ISSN: 2710-0588

An Improved YOLOv5s Method for Light Guide Plate Defect Detection

Huanyu Zhang

School of Electrical Information Engineering, Northeast Petroleum University, Daqing 163318, China

Abstract

Aiming at the problems of over 100 million pixels in a single image of a large-size light guide plate, only a dozen pixels in the defect size, uneven brightness, and slow detection speed, this paper designs an improved YOLOv5s method for light guide plate defect detection. YOLOv5s is used as the basic model, and the attention mechanism CBAM is added to the backbone feature extraction network (backbone) part; PAN-Net is replaced with Bi-FPN structure in the feature fusion network (Neck) part, and a small target detection layer is added at the same time; YOLOv5s Obtain higher accuracy and improve algorithm robustness. Finally, a comparative analysis is performed based on the self-built dataset LGPDD. The experimental results show that compared with the original YOLOv5s model, the average accuracy mAP is increased by 1.4%; the model detection speed can meet the production requirements, and has a good application prospect in industrial deployment.

Keywords

Light Guide Plate Defect Detection; Improve YOLOv5s; Aattention Mechanism.

1. Introduction

In recent years, many deep learning defect detection methods have been widely used in industrial fields such as LCD. And some preliminary attempts have also been started in the field of light guide plate defect detection. In 2021, Li et al[1] proposed an end-to-end multi-task learning network architecture for mobile phone light guide plate defect detection, which obtained multi-scale features through a U-shaped coding structure, and used feature fusion to interact with multi-scale features[2] proposed a dense bilinear convolutional neural network (BCNN), an end-to-end defect detection network. Generally, defect detection based on deep learning classification networks can only obtain rough locations. Relatively speaking, the target detection network can obtain the accurate location and category information of the target. Object detection models can be divided into two broad categories: two-stage and one-stage networks. The two-stage network needs to generate candidate regions that may contain defects before object detection, mainly including R-CNN[3], Faster R-CNN[4]and other networks. The one-stage network directly performs defect detection and identification through features extracted from the network, mainly including networks such as YOLO[5-6]series and SSD [7]. The detection speed is faster, and good results have been achieved in the problem of light guide plate defect detection. In 2022, Yao[8] proposed an AYOLOv3-Tiny network for LGP defect detection. The convolution operation of the traditional YOLOv3-Tiny backbone is improved to meet the needs of industrial inspection and verify the excellent performance of the YOLO model on the task of light guide plate defect detection.

ISSN: 2710-0588

Therefore, this paper takes YOLOv5s as the basic structure, and combines the defect characteristics and detection requirements of light guide plates to appropriately improve it, and proposes an improved YOLOv5s method for light guide plate defect detection.

2. Improved YOLOv5s Algorithm

Considering the real-time nature of the algorithm, this paper chooses YOLOv5s as the infrastructure for improvement. By analyzing the defect image of the light guide plate, it can be found that the defect detection of the light guide plate can be attributed to the problem of small target detection under the large-size image. The defect only occupies about ten pixels on the image. At the same time, the defect detection and classification are performed directly on the production line Therefore, the problem of detection time must be solved. It is difficult for traditional defect detection methods to quickly and accurately extract the characteristics of various defects, resulting in low detection accuracy. Based on this, this paper proposes an improved YOLOv5s rapid detection method for light guide plate defects: first, a small target detection layer is added on the basis of the YOLOv5s structure, then the attention mechanism CBAM is integrated into YOLOv5, and finally PAN-Net is replaced with Bi-FPN structure. Figure 1 shows the improved YOLOv5s network structure diagram, which consists of Input, Backbone, Neck, and Head.

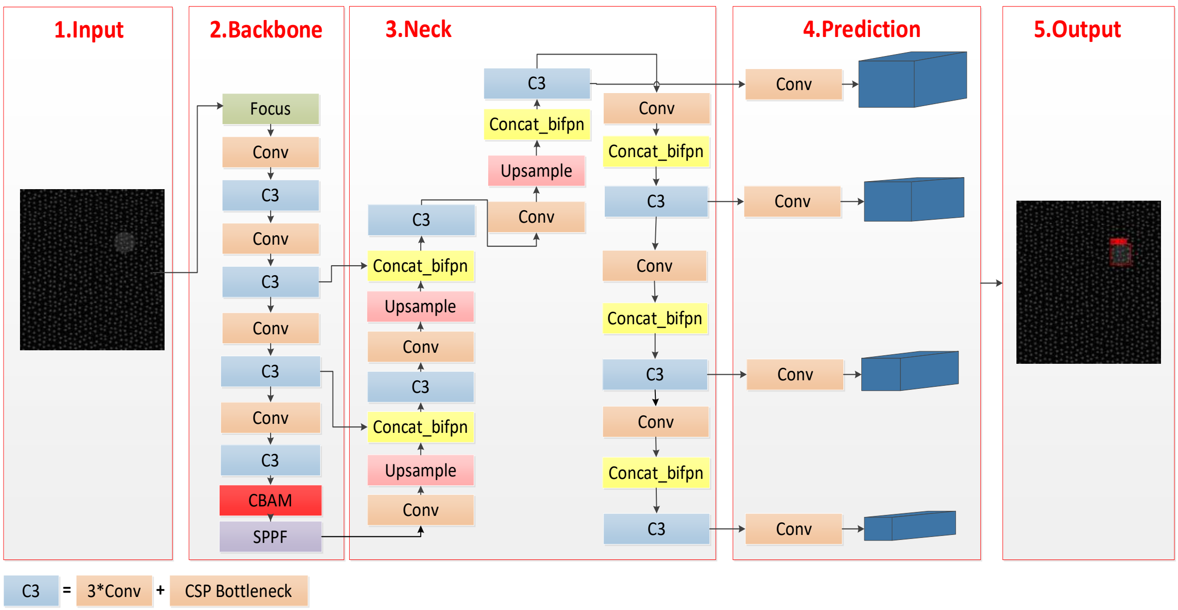

Figure 1. AYOLOv5s network structure

2.1. Add Attention Mechanism

CBAM is a combination of channel attention mechanism and spatial attention mechanism, which can achieve better results than SENet's attention mechanism that only focuses on channels. The schematic diagram of its implementation is shown in Figure 2 below. CBAM will process the channel attention mechanism and the spatial attention mechanism respectively for the input feature layer.

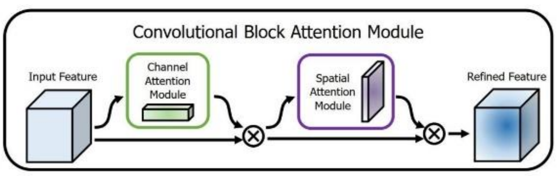

Figure 2. Schematic diagram of CBAM structure

ISSN: 2710-0588

Figure 3 below shows the specific implementation of the channel attention mechanism and the spatial attention mechanism: the upper part of the image is the channel attention mechanism, and the implementation of the channel attention mechanism can be divided into two parts. layers, which perform global average pooling and global max pooling, respectively. After that, the results of average pooling and max pooling are processed by using the shared fully connected layer. We will add the two processed results, and then take a sigmoid. At this time, we obtain each channel of the input feature layer. weight (between 0 and 1). After obtaining this weight, we can multiply this weight by the original input feature layer. The lower part of the image is the spatial attention mechanism. We will take the maximum value and average value on the channel of each feature point for the input feature layer. Then stack the two results, adjust the number of channels using a convolution with a channel number of 1, and then take a sigmoid. At this time, we obtain the weight of each feature point of the input feature layer (between 0-1). After obtaining this weight, we can multiply this weight by the original input feature layer.

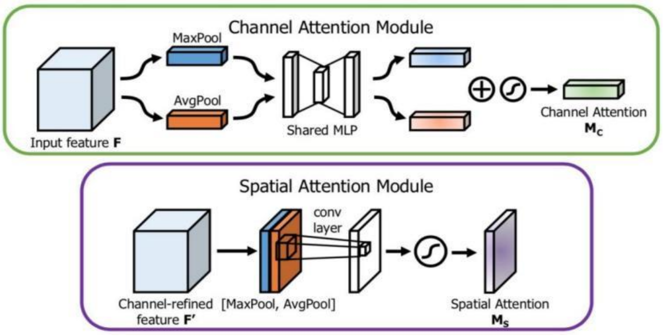

Figure 3. Schematic diagram of CBAM implementation

The attention mechanism is a plug-and-play module that can be seamlessly integrated into any CNN architecture with little overhead and can be trained end-to-end with the CNN network. Suppose we have an intermediate feature F (H×W×C) as the output of a convolutional layer. Depending on the depth of the layers, each such feature layer captures useful information, such as simple edges, shapes, etc., to obtain a more complex semantic representation of the input. We want the network to pay more attention to (i.e. focus on) the important parts of these feature maps. In this paper, CBAM is inserted into YOLOv5, and the implementation process is shown in Figure 4 below. Attention regions can be extracted using CBAM to help YOLOv5 resist confusing information and focus on useful target objects.

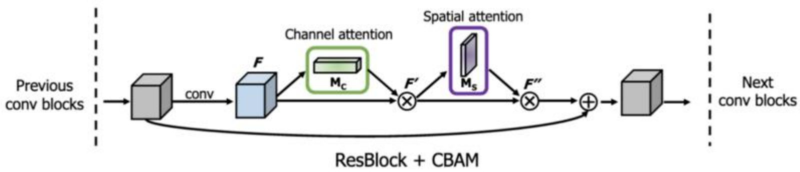

Figure 4. Schematic diagram of the integrated CABM structure

2.2. Replace PAN-Net with Bi-FPN

Since it is difficult to detect small objects in target detection, because during the convolution process, large objects have many pixels, and small objects have few pixels. With the deepening of convolution, the features of large objects are easily retained, and small objects The later the feature is, the easier it is to be ignored. So the FPN structure was created. Figure 5 below is the FPN structure diagram. After continuously down-sampling the feature points, it has a bunch of feature layers with high semantic content, and then up-sampling again, so that the length and

ISSN: 2710-0588

width of the feature layer become larger again, and a large-size feature map is used to detect small targets. Of course, it is not possible to simply upsample, because the results of such upsampling will not be clear about the characteristics and information of the small target, so we can stack the feature layers with the same length and width as the upsampling in the downsampling, which can ensure Characteristics and information of small targets. It is very similar to the U-net structure. But there is more stacking process.

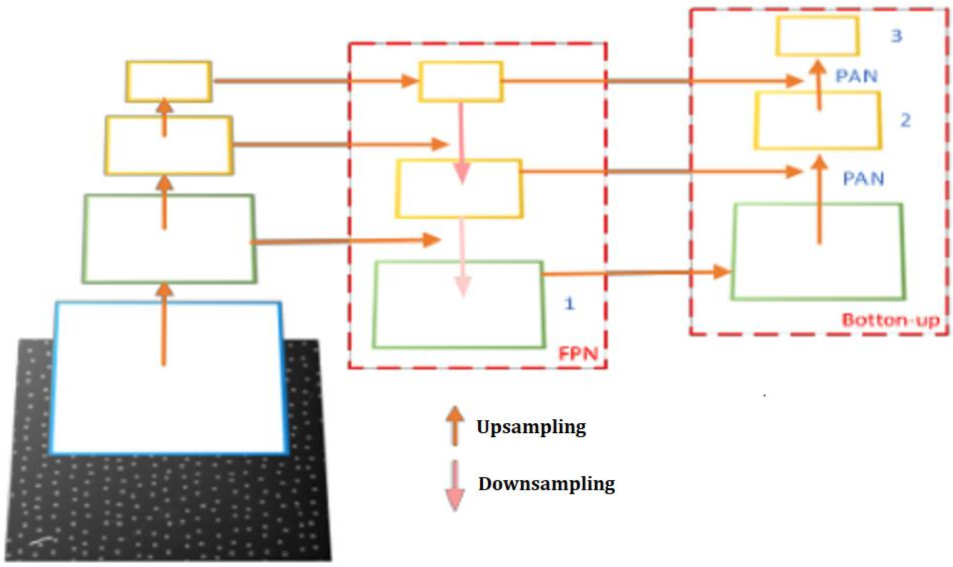

Figure 5. Schematic diagram of FPN structure

A weighted bidirectional feature pyramid network is proposed in EfficientDet [12], which allows simple and fast multi-scale feature fusion. The author's purpose is to pursue a more efficient multi-scale fusion method. In the past feature fusion, features of different scales were treated equally, and the author introduced weights (similar to attention) to better balance the feature information of different scales. In the paper, the author also has a comparison with other FPNs. The comparison is shown in Figure 6 below.

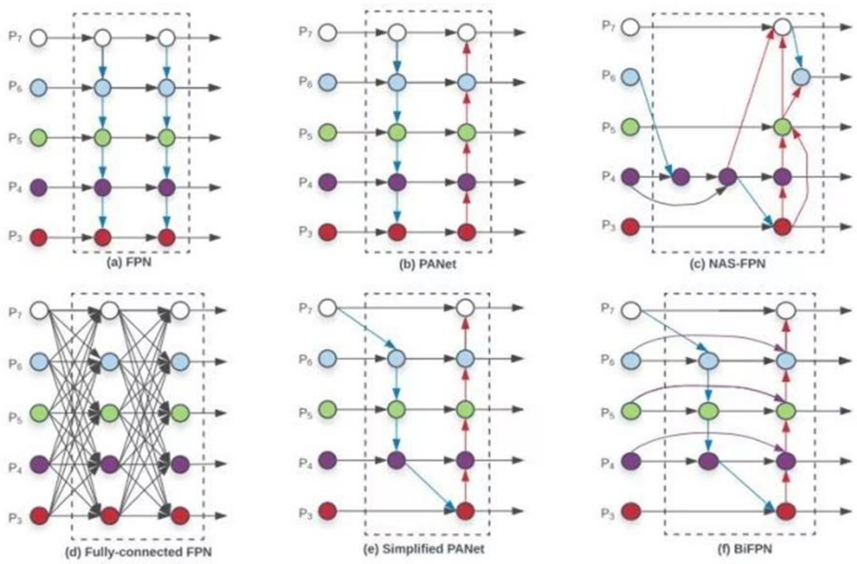

Figure 6. A comparison with other FPNs

ISSN: 2710-0588

3. Analysis of Experimental Results

3.1. Light Guide Plate Defects Dataset

This paper mainly studies the defects of the light guide plate, such as bright spots, crushing, scratches, dirt, foreign objects, black spots, and curved surface spots. Since the classification of these light guide plates is highly subjective, there is no obvious boundary between defects in actual imaging. Therefore, this paper divides these defects into two categories according to the technical requirements of the manufacturer, the work experience of the front-line workers and the imaging characteristics of the light guide plates. There are two types of point defects and line defects. Figure 7 shows the classification of light guide plate defects.

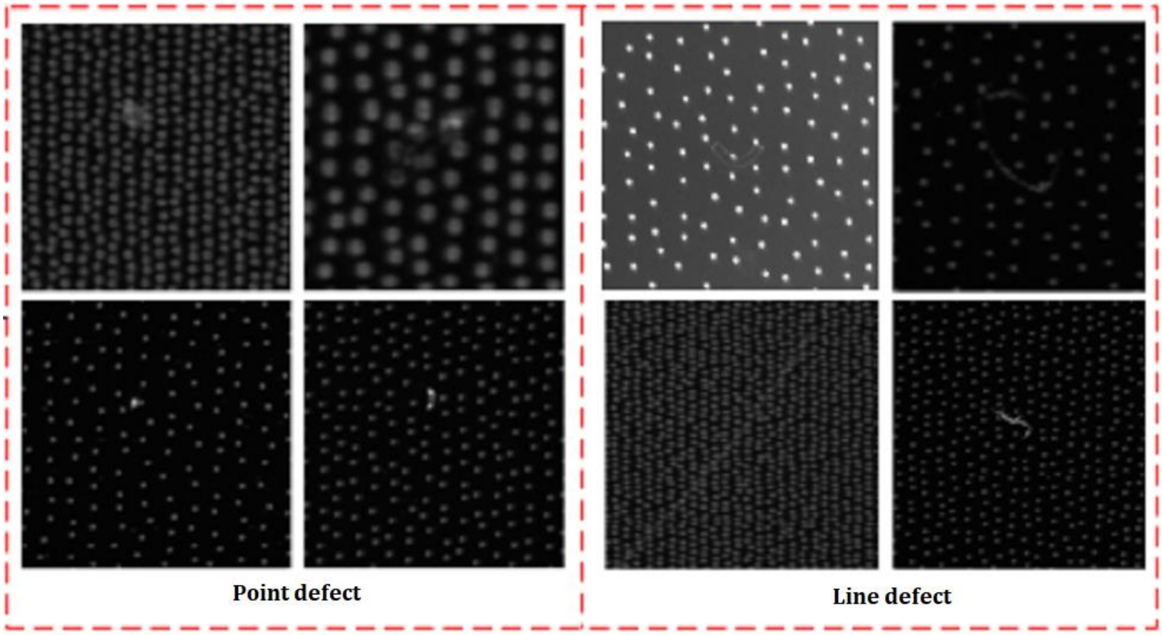

Figure 7. Defect classification of light guide plate

The Light Guide Plate Defect Dataset (LGPDD) comes from a real industrial domain. Contains a total of 500 light guide plates within 12-17 inches. After preprocessing, a 640*640 small image is obtained and a data set is made. This paper uses data expansion to balance the number of two defects. Data augmentation includes rotation with 50% probability, and 120-150% brightness boost. The specific parameters of the expanded dataset are shown in Table 1.

Table 1. LGPDD dataset parameters

Defect type	Label Name	Quantity
Point defect	dot	2362
Line defect	line	2140

3.2. Experimental Environment and Model Training

The specific configuration used in this experiment is: Windows10 operating system, single RTX2060 graphics card, python 3.7 virtual environment, pytorch 1.7 deep learning open source framework, and CUDA11 for acceleration. This article compares the improved YOLOv5s model with the original YOLOv5s on the self-built data set. The code refers to the yolov5 official code version 6.1. The number of iterations is 300 epoch, the optimizer selects SGD, the initial learning rate is 0.01, the momentum parameter is 0.937, and the warmup method is used for pre-learning with epoch 3 and momentum parameter 0.8. In the warm-up stage, one-dimensional

ISSN: 2710-0588

linear interpolation is used to update the learning rate. , and the learning rate is updated by cosine annealing after pre-learning. A validation phase occurs every time an iteration is completed in the training process. The weights file is automatically updated after each iteration.

3.3. Evaluation Indicators and Analysis of Experimental Results

This paper studies the use of mAP, Precision, Recall, and FPS as the measurement criteria for the defect detection model of the light guide plate. Among them, mAP is the average accuracy of the model, Precision refers to the number of correctly predicted positive samples in the prediction data set divided by the number of positive samples predicted by the model; Recall refers to the number of correctly predicted positive samples in the predicted data set divided by the actual number of positive samples is the number of positive samples. FPS represents the number of detected images per second. AP in the mAP calculation is the area under the P-R curve. The specific calculation is based on formulas (1), (2), (3), (4):

$$mAP = \frac{\sum_{i=1}^{N} AP_i}{N} \tag{1}$$

$$\mathrm{AP} = \int_0^1 P(R)dR \tag{2}$$

$$\mathrm{Pr}ecision = \frac{TP}{TP + FP} \tag{3}$$

$$\mathrm{Re\,call} = \frac{TP}{TP + F\mathrm{N}} \tag{4}$$

In the formula, the AP value refers to the area of the P-R curve; the mAP value is obtained by averaging the APs of all categories; N represents the total number of detected categories. In this experiment, N=2, and the larger the mAP value, the better the detection effect of the algorithm. , the recognition accuracy is higher; TP, FP and FN represent the number of correct detection frames, false detection frames and missed detection frames, respectively.

The improved YOLOv5s network is compared with object detection networks such as SSD, YOLOv3, Faster R-CNN, etc. by mAP and FPS. Table 3 summarizes the experimental results. As can be seen from Table 3, the improved YOLOv5s has a significant improvement in the accuracy of point and line defects, but a slight decrease in FPS. Compared with YOLOv5s, the improved YOLOv5s improves the average detection accuracy of defects by 1.7%, but reduces the FPS by 56.5. Due to the higher complexity of the network, the increase in volume results in a decrease in speed. Although the detection speed of the improved YOLOv5s is lost, it can still meet the detection requirements based on the light guide plate production process.

Table 2. Summary of experimental results

Network Model	mAP	FPS
SSD	96.97%	23.1
YOLOv3	89.19%	62
YOLOv5s	97.5%	102
Faster r-cnn	81.65%	17.5
ours	98.9%	45.5

ISSN: 2710-0588

4. Conclusion

This paper mainly focuses on the abnormal detection of the light guide plate as a whole, and proposes an improved YOLOv5-based light guide plate defect detection and identification algorithm. The algorithm in this chapter first adds a small target detection layer on the basis of the YOLOv5 structure, and then integrates the attention mechanism CBAM into YOLOv5, and finally uses Bi-FPN replaces PAN-Net structure. Finally, according to the original data collected in real time on the light guide plate production line, a data set that can be trained for the network is constructed after cutting. The experiments show that the improved YOLOv5s algorithm has a better detection rate in the task of light guide plate defect detection and recognition, and Inference time can meet production demands.

References

[1] Li Y, Li J. An end-to-end defect detection method for mobile phone light guide plate via multitask learning[J]. IEEE Transactions on Instrumentation and Measurement, 2021, 70: 1-13.

[2] Hong L, Wu X, Zhou D, et al. Effective Defect Detection Method Based on Bilinear Texture Features for LGPs[J]. IEEE Access, 2021, 9: 147958-147966.

[3] Girshick R, Donahue J, Darrell T, et al. Rich feature hierarchies for accurate object detection and semantic segmentation[C]//Proceedings of the IEEE conference on computer vision and pattern recognition. 2014: 580-587.

[4] Ren S, He K, Girshick R, et al. Faster r-cnn: Towards real-time object detection with region proposal networks[J]. Advances in neural information processing systems, 2015, 28.

[5] Redmon J, Farhadi A. Yolov3: An incremental improvement[J]. arXiv preprint arXiv:1804.02767, 2018.

[6] Bochkovskiy A, Wang C Y, Liao H Y M. Yolov4: Optimal speed and accuracy of object detection[J]. arXiv preprint arXiv:2004.10934, 2020.

[7] Liu W, Anguelov D, Erhan D, et al. Ssd: Single shot multibox detector[C]//European conference on computer vision. Springer, Cham, 2016: 21-37.

[8] Yao J, Li J. AYOLOv3-Tiny: An improved convolutional neural network architecture for real-time defect detection of PAD light guide plates[J]. Computers in Industry, 2022, 136: 103588.