proportion and the dense aggregation of small objects. Therefore, to compensate for the shortcomings of existing data augmentation methods, the sample distribution in the actual scene is simulated, considering the prior accuracy, quantity, and shape size of each category comprehensively. The specific implementation procedure is described below.

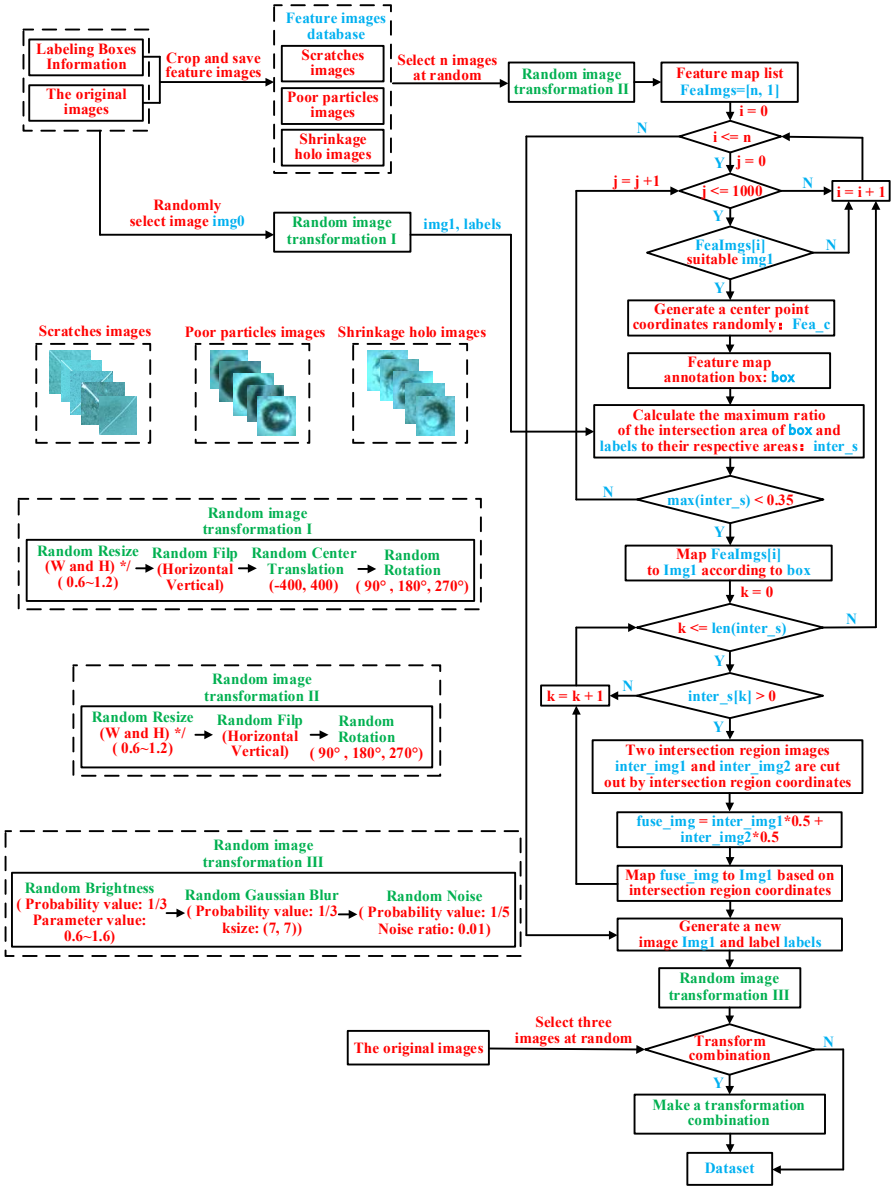


Fig. 3. Operating principle of FIRAWM algorithm.

(1). Step 1: [24] and [25] have confirmed that for the visual system, some objects that are difficult to detect can be identified by their context information, and it is useful for the system to consider the context information around the object during detection. Assuming that *img0* is the original image that needs to be augmented, where the background environment and the features (feature category and feature quantity) need to be analyzed, and filter some feature images that should not exist in a particular environment through the existing categories in *img0* (for example, in the scene of the living room of a house, there should be no airplanes and trains in real life), so that the content of the data augmentation conforms to the actual situation.

(2). Step 2: The YOLOv5 network is used as the baseline, taking the pre-training set as the training data, and the test set for testing, and the recognition accuracy of each coating defect category is obtained under the mAP_{50} metrics.

(3). Step 3: The categories in *img0* are denoted as $c = [c_1, c_2, \ldots, c_n]$ and the pre-training accuracy of each category is $P = [p_i]_{i=1}^{n}$, the number of *FeaImgs* of each corresponding feature images randomly selected from the feature image database is determined as

$$s_i = \frac{v \times (1 - p_i)}{\sum_{i=1}^{n} (1 - p_i)} \times m \tag{1}$$

where n is the total number of categories, v is the hyperparameter, m is the total number of feature images contained in *img0*, and the number of categories can be dynamically adjusted to enrich the number of specific categories.

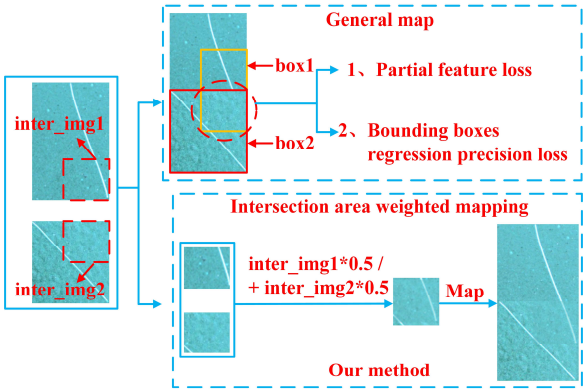


Fig. 4. Comparison of effect before and after improvement.

Algorithm 1 Weighted Mapping Strategy for Intersection Regions.

Input: *img0* and corresponding *labels,* feature images *FeaImg* , and mapping center point coordinates *Fea_c*

boxes ← Calculate the annotation boxes position of *FeaImg* in *imgs0*;
inters ← Calculate the ratio of the intersection area of *boxes* and *labels* to their respective areas;
if *max*(*inters*) > 0.35 **then**
  pass
else
  Map *FeaImg* to *img0*; // According to the *boxes*
  labels ← Add *boxes* to *labels*;
  for *s* ∈ *interns* **do**
  if *s* > 0 **then**
    inter_img1, *inter_img2* ← Crop out the respective images of the two intersection areas;
    fuse_img ← *inter_img1* × 0.5 + *inter_img2* × 0.5;
    Map *fuse_img* to *img1*;

Output: *img1*, *labels*

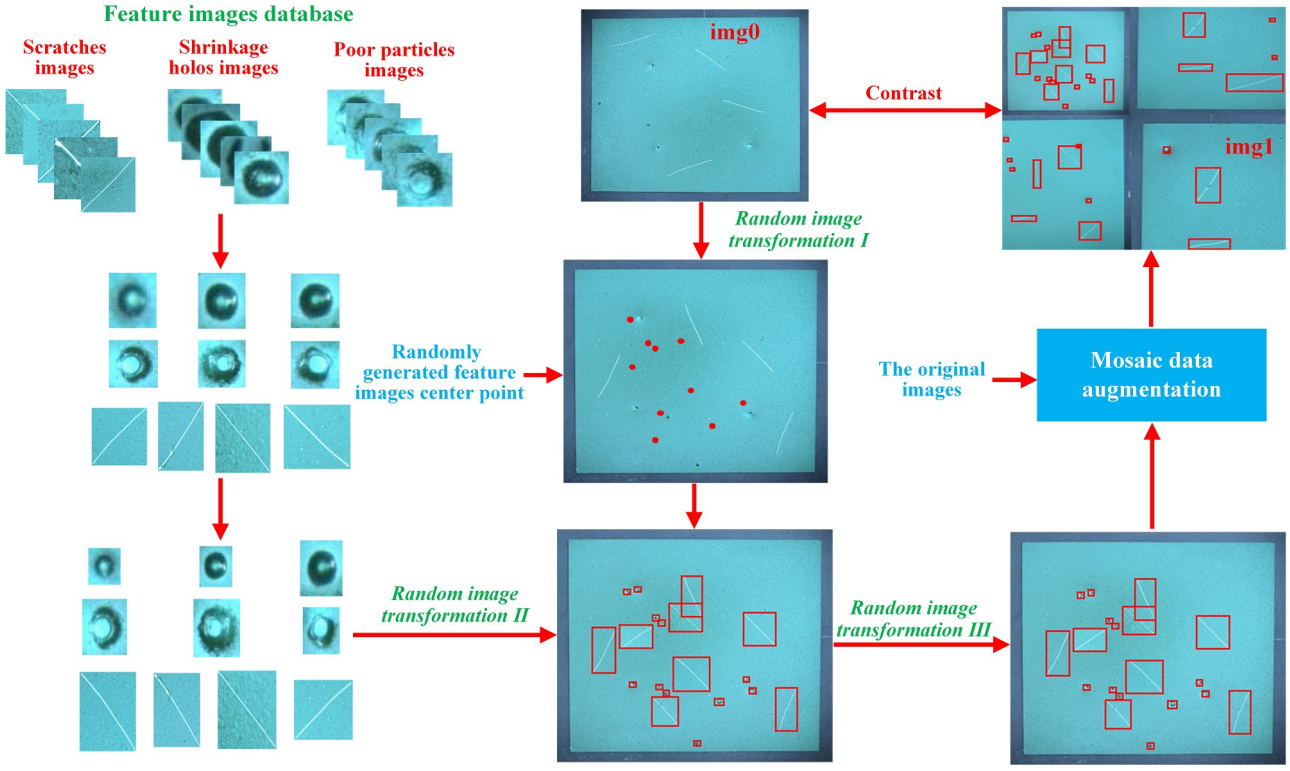


Fig. 5. Implementation process of FIRAWM algorithm.

(4). Step 4: Within a certain range of *img0*, randomly select the mapping center point *Fea_cs* of the feature images, perform *Random image transformation I* transformation on *img0* to obtain *img1*, then perform *Random image transformation II* transformation on *FeaImgs*, and the feature images *FeaImgs* are screened by the size of *Fea_cs* (height and width of the image) and *FeaImgs*, and the final effective feature images *FeaImgs* and the mapping location *Fea_cs* are obtained. The details are shown in Fig. 3.

3) Weighted Mapping Strategy for Intersection Regions: To better map and fuse the feature images *FeaImgs* and *img1*, and to solve the problems of partial features and bounding boxes regression precision loss in existing data augmentation methods, a weighted mapping strategy for the intersection regions of feature images is proposed. The complete process of the algorithm is shown in Algorithm 1, and the comparison of the proposed weighted mapping strategy for intersection regions with other methods is presented in Fig. 4. As shown in Fig. 4, compared with other methods of directly mapping feature images, the steps of the proposed method are as follows: first extract the intersection regions and positions of feature images, then multiply the extracted intersection regions by weight factors respectively for refusion, and finally remap the intersection regions after weighted mapping fusion back to the

original image according to the position information. The proposed method is more effective for coating defect data augmentation.

4) Transform Combination: Random image transformation III transform is performed on the mapped *img1* (the detailed operation is shown in Fig. 3). Then, there is a 1/3 probability that *img1* will be spliced with other images randomly selected from the pre-training set. The mosaic algorithm [18] is referenced. First, an empty image is constructed with twice the size of the network input. Then, the coordinate of the splice point is randomly selected in the empty image, divide the empty image into four blocks by the x-axis and y-axis of the coordinates of the splice point, adjust the size of the images to be spliced, and randomly map them onto the four empty images. Finally, the mapped image is scaled and offset, and the label information is adjusted to obtain a new image *img1* and label *labels*. The entire transformation process of the FIRAWM algorithm is illustrated in Fig. 5.

5) Data Analysis: After processing by the FIRAWM algorithm, 780 images in the pre-training set are expanded to 1400 images. Statistical analysis is performed on the expanded coating defect dataset, and the results are shown in Fig. 6.

In summary, compared with the existing data augmentation methods, the proposed FIRAWM algorithm has the following advantages: 1) It can be seen from Fig. 5 that the image after the FIRAWM algorithm has little correlation with the original image; 2) According to the actual needs, the hyperparameters and functions can be dynamically adjusted to enrich the number of feature images and avoid the phenomenon of network over-fitting; 3) Compared with other copy-paste data augmentation methods, the proposed method solves the problems such as class-imbalance, partial features and bounding boxes regression precision loss in the existing data augmentation methods, and achieves a better feature fusion effect. Its experimental verification is shown in the experimental section below.

C. Optimization Design of the network

In recent years, the methods based on deep learning have been introduced to defect detection tasks and have achieved satisfactory results. The YOLOv5 network has been widely used in various object detection tasks due to its fast recognition speed, high detection accuracy, and easy deployment [26], [27]. However, for the complex task of coating defect detection, it is difficult to achieve satisfactory results with a simple application. To compensate for the deficiency of the YOLOv5 network in detecting some coating defect features and to further improve the accuracy and robustness of the coating defect detection system, this paper improves the YOLOv5 network and proposes a coating defect detection network, referred to as AFC-YOLO.

The YOLOv5 network is a universal and effective object detector of the YOLO series. Its network architecture is mainly divided into three parts: backbone, neck, and head. The backbone is primarily composed of CSP and CBS modules, and the feature information in the image is extracted through the CSPDarknet53 network [18]. The neck adopts the Path Aggregation Network (PANet) [28] to fuse the features of different feature layers extracted from the backbone so that the features learned by the network have more diverse and richer semantic expression ability. The head mainly predicts the category and position of the object for the output feature layer in the neck. The YOLOv5 network can be divided into four parameter configurations from small to large: YOLOv5s, YOLOv5m, YOLOv5l and YOLOv5x. Considering the algorithm deployment and the requirement of detection speed, this paper improves the YOLOv5s according to the characteristics of coating defects, such as large morphological differences among defects, many small objects, dense aggregation, and so on. The improved network structure is shown in Fig. 7, and the details are as follows.

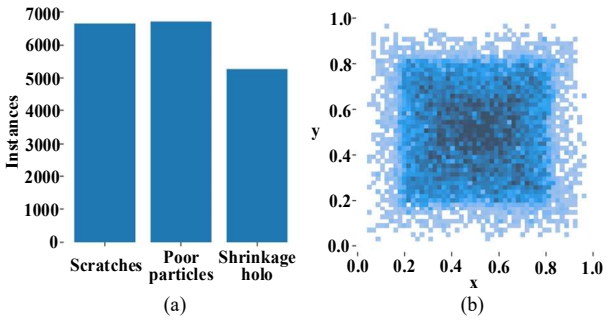


Fig. 6. Visual analysis of the coating defect dataset. (a) Quantity statistics of different defect categories, and (b) location distribution of defect features in the images.

1) Add an Additional Detection Layer: By examining the coating defect dataset, it is found that it contains many extremely small objects that are densely clustered together. Due to the characteristics of the network structure, the detailed features of small objects are lost after being processed by each layer of the convolutional backbone network. In the existing research on improving the detection ability of small objects, some researchers [29]-[31] have proved that the high-resolution feature map at the bottom of the network contains more detailed features, which is beneficial for the detection of small objects. Therefore, an additional high-resolution detection layer is added to preserve more detailed features of small objects and mitigate the negative effects caused by violent changes in object scale. As shown in Fig. 7, in order to improve the ability of the network to detect small objects, the neck and head structures of the YOLOv5 network are improved. Operations such as convolution and upsampling are performed to extend the neck, and the generated 160×160 high-resolution feature map is fused with CSP1_1 in the backbone, so that the fused feature map contains richly detailed features and semantic information, and serves as an additional detection layer. The Anchor size of the add additional detection layer is set to [5, 7; 9, 11; 8, 12] by the clustering algorithm, which makes the network more sensitive to microcoating defects. The final output sizes of feature map for the four detection heads are 20×20×24, 40×40×24, 80×80×24 and 160×160×24, respectively. The channel size includes the predicted bounding box information: defect category, bounding box coordinates and confidence value. Different scale detection layers can reduce the adverse impact caused by large scale variance of coating defects. Adding an additional detection layer increases the computational and memory cost. However, the performance of small object detection is greatly improved.

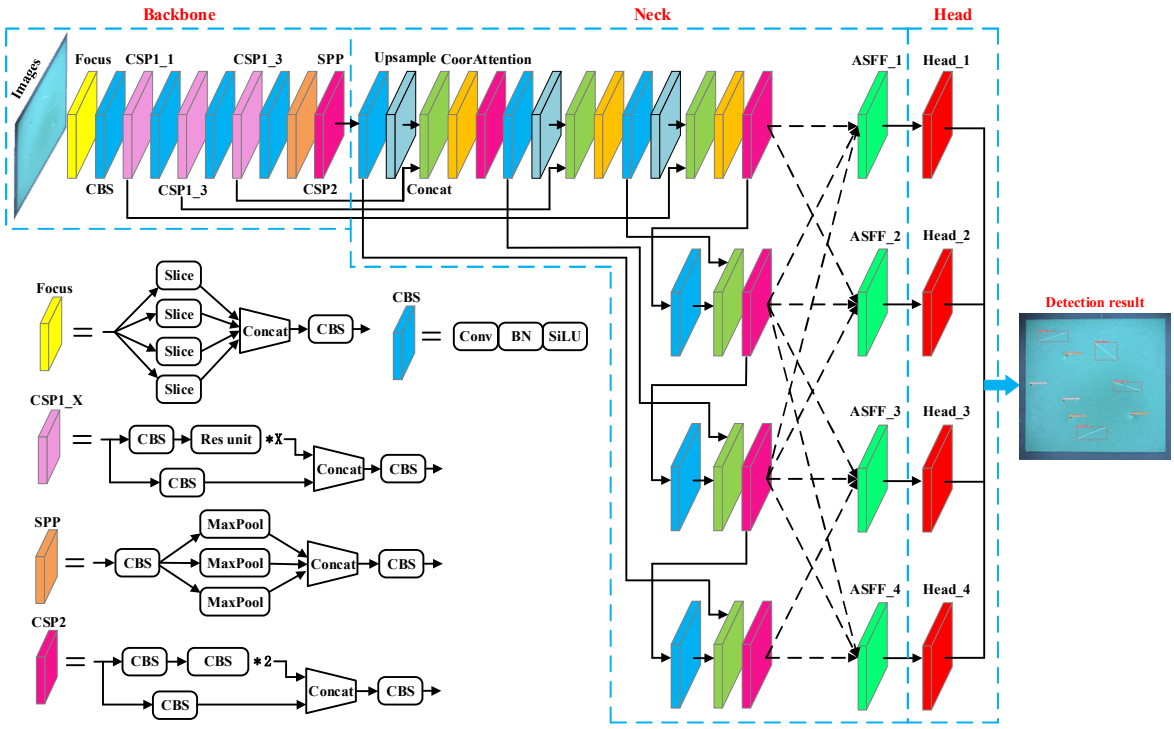


Fig. 7. Structure of the AFC-YOLO network.

2) Introduce the Coordinate Attention Module: Cross-scale fusion is widely used to improve the performance of object detection networks. However, due to the semantic differences between feature layers of different scales, the process of cross-scale fusion between different feature layers will cause the aliasing effect, and the fused feature map will contain a large amount of background noise and redundant features, which interfere with the recognition and localization tasks. To deal with the negative effects of feature fusion, researchers [32]-[36] have designed various types of attention mechanisms to make the network pay more attention to extracting useful features by establishing the spatial and channel interdependence of different features. Many studies have shown that the appropriate addition of attention mechanism can significantly improve the performance of the network.

The coordinate attention module [32] is an advanced attention mechanism. It decomposes channel attention into two one-dimensional feature coding processes. It captures the long-range dependency along one spatial direction while maintaining accurate positional information along the other spatial direction. Its structure is shown in Fig. 8. In particular, the coordinate attention module uses two spatial extents of the average pooling kernel (1, W) and (H, 1) to encode each channel along the horizontal and vertical coordinates, respectively, when given the input X. Therefore, the output of the c-th channel of width w can be determined as

$$z_c^w(w) = \frac{1}{H} \sum_{0 \le j < H} x_c(j, w) \tag{2}$$

Likewise, the output of the c-th channel of height h can be determined as

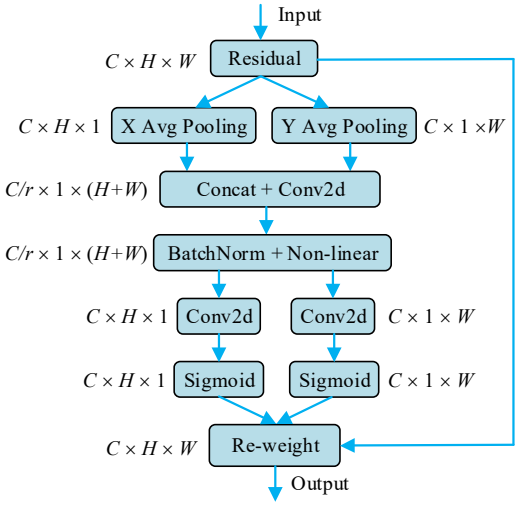


Fig. 8. Structure of the coordinate attention module.

$$z_c^h(h) = \frac{1}{W} \sum_{0 \le i < W} x_c(h, i) \tag{3}$$

Equations (2) and (3) aggregate features along two spatial directions, resulting in two directional feature maps that not only achieve a global receptive field but also encode precise position information. Then, the feature maps of the two spatial directions are concatenated and sent to the shared 1 × 1 convolution transform function F_1, resulting in

$$\mathrm{f} = \delta\left(F_1\left(\left[\mathrm{z}^h, \mathrm{z}^w\right]\right)\right) \tag{4}$$

Where $\mathrm{f} \in \mathbb{R}^{C/r\times(H+W)}$ is the intermediate feature map that encodes horizontal and vertical spatial information, δ is a non-linear activation function and r is the scaling ratio for module size control. Then, the coordinate attention module splits f into two independent tensors $\mathrm{f}^w \in \mathbb{R}^{C/r\times W}$ and $\mathrm{f}^h \in \mathbb{R}^{C/r\times H}$ along the spatial dimension. In addition, two 1×1 convolution transformations and Sigmoid activation functions are used to transform f^h and f^w into tensors with the same number of channels as input X.

Because the coordinate attention module aggregates features along two spatial directions and encodes location information into channel attention, its performance is superior to other attention modules, and there is almost no computational cost. Therefore, the coordinate attention module is integrated into the YOLOv5 network. As shown in Fig. 7, to mitigate the negative impact of the aliasing effect caused by the fusion of feature maps in the backbone and neck, this paper decides to insert a coordinate attention module after the connection of the backbone and neck features. A properly integrated lightweight coordinate attention module in the YOLOv5 network is beneficial for the network to locate objects of interest more accurately, thus achieving a balance between detection efficiency and performance.

3) Introduce Adaptively Spatial Feature Fusion (ASFF) Module: In the existing one-stage object detection algorithms, to deal with the shortcomings of multi-scale changes in object detection, various feature fusion methods (such as FPN, PANet, BiFPN [37], etc.) are usually employed to fuse different resolution and semantic information between different feature levels to improve the feature richness. However, due to the structural characteristics of the YOLO algorithm. YOLO will perform detection between feature levels of different scales. The low-level features are suitable for detecting small objects, and the high-level features are suitable for detecting large objects. To fully utilize features and semantic information of different scales between different detection layers, [38] proposed the Adaptively Spatial Feature Fusion (ASFF) module to perform an adaptive weighted fusion of features between different scales, and achieved remarkable performance in combination with YOLOv3 [39]. Since the network structure of YOLOv5 is similar to that of YOLOv3, we have partially modified the ASFF module to fit into our improved YOLOv5 network.

Specifically, the original ASFF module performs an adaptive weighted fusion of three detection layers. Since the improved YOLOv5 network has four detection layers, the current output layer is fused with the two nearest feature layers, taking into account the real-time detection speed and the effect of redundant features. This corresponds to the number of fusion layers of the original ASFF module, as shown in Fig. 7. For the feature layer ASFF_l to be fused, the features of the resolution at ASFF_l ($l \in \{1, 2, 3, 4\}$) are denoted as x^l. For upsampling, to ensure consistency in the shape of the feature, the x^s is resized at the other ASFF_s ($s \neq l$, $s \in \{1, 2, 3, 4\}$) to match the shape of x^l, and then 1×1 convolution is used to compress the channel number of the feature to the same as ASFF_l. Among them, the 2-ratio or 4-ratio upsampling process is completed by interpolation. For the 1/2-ratio downsampling, we also use a 3×3 convolution with a stride of 2 to change the number of channels and the resolution at the same time. For the 1/4-ratio downsampling, we do not use the max pooling layer, but use two 3×3 convolutions with a stride of 2 to preserve more detailed features.

During the adaptive feature fusion phase, we denote $\mathrm{x}_{ij}^{s\to l}$ as the feature vector located at position (i, j) on the feature maps that have been resized from ASFF_s to ASFF_l. The corresponding feature fusion process of ASFF_l can be determined as

$$\mathrm{y}_{ij}^{l} = \alpha_{ij}^{l} \cdot \mathrm{x}_{ij}^{1\to l} + \beta_{ij}^{l} \cdot \mathrm{x}_{ij}^{2\to l} + \gamma_{ij}^{l} \cdot \mathrm{x}_{ij}^{3\to l} + \lambda_{ij}^{l} \cdot \mathrm{x}_{ij}^{4\to l} \quad (5)$$

Where y_{ij}^{l} indicates that the vector located at position (i, j) in the output feature maps y^l between channels. α_{ij}^{l}, β_{ij}^{l}, γ_{ij}^{l} and λ_{ij}^{l} are the spatial importance weights of four feature maps at different levels to ASFF_l, which are defined using the Softmax function, and can be adaptively learned by the network. The farthest level from ASFF_l in α_{ij}^{l}, β_{ij}^{l}, γ_{ij}^{l} and λ_{ij}^{l} is 0, $\alpha_{ij}^{l} + \beta_{ij}^{l} + \gamma_{ij}^{l} + \lambda_{ij}^{l} = 1$ and $\alpha_{ij}^{l}, \beta_{ij}^{l}, \gamma_{ij}^{l}, \lambda_{ij}^{l} \in [0, 1]$.

By adaptively aggregating feature levels of different scales in the improved YOLOv5, the detailed features between different feature layers and the semantic information of different scales are fully utilized, and the output {y1, y2, y3, y4} is used as the input of the detection layer, thus improving the detection ability of the network for defects of different scales.

III. Experiment

In this section, this paper first presents the experimental setup, including implementation details and evaluation metrics, then compare the proposed data augmentation strategy with other popular methods, and uses ablation studies to prove the effectiveness of the AFC-YOLO network. Finally, to verify the versatility of the proposed method, a quantitative comparison with other advanced networks is performed on the dataset of different detection tasks, and industrial online detection experiments are performed on physical prototypes.

A. Experimental Details and Evaluation Metrics

1) Implementation Details: Experiments are conducted on a Windows10 machine with Intel(R) Core (TM) i7-9700 CPU @ 3.00GHz, 64GB and NVIDIA GeForce RTX 3090, 24GB memory. All networks are based on the PyTorch deep learning framework and are trained from scratch to ensure a fair comparison. In the training phase, the network parameters are updated using the SGD optimizer with a momentum of 0.937. The initial value of the learning rate is 0.01. The warmup strategy is used in the first 5 epochs and then use the cosine learning rate decay strategy to perform gradient descent. The batch size is set to 64. The datasets used are the coating defect dataset, the publicly available Pascal VOC dataset and the PCB dataset. The Pascal VOC dataset is a widely used computer vision dataset containing 20 different categories of objects (such as people, cars, cats, dogs, etc.), providing researchers with a standard platform for testing and comparing computer vision algorithms. The PCB dataset is a synthesized PCB dataset [40] containing 1386 images with 6 categories of defects, which is challenging for the detection algorithm.

TABLE I
DATA AUGMENTATION CONTRAST EXPERIMENTS ON COATING DEFECT DATASET

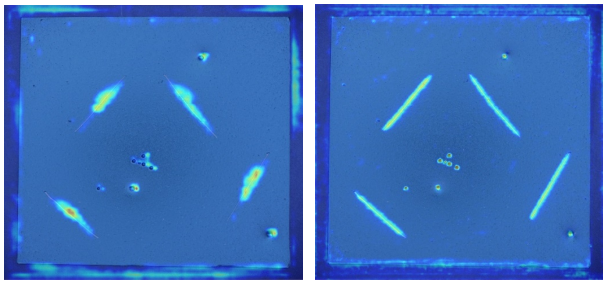
Methods	mAP_{50}
No Data Augmentation + YOLOv5s	91.6
Traditional Data Augmentation + YOLOv5s	92.3 (+0.7)
Mosaic + Traditional Data Augmentation + YOLOv5s	93.6 (+2.0)
FIRAWM + YOLOv5s	**95.6 (+4.0)**

TABLE II
DATA AUGMENTATION CONTRAST EXPERIMENTS ON PASCAL VOC DATASET

Methods	mAP_{50}
YOLOv5s	78
FIRAWM + YOLOv5s	**78.6 (+0.6)**

TABLE III
RESULTS OF ABLATION EXPERIMENTS WITH AFC-YOLO NETWORK

Methods	FPS	mAP_{50}
FIRAWM + YOLOv5s	91	95.6
FIRAWM + YOLOv5s + H4	78	96.1 (+0.5)
FIRAWM + YOLOv5s + H4 + coordinate attention	73	96.4 (+0.8)
FIRAWM + AFC-YOLO	**61**	**96.7 (+1.1)**



(a) (b)

Fig. 9. Comparison of the heat map before and after the introduction of the coordinate attention module. (a) Heat map without the coordinate attention module, and (b) heat map after the coordinate attention module is introduced.

2) Evaluation Metrics: The mAP_{50} is a common metric used to evaluate the performance of all categories. It is the average of all categories of AP when the IOU threshold is set to 0.5, and AP can be calculated as the area of the Precision × Recall curve with axes. Relevant metrics are defined as

$$P = \frac{TP}{TP + FP} \tag{6}$$

$$R = \frac{TP}{TP + FN} \tag{7}$$

$$\mathrm{AP} = \int_0^1 PR \, \mathrm{d}R \tag{8}$$

$$\mathrm{mAP} = \frac{1}{m} \sum_{m=1}^{m} \mathrm{AP}_m \tag{9}$$

where m is the total number of categories, AP is the evaluation metric of the detection accuracy of each category, P is the accuracy rate, R is the recall rate, TP is the number of correctly classified positive samples, FP is the number of incorrectly classified positive samples, FN is the number of falsely classified negative samples.

B. Data Augmentation Experiment

To verify the effectiveness of the proposed data augmentation strategy, the YOLOv5s in the YOLOv5 network is used as a baseline, and a series of comparative experiments are conducted on the coating defect dataset by changing the data augmentation method. Based on the characteristics of coating defects, the following popular data augmentation methods are selected: the proposed FIRAWM, No Data Augmentation, Traditional Data Augmentation, Mosaic + Traditional Data Augmentation. Table I shows the experimental results. The proposed FIRAWM performs better than traditional data augmentation methods and state-of-the-art data augmentation methods for coating defect detection, with an accuracy of 95.6 mAP_{50}. In particular, the proposed FIRAWM is 3.3 mAP_{50} higher than Traditional Data Augmentation and 2.0 mAP_{50} higher than Mosaic + Traditional Data Augmentation. To verify the universality of the proposed data augmentation strategy, experiments are conducted on the public Pascal VOC dataset, and the results are shown in Table II. The performance of the FIRAWM is 0.6 mAP_{50} higher than the data augmentation method provided by the YOLOv5s. The results of the experiment show that the proposed data augmentation strategy is successful in improving the accuracy of coating defect detection. Furthermore, it has a certain degree of universality that can be applied to other detection tasks.

C. Network Optimization Design Experiment

1) Ablation Study: According to the characteristics of coating defects, a series of improvements are made to the YOLOv5s network by adding an additional detection layer (denoted by H4), and introducing the coordinate attention module and the ASFF module. To verify the contribution of these modules to coating defect detection, a series of ablation experiments are conducted on the coating defect dataset after the proposed FIRAWM data augmentation strategy, and the experimental results are shown in Table III.

The results from Table III show that all the improved modules contribute to the detection results, and the YOLOv5s provides a detection mAP_{50} of 95.6. Adding an additional detection layer increases the mAP_{50} of the YOLOv5s by 0.5 mAP_{50}. Although the computational cost is increased, the ability to detect small defects is significantly improved. With the introduction of the coordinate attention and ASFF modules, the AFC-YOLO achieves an accuracy of 96.7 mAP_{50} for coating defects, which is 1.1 mAP_{50} higher than the YOLOv5s. Despite sacrificing some of the detection speed of the APC-YOLO, it still close to the real-time detection requirement with a detection speed of 61 FPS. Then, the feature maps before and after the introduction of the coordinate attention module are analyzed by a heat map, as shown in Fig. 9. It can be seen that after the introduction of the coordinate attention module, the network filters the redundant features through the attention mechanism. This allows the network to better locate the defective region and assign higher confidence. Fig. 10 shows the comparison between the YOLOv5s and the proposed network on the detection results of densely clustered coating defects. It is evident that the YOLOv5s has missed detection, while the

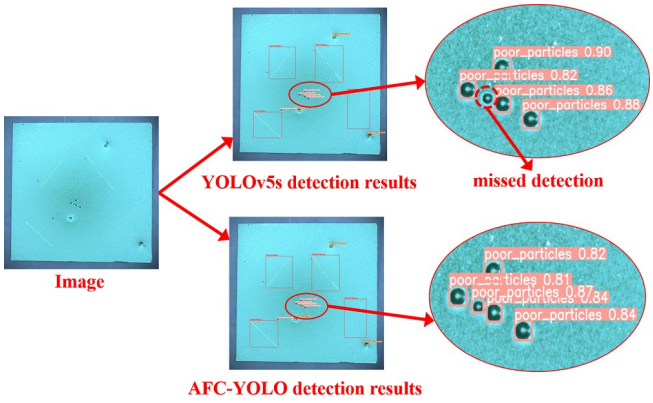


Fig. 10. Visual comparison of the YOLOv5s and the AFC-YOLO detection results.

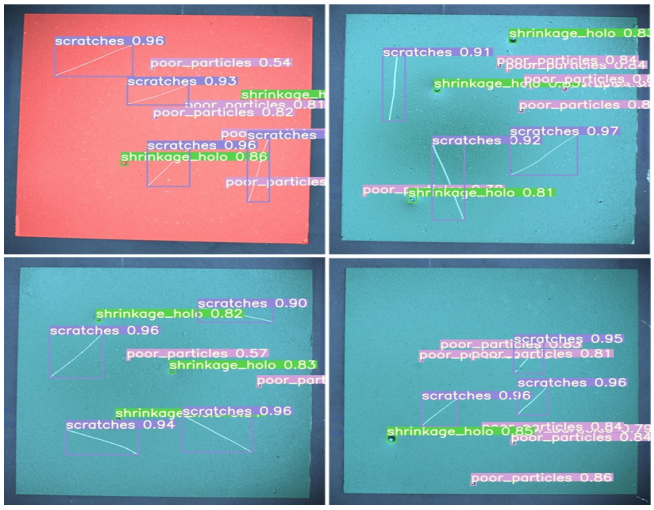


Fig. 11. Visualization of the AFC-YOLO network detection results.

proposed method has almost no missed and false detections, indicating that the proposed method achieves remarkable performance in coating defect detection. Fig. 11 shows the visualization of the proposed method. It can be seen that the proposed method successfully detects different sizes and types of coating defects with high detection accuracy.

2) Comparison of Different Methods: To further demonstrate the advantages of the proposed AFC-YOLO on the coating defect dataset, a comparison with advanced networks such as YOLOv7 [41], YOLOv5x, YOLOv5s, YOLOv3 is performed on the coating defect dataset, and the state-of-the-art results are obtained. As shown in Table IV, the AFC-YOLO is 0.1 mAP_{50} and 0.2 mAP_{50} higher than YOLOv5x and YOLOv7, respectively, but the AFC-YOLO parameters and size are much smaller than them, which is more suitable for deployment on mobile terminals. Compared to the YOLOv5s without FIRAWM data augmentation, the proposed method improves the mAP_{50} by 3.1, demonstrating the effectiveness of the proposed data augmentation strategy and network model improvement. The AFC-YOLO is also 1.6 mAP_{50} higher than the industry standard YOLOv3 network. As a result, the proposed method outperforms some outstanding work in coating defect detection, achieving a balance between speed and accuracy.

TABLE IV
PERFORMANCE COMPARISON OF THE PROPOSED NETWORK WITH OTHER NETWORKS ON THE COATING DEFECT DATASET

Methods	parameters (M)	FLOPs (G)	Model (M)	mAP_{50}
YOLOv5s	7.1	16.4	14	93.6
FIRAWM + YOLOv3	62.6	154.7	122	95.1
FIRAWM + YOLOv5s	7.1	16.4	14	95.6
FIRAWM + YOLOv7	37.2	104.8	73	96.5
FIRAWM + YOLOv5x	87.3	217.4	683	96.6
FIRAWM + AFC-YOLO	**18.6**	**44.3**	**36**	**96.7**

TABLE V
THE PERFORMANCE OF THE PROPOSED METHOD IS COMPARED WITH OTHER ADVANCED METHODS ON PASCAL VOC DATASET

Methods	mAP_{50}
YOLOv5s	78
FIRAWM + YOLOv5s	78.6 (+0.6)
FIRAWM + AFC-YOLO	**79.4 (+1.4)**

TABLE VI
THE PERFORMANCE OF THE PROPOSED METHOD IS COMPARED WITH OTHER ADVANCED METHODS ON PCB DATASET

Methods	mAP_{50}
YOLOv5s	95.2
AFC-YOLO	**95.7**

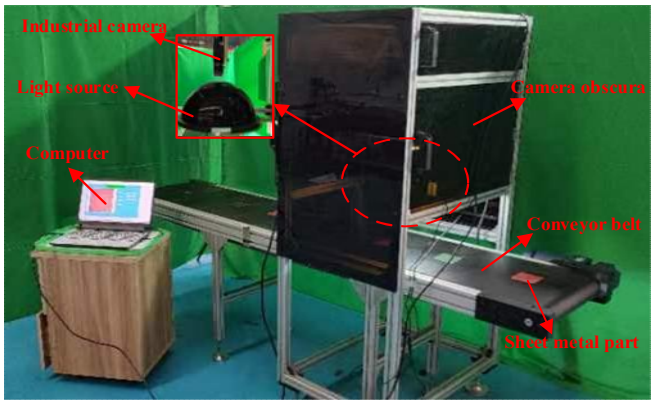


Fig. 12. Physical prototype diagram.

3) Robustness analysis on different detection tasks: To verify the versatility and effectiveness of the proposed method for other detection tasks, comparative experiments are performed on the publicly available Pascal VOC dataset, and the results are shown in Table V. The results show that the proposed FIRAWM data augmentation strategy improves the YOLOv5s by 0.6 mAP_{50}. Using the same FIRAWM data augmentation strategy, the AFC-YOLO improved by 0.8 mAP_{50} over the YOLOv5s to reach 79.4 mAP_{50}. To further validate the robustness of the AFC-YOLO for other defect detection tasks, experiments are performed on the PCB dataset. Table VI shows that the AFC-YOLO is still 0.5 mAP_{50} higher than the YOLOv5s in other defect detection tasks. In summary, the experimental results on different datasets show that the proposed method has certain effectiveness and robustness for different detection tasks.

D. Physical Prototype Experiment

The proposed coating defect detection method is deployed and a physical prototype of an online defect detection experiment is built, as shown in Fig. 12. The whole detection system is composed of hardware and software, which simulates the actual industrial detection environment and conditions. 80 sheet metal parts with a mixture of different coating defects are prepared and manually placed on the conveyor belt for online detection. Through statistical analysis of the experimental data, it is found that in the actual dynamic detection scene, the speed of the conveyor belt, the change of the external lighting environment and the parameter setting of the industrial camera will have different degrees of influence on the quality of the acquired images, thereby affecting the detection accuracy. When the algorithm post-processing threshold is set to 0.4 and the conveyor speed is 0.36 m/s, the coating defect detection accuracy reached the highest 96.4%. Experimental results show that the proposed method has certain universality in different environments, and achieves the balance of speed and accuracy, which meets the requirements of practical applications.

IV. Conclusion

To improve the accuracy of coating defect detection, this paper analyzes the difficulties of existing coating defect detection and proposes a coating defect detection method based on data augmentation and network optimization design. By analyzing the causes and characteristics of coating defects, a coating defect dataset with practical application value is constructed, and a new data augmentation strategy is proposed, which effectively improves the quality of the dataset without adding additional computational overhead and cost. To further improve the performance of the coating defect detection algorithm, this paper improves the YOLOv5 network by adding an additional detection layer and introducing the coordinate attention module and ASFF module, and the accuracy of the coating defect detection reaches 96.7 mAP_{50}. Finally, the effectiveness and versatility of the proposed method are verified by conducting experimental tests and comparisons with other advanced methods.

In future work, the lightweight design of the network backbone will be considered to further improve the detection speed while maintaining the detection accuracy, and to combine with the spraying robot to repair some of the detected defects.

References

[1] H. Yang, Y. Wang, J. He, Z. Yao, and Q. Bi, "Segmentation of track surface defects based on machine vision and neural networks," *IEEE Sensors J.*, vol. 22, no. 2, pp. 1571–1582, Jan. 2022.

[2] L. Xiao, B. Wu, and Y. Hu, "Surface defect detection using image pyramid," *IEEE Sensors J.*, vol. 20, no. 13, pp. 7181–7188, Jul. 2020.

[3] K. K. Kieselbach, M. Nöthen, and H. Heuer, "Development of a visual inspection system and the corresponding algorithm for the detection and subsequent classification of paint defects on car bodies in the automotive industry," *J Coat Technol and Res*, vol. 16, no. 4, pp. 1033–1042, Feb. 2019.

[4] J. Xu, J. Zhang, K. Zhang, T. Liu, D. Wang, and X. Wang, "An APF-ACO algorithm for automatic defect detection on vehicle paint," *Multimed Tools Appl*, vol. 79, no. 35–36, pp. 25315–25333, Jul. 2020.

[5] L. Taylor and G. Nitschke, "Improving deep learning using generic data augmentation," in *Proc. IEEE Symp. Ser. Comput. Intell. (SSCI)*, Aug. 2018, pp. 1542–1547.

[6] Z. Gao, G. Yang, and Z. Liang, "Novel feature fusion module-based detector for small insulator defect detection," *IEEE Sensors J.*, vol. 21, no. 15, pp. 16807–16814, Aug. 2021.

[7] S. Yun, D. Han, S. Chun, S. J. Oh, Y. Yoo, and J. Choe, "CutMix: Regularization strategy to train strong classifiers with localizable features," in *Proc. IEEE/CVF Int. Conf. comput. Vis. (ICCV)*, Otc. 2019, pp. 6022–6031.

[8] P. Chen, S. Liu, H. Zhao, and J. Jia, "GridMask data augmentation," 2020, *arXiv:2001.04086.*

[9] H. Zhang, M. Cisse, Y. N. Dauphin, and D. Lopez-Paz, "mixup: Beyond empirical risk minimization," 2018, *arXiv:171009412.*

[10] K. Chen, Z. Zeng, and J. Yang, "A deep region-based pyramid neural network for automatic detection and multi-classification of various surface defects of aluminum alloys," *J Build Eng*, vol. 43, p. 102523, Nov. 2021.

[11] G. Wen, Z. Gao, Q. Cai, Y. Wang, and S. Mei, "A novel method based on deep convolutional neural networks for wafer semiconductor surface defect inspection," *IEEE Trans. Instrum. Meas.*, vol. 69, no. 12, pp. 9668–9680, Dec. 2020.

[12] W. Zhao, F. Chen, H. Huang, D. Li, and W. Cheng, "A new steel defect detection algorithm based on deep learning," *Comput Intel Neurosc*, vol. 2021, pp. 1–13, Mar. 2021.

[13] F. Chang, M. Liu, M. Dong, and Y. Duan, "A mobile vision inspection system for tiny defect detection on smooth car-body surfaces based on deep ensemble learning," *Meas. Sci. Technol.*, vol. 30, no. 12, p. 125905, Sep. 2019.

[14] M. Xu, S. Yoon, A. Fuentes, and D. S. Park, "A comprehensive survey of image augmentation techniques for deep learning," 2022, *arXiv:2205.01491.*

[15] C. Shorten and T. M. Khoshgoftaar, "A survey on image data augmentation for deep learning," *J Big Data*, vol. 6, no. 1, p. 60, Jul. 2019.

[16] T. DeVries and G. W. Taylor, "Improved regularization of convolutional neural networks with cutout," 2022, *arXiv:1708.04552.*

[17] X. Liu, F. Shen, J. Zhao, and C. Nie, "RandomMix: A mixed sample data augmentation method with multiple mixed modes," 2022, *arXiv:220508728.*

[18] A. Bochkovskiy, C.-Y. Wang, and H.-Y. M. Liao, "YOLOv4: Optimal speed and accuracy of object detection," 2020, *arXiv:2004.10934.*

[19] L. Chen *et al.*, "Detection perceptual underwater image enhancement with deep learning and physical priors," in *Proc. IEEE Trans. Circuits. Syst. Video. Technol.*, vol. 31, pp. 3078–3092, Aug. 2021.

[20] L. Liu, D. Cao, Y. Wu, and T. Wei, "Defective samples simulation through adversarial training for automatic surface inspection," *Neurocomputing*, vol. 360, pp. 230–245, Sep. 2019.

[21] H. Lu, M. Du, K. Qian, X. He, and K. Wang, "GAN-Based data augmentation strategy for sensor anomaly detection in industrial robots," *IEEE Sensors J.*, vol. 22, no. 18, pp. 17464–17474, Sep. 2022.

[22] E. D. Cubuk, B. Zoph, J. Shlens, and Q. V. Le, "Randaugment: Practical automated data augmentation with a reduced search space," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2020, pp. 702–703.

[23] G. Ghiasi *et al.*, "Simple copy-paste is a strong data augmentation method for instance segmentation," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2021, pp. 2918–2928.

[24] H.-S. Fang, J. Sun, R. Wang, M. Gou, Y.-L. Li, and C. Lu, "InstaBoost: Boosting instance segmentation via probability map guided copy-pasting," in *Proc. IEEE Int. Conf. Comput. Vis. (ICCV)*, Aug. 2019, pp. 682–691.

[25] N. Dvornik, J. Mairal, and C. Schmid, "Modeling visual context is key to augmenting object detection datasets," in *Proc. Eur. Conf. Comput. Vis. (ECCV)*, 2018, pp. 375–391.

[26] X. Xiang, Z. Wang, and Y. Qiao, "An improved YOLOv5 crack detection method combined with transformer," *IEEE Sensors J.*, vol. 22, no. 14, pp. 14328–14335, Jul. 2022.

[27] Y. Song, Z. Xie, X. Wang, and Y. Zou, "MS-YOLO: Object detection based on YOLOv5 optimized fusion millimeter-wave radar and machine vision," *IEEE Sensors J.*, vol. 22, no. 15, pp. 15435–15447, Aug. 2022.

[28] S. Liu, L. Qi, H. Qin, J. Shi, and J. Jia, "Path aggregation network for instance segmentation," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2018, pp. 8759–8768.

[29] X. Zhu, S. Lyu, X. Wang, and Q. Zhao, "TPH-YOLOv5: Improved YOLOv5 based on transformer prediction head for object detection on drone-captured scenarios," in *Proc. IEEE/CVF Int. Conf. Comput. Vis. Workshops (ICCVW)*, Oct. 2021, pp. 2778–2788.

[30] Y. Shi *et al.*, "Rethinking the Detection Head Configuration for Traffic Object Detection," 2022, *arXiv:2210.03883.*

[31] C. Wang, A. Bochkovskiy, and H. Liao, "Scaled-yolov4: Scaling cross stage partial network," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2021, pp. 13029–13038.
[32] Q. Hou, D. Zhou, and J. Feng, "Coordinate attention for efficient mobile network design," in *P*roc. *IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2021, pp. 13713–13722.
[33] Y. Li, M. Yang, J. Hua, Z. Xu, J. Wang, and X. Fang, "A channel attention-based method for micro-motor armature surface defect detection, " *IEEE Sensors J.*, vol. 22, no. 9, pp. 8672–8684, May. 2022.
[34] Z. Zhong *et al*., "Squeeze-and-Attention networks for semantic segmentation, " in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2020, pp. 13065–13074.
[35] X. Pan *et al*., "On the integration of Self-Attention and convolution, " in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2022, pp. 815–825.
[36] M. Lin, Q. Zhu, Y. Yin, Y. Fan, Z. Su, and S. Zhang, "Damage detection method of mining conveyor belt based on deep learning," *IEEE Sensors J.*, vol. 22, no. 11, pp. 10870–10879, Jun. 2022.
[37] M. Tan, R. Pang, and Q. V. Le, "EfficientDet: Scalable and efficient object detection," in *P*roc. *IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2020, pp. 10781–10790.
[38] S. Liu, D. Huang, and Y. Wang, "Learning spatial fusion for single-shot object detection," 2019, *arXiv: 1911.09516.*
[39] J. Redmon and A. Farhadi, "YOLOv3: An incremental improvement," 2018, *arXiv: 1804.02767*.
[40] W. Huang, P. Wei, "A PCB dataset for defects detection and classification," 2019, *arXiv: 1901.08204.*
[41] C.-Y Wang, A. Bochkovskiy, and H.-Y. M. Liao, "YOLOv7: Trainable bag-of-freebies sets new state-of-the-art for real-time object detectors,'' 2022, *arXiv:2207.02696.*

Kai Tang is currently pursuing the Ph.D. degree with the College of Mechanical Engineering, Hefei University of Technology, China. His current research interests include coating defect detection, machine learning, and computer vision.

Bin Zi received the Ph.D. degree from The Xidian University in 2007. He is currently a professor, doctoral supervisor, and the Dean of the School of Mechanical Engineering, Hefei University of Technology.

He has received the National Science Fund for Distinguished Young Scholars of China and presided over more than 20 projects including the National Natural Science Foundations of China and National Key Research and Development Program on Intelligent Robots. He is the associate editor of ASME Journal of Mechanical Design, and editorial board member of the Chinese Journal of Mechanical Engineering, Mechanical Sciences, Robot, and Control and Decision. His research interests include the theory, technology and equipment of flexible driven robots, machine learning, control and automation of intelligent manufacturing system.

Feng Xu received the M.S. degree in safety engineering from China University of Petroleum (Beijing), China, in 2016. He is currently pursuing the PH.D. degree in Advanced Manufacturing from Hefei University of technology, china. His current research interests include motion planning and control of spray-painting robot.

Weidong Zhu received the dual B.S. degrees in mechanical engineering and computational science from Shanghai Jiao Tong University, Shanghai, China, in 1986, the M.S. degree from Arizona State University, Tempe, AZ, USA in 1988, and the Ph.D. degree from the University of California, Berkeley, CA, USA, in 1994, both in mechanical engineering.

From 1994 to 1997, he was an Assistant Professor with the Chinese University of Hong Kong, Hong Kong. From 1997 to 1999, he was an Assistant Professor with the University of North Dakota, Grand Forks, ND, USA. Since 2007, he has been a Professor with the Department of Mechanical Engineering, University of Maryland Baltimore County, Baltimore, MD, USA. His research interests include the fields of dynamics, vibration, control, applied mechanics, structural health monitoring, metamaterials, and wind energy, and involves analytical development, numerical simulation, experimental validation, and industrial application.

Prof. Zhu is a Fellow of the American Society of Mechanical Engineers. He has served as an Associate Editor for the ASME Journal of Vibration and Acoustics and the ASME Journal of Dynamic Systems, Measurement and Control, and as a Subject Editor for the Journal of Sound and Vibration and Nonlinear Dynamics.

Kai Feng received the B.S. degree in mechatronics engineering from Anhui University, Hefei, Anhui, China, in 2017, and the M.S. degree in electromechanical engineering from the Hefei University of Technology, Hefei, in 2020. He is currently pursuing the Ph.D. degree in electromechanical engineering with the University of Macau, Macau, China. His main research interests include teleoperated robot systems, human-machine interaction and biological micromanipulation.