Efficient Real-time Detection of Complex Tire Cord Defects on Air-jet Loom

1st Juan Du
School of Automation Science and Engineering
South China University of Technology
Guangzhou, China
dujuan@scut.edu.cn

2nd Yuhan Li
School of Automation Science and Engineering
South China University of Technology
Guangzhou, China
auliyuhan@mail.scut.edu.cn

2024 9th International Conference on Automation, Control and Robotics Engineering (CACRE) | 979-8-3503-5030-2/24/$31.00 ©2024 IEEE | DOI: 10.1109/CACRE62362.2024.10635081

***Abstract*—The paper presents an improved YOLOv8 model tailored for the detection of defects in air-jet loom tire cord fabric, a critical component in the automotive industry. Current defect detection methods, predominantly manual, are inefficient and prone to errors, necessitating an automated solution. Leveraging advancements in machine vision and deep learning, this study presents an vision-based defect detection model that tackles the distinctive obstacles posed by the complex background, variations in target sizes, and indistinct defect features of tire cord fabric defects. The proposed model integrates a novel BiFPN (Bidirectional Feature Pyramid Network) into the YOLOv8 framework to enhance feature fusion and employs a Dynamic Head mechanism to improve scale, spatial, and task awareness. Additionally, a new loss function, Wise-Inner-Shape IoU, is introduced to refine the localization of defects. The model's effectiveness is assessed utilizing a custom dataset of air-jet loom tire cord fabric, demonstrating a notable 4.0% enhancement over the original YOLOv8 architecture in the mean Average Precision (mAP) at a threshold of 0.5. The results show that the model is effective in achieving precise defect detection,which is essential for upholding quality standards of tire cord fabric in industrial production.**

***Index Terms*—Air-jet loom tire cord, Defect detection, YOLOv8 architecture, Industrial manufacturing**

I. Introduction

The thriving development of the automotive industry is closely linked to the increasing demand for travel, leading to a continuous rise in the per capita ownership of automobiles. As cars serve as the primary means of transportation, tires, as a crucial component, attract significant attention in the automotive industry. The quality of the cord fabric layer significantly influences the strength of the tire. Therefore, achieving efficient quality control during the production of cord fabric has become a widely discussed focus. Currently, most weaving workshops rely on manual visual inspections, which are inefficient, costly, demand high expertise, and are prone to errors, making it challenging to meet industrial requirements. Hence, there is an urgent need for automated solutions.

This work was supported by the Fundamental Research Funds for the Central Universities (Grant No. 2023ZYGXZR009), the National Natural Science Foundation of China (Grant No. 62173102), and the Guangzhou Science and Technology Project (Grant No. 2023B01J0037).

The development of machine vision technology has made image-based defect detection techniques crucial for enhancing production efficiency in various industries. In recent years, CNNs have been widely applied in image processing for their excellent feature extraction and generalization. Object detection methods, including both single-stage algorithms like YOLO [1], known for their efficiency and real-time capabilities, and two-stage methods like Faster R-CNN [2], which use a two-step strategy, have been integrated into defect detection by numerous researchers. These models are being applied in various industrial settings. For instance, Xie et al. [3] introduced a feature-enhanced algorithm FE-YOLO for industrial surface detection. Xu et al. [4] presented a technique for identifying flaws in metal surfaces using a modified YOLO model that incorporates features for small defects. Hu and Wang [5] proposed a Faster R-CNN based detection model for PCB defect detection. A two-stage algorithm for PCB defect identification was proposed by Hu and Wang [5]. Gao et al. [6] proposed a FCN-RCNN model for efficient and accurate detection of multiple tunnel defects in the complex subway environment.

To our knowledge,there is little research available on detecting cord defects in air-jet loom machine cord fabric. Tang et al. [7] introduced a method utilizing Gabor filters for identifying defects in cord fabric yarn. Considering the similarities in defect features between fabric defects and cord fabric defects, our research requires the exploration of pertinent fabric defect detection algorithms. Li et al. [8] introduced a broad yet efficient network for detection,which includes multi-scale analysis,filter factorization,pooling at multiple locations,and reducing parameters. Jun et al. [9] presented a framework for automatic fabric inspection through local defect prediction and global defect recognition steps. In order to tackle the problem of data imbalance in detecting fabric defects, Jing et al. [10] introduced Mobile-Unet for defect segmentation, using a median frequency balancing loss function.

Despite the notable results achieved in defect detection, there are still challenges to be addressed. Firstly, the defects in air-jet loom cord fabric exhibit wave-like features concentrated in the vertical direction of a single yarn and are often small in size. Secondly, to achieve reliable results, these algorithms necessitate a substantial quantity of annotated images, and

acquiring defect images in industrial settings poses difficulties.

Building upon these points, this paper addresses the practical challenge of defect detection in air-jet loom tire cord fabric and introduces an object detection model leveraging the YOLOv8 framework. Fig. 1 illustrates the setup of an Air-Jet Loom Tire Cord Inspection System in a factory environment. In order to maintain quality control throughout production, this system is made to monitor and identify flaws in tire cord textiles in real time. By leveraging our algorithm and detection system, the system enables accurate assessment of tire cord fabric quality, thereby enhancing product quality and production efficiency.

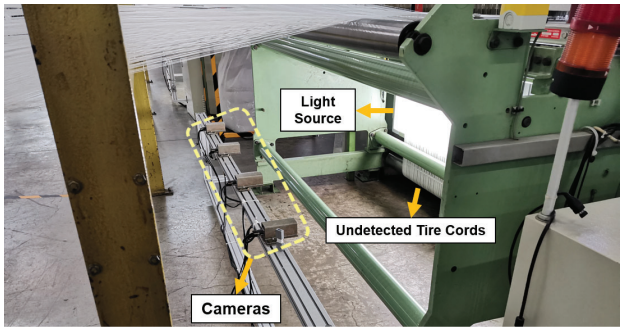

Fig. 1. Air-Jet Loom Tire Cord Inspection System Setup in Factory Environment

II. Method

A. Problem analysis

Air-jet loom tire cord is a special type of textile product, whose defect types are significantly different from those of common fabric defects, making the defect detection of tire cord fabric an important complement to the fabric defect detection area.

From the captured images, a series of cords can be observed, neatly arranged and extending from the left to the right of the image in a vertical direction, with tightly spaced intervals between the cords. Overall, the detection system must target seven main categories of defects, including abnormal threads, entangled threads, hand joints, air splices, spliced joints, insects, and cotton fluff, to ensure the tire cord fabric meets high-quality standards. Fig. 2 presents the defect categories and image characteristics observed in tire cord inspection. To provide readers with a clearer understanding of the characteristics of each defect, we selected images with more prominent features. It is crucial to note that defect images in practical situations may be more intricate. In the image's right portion, we showcase the challenges encountered in defect image detection.

The following challenges are encountered in designing algorithms for these cord defects:

- Complex Background: The tire cord fabric yarns are twisted from multiple fine yarns, horizontally arranged and dense, sometimes even partially overlapping, making the background information in the images exceptionally complex.
- Large Variations in Target Sizes: Some defect images may only exhibit nodal features at the endpoints of the yarn, while others may define the entire yarn as a defective area. Due to the high image resolution, there is a significant variation in the sizes of different defect targets, making it difficult to simply define defect areas using rectangular boxes.
- Indistinct Defect Features: The differences between defects and the background are subtle. For example, defects like entangled threads may only exhibit nodal features in small areas and are easily confused with other yarns. Even to the naked eye, it is challenging to accurately identify defects.

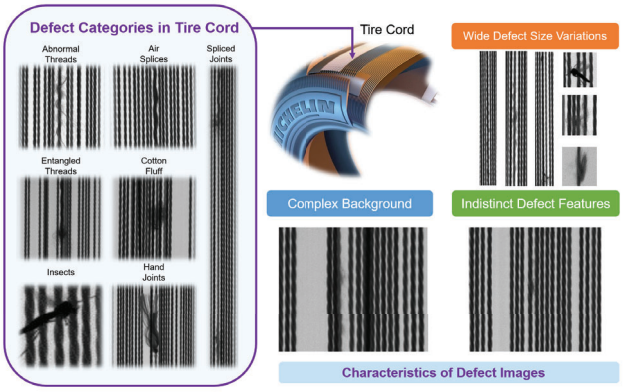

Fig. 2. Defect Categories and Image Characteristics in Tire Cord

Besides the challenges posed by defect features, the dataset dilemma is also a significant issue. Currently, there is no publicly available dataset for air-jet loom machine tire cord fabric, thus defect image data can only be obtained by continuously capturing images during the production process and manually selecting defective images from a large number of images. Moreover, accurate annotation of the data is required by process experts. Additionally, the probability of occurrence varies for each type of defect, leading to potential issues of sample imbalance in scarce data situations.

To tackle the previously listed obstacles and meet the high requirements of industrial defect inspection systems for real-time performance, this paper will introduce the YOLOv8 model as the foundational algorithm. Considering the specificity and complexity of this task, we have made corresponding improvements to YOLOv8 to achieve high-precision defect detection for tire cord fabric.

B. Improved YOLOv8 model

The YOLO series is widely used in the industry due to its accuracy and efficiency. YOLOv8 is the recent iteration of the YOLO series. The three primary components of YOLOv8 are the loss function, head, neck, and backbone.

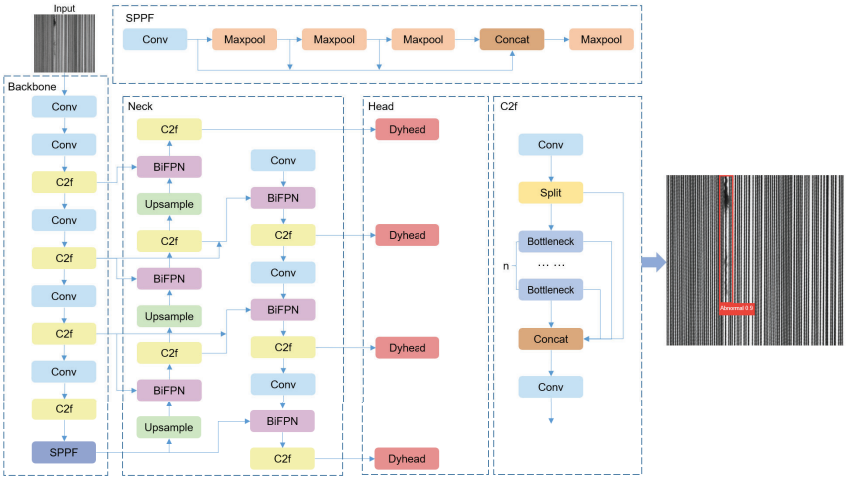

Fig. 3. Improved YOLOv8 network architecture

Firstly, the backbone part replaces the C3 module in YOLOv5 with the more efficient feature extraction C2f module, optimizing network size and performance through residual connections and bottleneck structures.

The neck part handles multi-scale feature maps from the backbone, including an SPPF module, a PAA module, and two PAN modules, constructing a feature pyramid through the combination of FPN and PAN.

The head part processes the three-layer feature maps from the Neck. The initial detection head splits into a detection and a classification branch, each with its own loss function. The classification branch adopts Varifocal Loss, while the regression branch uses CIoU Loss + Distribution Focal Loss format.

YOLOv8 offers models in the n/s/m/l/x scales. Considering the real-time requirements in industrial defect detection scenarios, the proposed method is based on the version of YOLOv8n. To tackle the intricate problems associated with air-jet loom tire cord defect inspection, this paper introduces Bi-directional Feature Pyramid Network (BiFPN) [11] into the neck part of YOLOv8n to construct a feature pyramid, enhancing the model's feature fusion capability. Meanwhile, the head part introduces Dynamic Head (Dyhead) [12] to strengthen the model's capabilities in Scale-awareness, Spatial-awareness, and Task-awareness. Furthermore, by combining the characteristics of Wise-IoU [13], Inner-IoU [14], and Shape-IoU [15], this paper introduces the Wise-Inner-Shape IoU loss function to achieve more precise defect localization. Fig. 3 displays the enhanced network architecture, showcasing the modifications and improvements made to the original design.

C. Improved Neck and Head

a) Intergrating BiFPN into neck layer: As mentioned earlier, YOLOv8 integrates FPN and PAN [16] to construct a feature pyramid, thus fully integrating multi-scale information. However, FPN and PAN alone cannot efficiently integrate features of different scales. PAN may miss certain small-sized target information when processing features from FPN, which would lead to inadequate small-sized target detection. BiFPN is an improvement based on PAN, which reconstructs the pathways and inserts two additional lateral connection paths, achieving more effective fusion and promoting the propagation of features' context. Its adaptive feature adjustment mechanism adjusts features between different levels according to task differences, thereby improving the effectiveness of feature fusion. Considering the significant differences in the sizes of various tire cord fabric defects, the introduction of BiFPN enables the model to fully understand defects of different sizes, improve the contextual understanding of defects, and avoid false positives or false negatives. Moreover, its simpler parameters and computational load make it suitable for the scenarios requiring rapid response and immediate processing. Details about the design of BiFPN are illustrated in Fig. 3

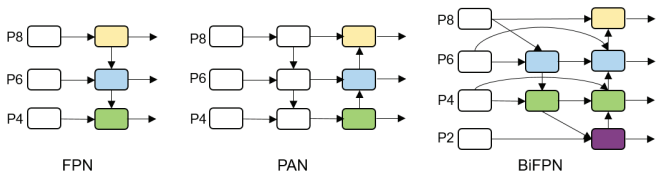

Fig. 3. Different feature network design: FPN, PAN and BiFPN

b) Intergrating Dyhead into head layer: Although the Decoupled-Head structure adopted by YOLOv8 improves detection accuracy, in complex detection scenarios of air-jet loom tire cord fabric, challenges such as fabric scale changes and possible overlapping adjacent threads exist. At this point, YOLOv8's detection head cannot fully perceive the multiscale, spatial, and task information.

To better adapt the model to the diversity of feature scale caused by different target sizes and learn the unique positional relationship features possibly implied by different shapes and

spatial positions of different targets, this paper introduces the Dynamic Head framework (Dyhead) into YOLOv8. Given the feature tensor $\mathcal{F} \in \mathcal{R}^{L\times S\times C}$, Dyhead transforms three types of attention into sequential arrangements, each attention focusing on one model, thus forming a nested attention function, described by (1):

$$W(\mathcal{F}) = \pi_c\left(\pi_s\left(\pi_L(\mathcal{F})\cdot\mathcal{F}\right)\cdot\mathcal{F}\right)\cdot\mathcal{F} \tag{1}$$

where W is the attention function, $\pi_L(\cdot)$, $\pi_S(\cdot)$, and $\pi_C(\cdot)$ represent the scale-aware attention module, spatial-aware attention module, and task-aware attention module, respectively. Fig. 4 depicts the structure of Dyhead.

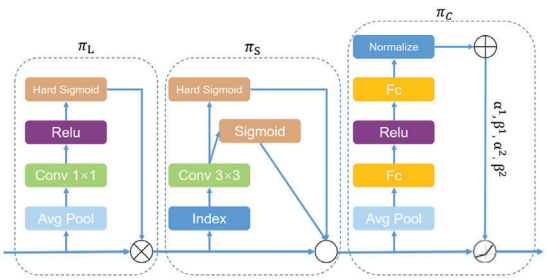

Fig. 4. Overview of Dynamic Head framework.

Dyhead integrates various attention mechanisms into the head,enhancing the detection head's perception capabilities in terms of scale,spatial location,and multitasking. This enables the detection of small-scale targets to be more flexible and precise. Enhancing the model's scale perception ability can better cope with the large differences in defect sizes caused by the ambiguity of defect regions and category differences. Spatial awareness refers to the deformation and rotation of objects under different viewpoints, with changes in their contours and positions. Strengthening spatial awareness can enhance the detector's generalization ability. Enhancing the model's ability to perceive tasks allows it to adjust to the various representations of different defects. To sum up, by introducing Dyhead, the model's adaptability to tire cord fabric defect detection tasks is enhanced, leading to improved accuracy.

D. Improved Loss Function

The regression loss function has a major effect on how well object detectors localize. IoU(Intersection over Union) measures bounding box accuracy against ground truth. In the regression loss of YOLOv8, CIoU is used instead, as illustrated in Fig. 5, where $l = \rho^2(b, b^{gt})$ represents the Euclidean distance between the centers of the two boxes, and c represents the diagonal distance of their minimum enclosing rectangle.

$$CIoU\ \ Loss = 1 - CIoU = 1 - \left(IoU - \frac{\rho^2\left(b, b^{gt}\right)}{c^2} - \alpha v\right) \tag{2}$$

$$v = \frac{4}{\pi^2}\left(\arctan\frac{w^{gt}}{h^{gt}} - \arctan\frac{w}{h}\right)^2 \tag{3}$$

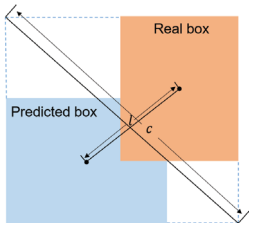

Fig. 5. Schematic diagram of CIoU

$$\alpha = \frac{v}{(1 - IoU) + v} \tag{4}$$

The calculation formulas are shown in Equations (2), (3), and (4). Here, w^{gt} and h^{gt} are the width and height of the ground truth box, while w and h are the width and height of the predicted box. v is a penalty term representing the disparity between two boxes, with a larger discrepancy leading to a larger v. α is the weight of v.

However, CIoU does not fully consider the influence of sample quality and may not sufficiently characterize the shape and size of defective target boxes. Therefore, introducing a loss function designed reasonably for task requirements can greatly improve the model's performance.

a) Wise-IoU: Complex detection needs may reduce localization effectiveness if loss functions are merely improved for high-quality sample fitting. In this regard, Wise-IoU introduces a dynamic non-monotonic focusing mechanism (FM). By more reasonably allocating small gradient gains, Wise-IoU enables bounding box regression by emphasizing anchor boxes of average quality, consequently minimizing the influence of low-quality samples on bounding box regression. Wise-IoU comprises three versions. The version we are utilizing here is Wise-IoU v3. Equations (5) and (6) provide the formula for WIoU:

$$\mathcal{L}_{WIoUv3} = r\mathcal{R}_{WIoU}\mathcal{L}_{IoU} \tag{5}$$

$$\mathcal{R}_{WIoU} = \exp\frac{(x - x_{gt})^2 + (y - y_{gt})^2}{(W_g^2 + H_g^2)^*} \tag{6}$$

The variables within $\mathcal{R}_{WIoU}$ are defined as follows: x and y represent the expected center coordinates of the predicted box, while x_{gt} and y_{gt} denote the coordinates of the real box's center. W_g and H_g represent the dimensions of the GT box, respectively. The r-values applied to $\mathcal{R}_{WIoU}$ represent non-monotonic focusing coefficients.

b) Inner-IoU: Inner-IoU considers the specificity of the target and flexibly uses supplementary bounding boxes with varying scales for various samples, thereby improving regression convergence efficiency. By using small auxiliary bounding boxes for samples with high IoU and larger auxiliary bounding boxes for samples with low IoU, the model's generalization ability for complex samples is enhanced. This approach is adopted to expedite network convergence. The formula can be found in Equations (7) to (13).

$$b_l^{gt} = x_c^{gt} - \frac{w^{gt} * ratio}{2}, b_r^{gt} = x_c^{gt} + \frac{w^{gt} * ratio}{2} \tag{7}$$

$$b_t^{gt} = y_c^{gt} - \frac{h^{gt} * ratio}{2}, b_b^{gt} = y_c^{gt} + \frac{h^{gt} * ratio}{2} \tag{8}$$

$$b_l = x_c - \frac{w * ratio}{2}, b_r = x_c + \frac{w * ratio}{2} \tag{9}$$

$$b_t = y_c - \frac{h * ratio}{2}, b_b = y_c + \frac{h * ratio}{2} \tag{10}$$

$$inter = (min(b_r^{gt}, b_r) - max(b_l^{gt}, b_l)) * (min(b_b^{gt}, b_b) - max(b_t^{gt}, b_t) \tag{11}$$

$$union = (w^{gt} * h^{gt}) * (ratio)^2 + (w * h) * (ratio)^2 - inter \quad (12)$$

$$IoU^{inner} = \frac{inter}{union} \quad (13)$$

Where the GT box and predicted box are denoted as b^{gt} and b respectively. (x_c^{gt}, y_c^{gt}) represents the center point of the GT box, while (x_c, y_c) denotes the center point of the anchor. The widths and heights of the GT box and anchor are represented by w^{gt}, h^{gt}, w, and h respectively. The variable ratio corresponds to the scaling factor, typically within the range of values [0.5, 1.5].

c) Shape-IoU: Although CIoU considers the shape similarity between ground truth and anchor boxes by adding a shape loss term on top of IoU, it still fails to fully characterize the effects of bounding box form and scale for different objects. Consequently, Shape-IoU calculates the loss by focusing on these impacts, aiming to improve localization accuracy. The formula for shape-iou is shown in Equation (14) to (20).

$$ww = \frac{2 \times (w^{gt})^{scale}}{(w^{gt})^{scale} + (h^{gt})^{scale}} \quad (14)$$

$$hh = \frac{2 \times (h^{gt})^{scale}}{(w^{gt})^{scale} + (h^{gt})^{scale}} \quad (15)$$

$$ww = \frac{2 \times (w^{gt})^{scale}}{(w^{gt})^{scale} + (h^{gt})^{scale}} \quad (16)$$

$$distance^{shape} = hh \times (x_c - x_c^{gt})^2 / c^2 + ww \times (y_c - y_c^{gt})^2 / c^2 \quad (17)$$

$$\Omega^{shape} = \sum_{t=w,h} (1 - e^{-\omega_t})^{\theta}, \theta = 4 \quad (18)$$

$$\omega_w = hh \times \frac{|w - w^{gt}|}{\max(w, w^{gt})} \quad (19)$$

$$\omega_h = ww \times \frac{|h - h^{gt}|}{\max(h, h^{gt})} \quad (20)$$

Here, the term "scale" denotes the scale factor, which correlates with the defects' scale in the dataset. Additionally, "ww" and "hh" represent the weight coefficients in the both directions, respectively, with values determined by the shape of the ground truth box. The following is the corresponding loss formula for it:

$$L_{Shape-IoU} = 1 - IoU + distance^{shape} + 0.5 \times \Omega^{shape} \quad (21)$$

d) Wise-Inner-Shape-IoU: We proposes a new loss function called Wise-Inner-Shape IoU, which combines three methods: Wise-IoU, Shape-IoU, and Inner IoU. Firstly, the use of Wise-IoU reduces the influence of low-quality samples on bounding box regression. This is crucial for handling common defects in cord fabrics. Secondly, Shape-IoU considers the influence of the shape and scale of the target box on localization accuracy. In cord fabric defect detection, considering shape similarity is crucial for accurate positioning due to the large differences in defect shapes and sizes. Finally, Inner-IoU introduces the concept of auxiliary bounding boxes, allowing the model to flexibly select appropriate scales to handle targets. This is highly effective for dealing with situations where there are significant scale differences in cord fabric defects because different defects may have different scales, requiring flexible adjustment.

By considering these three methods comprehensively, the Wise-Inner-Shape IoU loss function can better handle the diversity of scale, shape, and features in cord fabric defect detection tasks, thereby achieving more precise localization. Our suggested loss function is as described below:

$$\mathcal{L} = r \times R_{WIoU} \times (1 - IoU^{inner} + distance^{shape} + 0.5 \times \Omega^{shape}) \quad (22)$$

In the equation, r, R_{WIoU} , IoU^{inner}, $distance^{shape}$, $omega^{shape}$ have been previously introduced in the preceding text and are not reiterated here.

III. Experiment

A. Dataset and Experimental Setup

a) Dataset: As an essential part of factory production, the quality of dataset directly affects the final quality of the textiles. However, up to now, there has been no dataset specifically for air-jet loom tire cord fabric. In this study, we have constructed a self-built dataset focused on tire cord fabric.

The construction of the dataset was complex and challenging. Firstly, the scarcity of defective samples necessitates manual selection of a large proportion of unknown samples to gather sufficient defect images covering all categories. Secondly, professionals in the relevant domains were invited to annotate the photos in-depth in order to guarantee the correctness of the annotations and the quality of the dataset. As mentioned earlier, the constructed dataset includes seven categories. There are 722 high-quality photos total in the dataset, each with a resolution of 2048 × 1024 pixels. The dataset was split into training,validation,and test sets in an 8:1:1 ratio to evaluate the model's performance and improve training efficiency.

b) Experimental Setup: The experimental platform is equipped with a 12th generation Intel(R) Core(TM) i5-12400F CPU and an NVIDIX GTX 3060 12GB GPU. The software environment is Ubuntu 20.04, 64-bit, with Python 3.10 and the Pytorch 2.0.1 framework, using Cuda 11.8. The training parameters are as follows(Table I):

TABLE I
Training Parameters

Training Parameters	Values
Initial Learning Rate	0.01
Optimizer	SGD
Optimizer Momentum	0.937
Optimizer Weight Decay Rate	0.0005
Number of Images per Batch	16
Number of Epochs	300

c) Evaluation Metrics: The following metrics are mainly used in this experiment to assess the performance of the improved YOLOv8 on the air-jet loom tire cord fabric dataset:

Precision: The proportion of accurately identified samples to all samples that were detected. $\mathrm{Precision} = \frac{TP}{TP+FP}$

Recall: The proportion of samples in the test set that were successfully identified to all samples in the test set. $\mathrm{Recall} = \frac{TP}{TP+FN}$

Accuracy: The proportion of correctly identified samples among all samples. $\mathrm{Accuracy} = \frac{TP+TN}{TP+TN+FP+FN}$

The Precision-Recall (P-R) curve is the graph created by using precision as the y-axis and recall as the x-axis. The Average Precision (AP), a performance evaluation indicator for a model's performance on the PASCAL VOC dataset, is the area under the curve. The average precision for each category over the whole dataset is represented by the mean, or mAP, which is computed in Equation (23):

$$mAP = \frac{1}{N} \sum_{i=1}^{N} AP_i \quad (23)$$

B. Experimental Result

This paper's experiments are all based on the YOLOv8n model. To confirm the efficacy of the approach suggested in this paper, we carry out several ablation studies. The effects of the modifications we made to different model components are displayed in Table II.

TABLE II
ABLATION STUDY

Base model	BiFPN	Dyhead	Wise-Inner-Shape-IoU	Precision	Recall	mAP50
YOLOv8n	×	×	×	0.443	0.619	0.602
YOLOv8n	×	×	×	0.408	0.642	0.615
YOLOv8n	×	✓	×	0.549	0.624	0.617
YOLOv8n	✓	✓	×	0.548	**0.646**	0.631
YOLOv8n	✓	×	✓	0.446	0.653	0.627
YOLOv8n	×	✓	✓	0.635	0.467	0.628
YOLOv8n	✓	✓	✓	**0.555**	0.605	**0.642**

From the results, it is evident that the introduction of BiFPN improves feature fusion effectiveness and accuracy by optimizing cross-scale connections, allowing for more precise detection and localization of various industrial defects. This led to a 1.3% gain in mAP50 over the baseline. Integrating Dyhead enhanced the model's information utilization across feature levels and spatial positions, improving its adaptability for tire cord fabric defect detection. The introduction of Dyhead led to improvements in Precision, Recall, and mAP50, with mAP increasing by 1.5%. When both BiFPN and Dyhead were introduced simultaneously, the model's mAP50 showed a comparable increase of 2.9%. Additionally, we carried out comparative experiments using various combinations of modules along with the upgraded loss function. The results demonstrated that when BiFPN, Dyhead, and our Wise-Inner-Shape-IoU loss function were introduced, the best performance was attained by the model. Eventually, our model enhances the mAP50 by 4.0% compared to the baseline. This result demonstrates the effectiveness of our proposed enhancements, achieving accurate detection for tire cord fabric defects.

TABLE III
PERFORMANCE OF DIFFERENT LOSS FUNCTIONS

Model	IoU	Precision	Recall	mAP50
YOLOv8+BiFPN+Dyhead	CIoU	0.548	0.646	0.631
	WIoU	0.476	0.588	0.632
	InnerIoU	0.516	0.515	0.594
	Wise-Inner-MDPIoU	0.353	0.684	0.544
	Wise-ShapeIoU	0.446	0.653	0.627
	Wise-Shape-FocalIoU	0.393	**0.708**	0.589
	Wise-Inner-ShapeIoU	**0.555**	0.605	**0.642**

To demonstrate the effectiveness of the combined Wise IoU, Inner IoU, and Shape IoU loss function, we utilized the improved YOLOv8 model integrated with BiFPN and Dyhead as the baseline, and compared the results using different loss functions, as shown in the Table III. The combinations with MDPIoU [17] and FocalerIoU [18] are also tested here. From the experimental results, it is evident that when employing the Wise-Inner-Shape IoU loss function, the model can enhance its representation and generalization capabilities for defects of various sizes, ultimately yielding the most significant improvement in mAP50.

IV. CONCLUSIONS

In conclusion, this paper has successfully addressed the critical challenge of defect detection in air-jet loom tire cord fabric by proposing an enhanced YOLOv8 model. The integration of the BiFPN has significantly improved the model's capability to fuse features from different scales, enabling it to effectively identify and localize defects against complex backgrounds. The introduction of the Dynamic Head mechanism has further refined the model's perception of scale, spatial relationships, and task-specific features, leading to a more accurate and robust detection of various defect types.

The novel Wise-Inner-Shape IoU loss function has played a crucial role, leading to a significant improvement in localization precision. Through extensive experiments conducted on a meticulously curated dataset of air-jet loom tire cord fabric, we have validated the superior performance of the improved model. Achieving a comparable increase of 4.0% in relation to the YOLOv8 baseline model, our method demonstrates substantial improvements in defect detection performance.

ACKNOWLEDGMENT

This work was supported by the Fundamental Research Funds for the Central Universities (Grant No. 2023ZYGXZR009), the National Natural Science Foundation of China (Grant No. 62173102), and the Guangzhou Science and Technology Project (Grant No. 2023B01J0037).

REFERENCES

[1] J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, "You only look once: Unified, real-time object detection," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2016, pp. 779–788.

[2] S. Ren, K. He, R. Girshick, and J. Sun, "Faster r-cnn: Towards real-time object detection with region proposal networks," *Advances in neural information processing systems*, vol. 28, 2015.

[3] Y. Xie, W. Hu, S. Xie, and L. He, "Surface defect detection algorithm based on feature-enhanced yolo," *Cognitive Computation*, vol. 15, no. 2, pp. 565–579, 2023.

[4] Y. Xu, K. Zhang, and L. Wang, "Metal surface defect detection using modified yolo," *Algorithms*, vol. 14, no. 9, p. 257, 2021.

[5] B. Hu and J. Wang, "Detection of pcb surface defects with improved faster-rcnn and feature pyramid network," *Ieee Access*, vol. 8, pp. 108 335–108 345, 2020.

[6] X. Gao, M. Jian, M. Hu, M. Tanniru, and S. Li, "Faster multi-defect detection system in shield tunnel using combination of fcn and faster rcnn," *Advances in Structural Engineering*, vol. 22, no. 13, pp. 2907–2921, 2019.

[7] X. Tang, S. Yang, and D. Chen, "Study on defect detection of warp knitting front curtain cloth based on gabor filter(in chinese)," *Journal of Mechanical & Electrical Engineering*, vol. 34, no. 7, 2017.

[8] Y. Li, D. Zhang, and D.-J. Lee, "Automatic fabric defect detection with a wide-and-compact network," *Neurocomputing*, vol. 329, pp. 329–338, 2019.

[9] X. Jun, J. Wang, J. Zhou, S. Meng, R. Pan, and W. Gao, "Fabric defect detection based on a deep convolutional neural network using a two-stage strategy," *Textile Research Journal*, vol. 91, no. 1-2, pp. 130–142, 2021.

[10] J. Jing, Z. Wang, M. Rätsch, and H. Zhang, "Mobile-unet: An efficient convolutional neural network for fabric defect detection," *Textile Research Journal*, vol. 92, no. 1-2, pp. 30–42, 2022.

[11] M. Tan, R. Pang, and Q. V. Le, "Efficientdet: Scalable and efficient object detection," in *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, 2020, pp. 10 781–10 790.

[12] X. Dai, Y. Chen, B. Xiao, D. Chen, M. Liu, L. Yuan, and L. Zhang, "Dynamic head: Unifying object detection heads with attentions," in *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, 2021, pp. 7373–7382.

[13] Z. Tong, Y. Chen, Z. Xu, and R. Yu, "Wise-iou: Bounding box regression loss with dynamic focusing mechanism," *arXiv preprint arXiv:2301.10051*, 2023.

[14] H. Zhang, C. Xu, and S. Zhang, "Inner-iou: more effective intersection over union loss with auxiliary bounding box," *arXiv preprint arXiv:2311.02877*, 2023.

[15] H. Zhang and S. Zhang, "Shape-iou: More accurate metric considering bounding box shape and scale," *arXiv preprint arXiv:2312.17663*, 2023.

[16] S. Liu, L. Qi, H. Qin, J. Shi, and J. Jia, "Path aggregation network for instance segmentation," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2018, pp. 8759–8768.

[17] M. Siliang and X. Yong, "Mpdiou: a loss for efficient and accurate bounding box regression," *arXiv preprint arXiv:2307.07662*, 2023.

[18] H. Zhang and S. Zhang, "Focaler-iou: More focused intersection over union loss," *arXiv preprint arXiv:2401.10525*, 2024.