2025 5th International Conference on Artificial Intelligence and Industrial Technology Applications (AIITA)

Improving YOLOv8s through Attention Mechanism for Steel Surface Defect Detection

Wenyuan Liu*, Yuhang Luo, and Ruijie Lv
Rongcheng Campus, Harbin University of Science and Technology, Shandong, China
*Corresponding author's e-mail: 13563109396@163.com

***Abstract*—With the development of deep learning technology, object detection algorithms based on the YOLO series have been widely used in steel surface defect detection. As a lightweight target detection model, YOLOv8s stands out among many YOLO series models due to its high efficiency and real-time performance. However, YOLOv8s has insufficient detection accuracy for small targets in complex backgrounds and poor adaptability to defects of different scales. To address the above problems, we propose the CW-YOLOv8s model. The model introduces the Convolutional Block Attention Module (CABM) attention mechanism, which enhances the model's ability to extract defect features by combining channel attention and spatial attention. In addition, CW-YOLOv8s adopts the Wise-IoU (WIoU) loss function, evaluates the quality of anchor boxes by "outlier degree", and provides a wise gradient gain allocation strategy, which can reduce the negative impact of low-quality samples on the model while focusing on anchor boxes of ordinary quality. Experimental results show that CW-YOLOv8s significantly improves detection accuracy and robustness on multiple datasets, verifying the effectiveness of the model.**

Keywords-computer vision; steel surface defect detection; YOLOv8s; attention mechanism; loss function

I. INTRODUCTION

In modern industrial production, steel, as an important basic material, is widely used in various fields. Its surface quality is directly related to the service life, corrosion resistance and safety of the product [1-2]. Therefore, fast and accurate detection and classification of steel surface defects is a key link to ensure product quality. Traditional defect detection methods mainly rely on manual inspection or contact measurement technology, which are inefficient and low precision, and it is difficult to meet the requirements of modern industrial production for high efficiency and high precision [3].

With the rapid development of computer vision and deep learning technology, image-based automatic defect detection technology has gradually become a research hotspot [4-5]. These technologies use high-resolution cameras to collect images of steel surfaces and use advanced image processing and analysis algorithms to automatically identify and classify defects. In recent years, convolutional neural networks (CNNs) in deep learning and their derived target detection algorithms, such as the YOLO series, have shown great potential in steel surface defect detection due to their high efficiency and accuracy. For instance, zhao *et al*. [6] design a dual feature pyramid network to enhance feature representation by adopting Res2Net blocks as the backbone network to expand the receptive field and extract multi-scale features. Tian *et al*. [7] expand the receptive field by expanding the feature enhancement module, introduces a new center weight function to improve the accuracy of key point estimation, and uses CIoU loss to optimize size regression. Cheng *et al*. [8] combine anchor optimization based on differential evolution search, a novel channel attention mechanism, and an adaptive spatial feature fusion module to improve detection accuracy and feature integration. Tan *et al*. [9] propose the EfficientDet target detector series, which significantly improves the model efficiency by optimizing the neural network architecture through the weighted bidirectional feature pyramid network and compound scaling method. These algorithms can process image data in real time, quickly locate and identify defects, and significantly improve detection efficiency and accuracy [10-11].

However, defects have various shapes and sizes, ranging from tiny scratches to large-scale corrosion, which requires the detection algorithm to have strong feature extraction capabilities and adaptability [12]. In order to improve the accuracy of steel surface defect detection, this paper proposes the CW-YOLOv8s model. By introducing the Convolutional Block Attention Module (CBAM) attention mechanism, combined with channel and spatial attention, this model significantly enhances the ability to extract defect features [13], enabling it to more effectively cope with the detection challenges of complex backgrounds and small target defects. In addition, CW-YOLOv8s adopts the Wise-IoU (WIoU) loss [14]. This improved IoU loss function evaluates the quality of anchor frames through a dynamic non-monotonic focusing mechanism and "outlier degree", optimizes the gradient gain allocation strategy, reduces the interference of low-quality samples, and focuses on anchor frames of ordinary quality, thereby improving the robustness and detection accuracy of the model. This paper's main contributions are as follows:

- The CW-YOLOv8s model is proposed. It significantly enhances the model's ability to extract features of steel surface defects by introducing the CBAM attention and combining channel attention and spatial attention.
- The WIoU loss is used to reduce the negative impact of low-quality samples on model training. This loss based on the dynamic non-monotonic focusing mechanism evaluates the quality of anchor boxes through outliers.
- The experiment verifies the performance improvement of CW-YOLOv8s on multiple data sets, proving its effectiveness and robustness in steel surface defect detection tasks.

The rest of this paper is organized as follows: Section II proposes the CW-YOLOv8s model, Section III conducts experiments, and Section IV concludes this paper.

979-8-3315-0976-7/25/$31.00 ©2025 IEEE

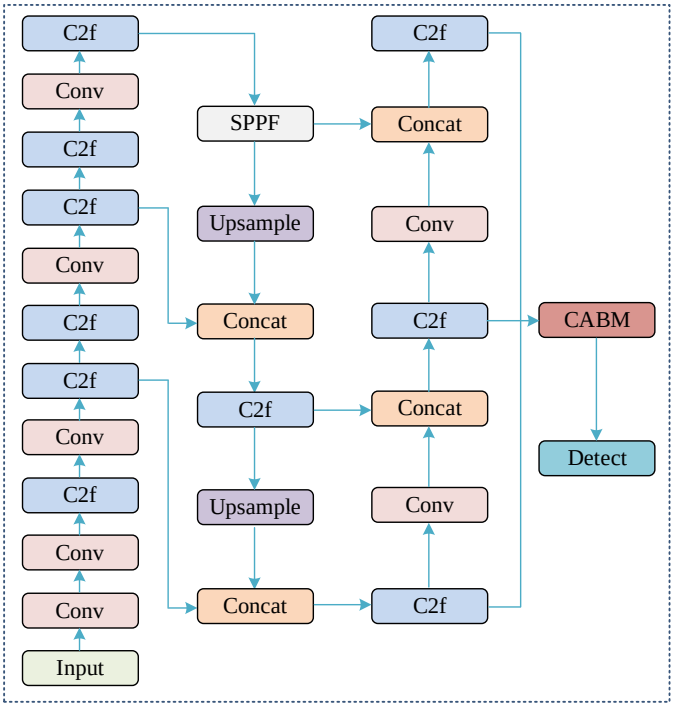


Figure 1. The structure of the proposed CW-YOLOv8s model.

II. METHODOLOGY

A. *YOLOv8s Network Architecture*

YOLOv8s is a lightweight version of the YOLO series of target detection models [15]. It inherits many innovations and improvements of YOLOv8 and aims to provide efficient and accurate real-time target detection capabilities. It adopts an anchor-free detection mechanism, abandoning the traditional anchor-based detection method, thereby improving the flexibility and accuracy of detection. As shown in Fig. 1, the backbone network of YOLOv8s introduces the C2f module, which can provide richer gradient flow information while maintaining lightweight, further improving the efficiency of feature extraction. In addition, its detection head adopts a decoupled head structure to separate the classification and detection tasks, making the network training and reasoning more efficient.

Although YOLOv8s has achieved significant improvements in performance and efficiency, as a general target detection model, it still has some limitations. For example, when faced with complex backgrounds or densely arranged targets, the model may have false detection or missed detection, especially in small target detection tasks. In addition, the performance of YOLOv8s depends heavily on the quality and diversity of the training data. If the data is not accurately labeled or the samples are unevenly distributed, it may lead to insufficient generalization ability of the model.

B. *The Proposed CW-YOLOv8s*

As a lightweight target detection model, YOLOv8s has significant advantages in speed and model size, but it still has some shortcomings in detection accuracy. Also, YOLOv8s

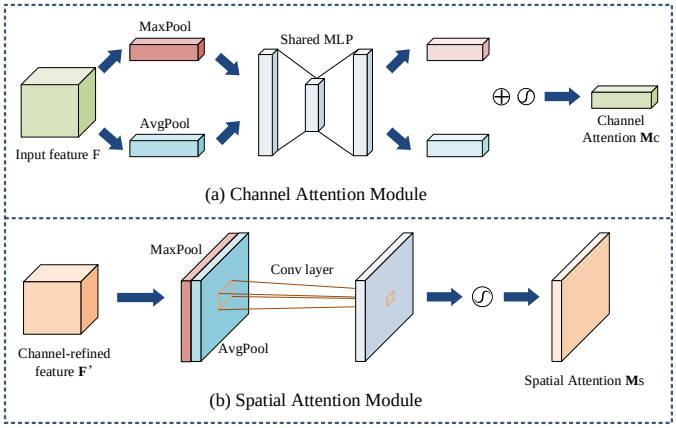


Figure 2. The structure of the CABM attention mechanism.

is highly dependent on the dataset and has poor adaptability to complex or small sample datasets. These shortcomings limit its application in high-precision detection tasks.

As shown in Fig. 2, the CABM attention mechanism is a dual attention module that combines channel attention and spatial attention [13], which aims to improve the performance of the model by enhancing the importance of key information in the feature map. In the YOLOv8s target detection model, the introduction of the CABM module can effectively make up for the lack of an explicit attention mechanism in the original architecture. Specifically, CABM extracts channel features through global average pooling and global maximum pooling, and uses a shared multi-layer perceptron to generate channel attention weights, thereby highlighting the features of important channels. Subsequently, it further performs weighted processing on the feature map in the spatial dimension, generates spatial attention weights through convolution operations, and emphasizes the key areas in the feature map. The output $M_c(F)$ of the channel attention module can be calculated by the following formula:

$$M_c(F) = \sigma\left(MLP(AvgPool(F)) + MLP(MaxPool(F))\right) \quad (1)$$

where F is the input feature map, *AvgPool* and *MaxPool* represent global average pooling and global maximum pooling operations respectively, *MLP* is a multi-layer perceptron, usually containing two fully connected layers, and σ is the Sigmoid activation function, which is used to normalize the output to the range [0, 1]. The output $M_s(F)$ of the spatial attention module can be calculated by the following formula:

$$M_s(F) = \sigma\left(f^{k\times k}([AvgPool(F); MaxPool(F)])\right), \quad (2)$$

where $f_{k\times k}$ is a convolution operation used to learn spatial attention weights from the concatenated average pooling and maximum pooling feature maps, and [;] represents the channel dimension concatenation of the feature maps. The CABM module combines channel attention and spatial attention. The specific calculation process is as follows:

$$\begin{aligned} F_{\text{channel}} &= F \times M_c(F), \\ F_{\text{output}} &= F_{\text{channel}} \times M_s(F_{\text{channel}}). \end{aligned} \quad (3)$$

This dual attention mechanism enables the model to focus more on the key features of the target and improve detection accuracy, especially in small target detection scenarios.

C. Loss Function

In the object detection task, the traditional IoU loss function and its variants (such as GIoU, DIoU, CIoU, etc.) are widely used to measure the similarity between the predicted box and the true box. However, these loss functions still have some limitations in practical applications. For example, when the predicted box does not overlap with the true box at all, the IoU loss returns 0, resulting in the disappearance of the gradient and ineffective guidance of model learning. In addition, these loss functions are easily disturbed when dealing with low-quality annotated data, and lack the ability to effectively distinguish between predicted boxes of different qualities, which makes it difficult for the model to balance the optimization of high-quality and low-quality predicted boxes during training.

To solve these problems, this paper adopts Wise IoU (WIoU) as an improved loss function [14]. The core idea of WIoU is to introduce a dynamic focusing mechanism to evaluate the quality of the anchor box by "outlier degree" instead of relying solely on the IoU value. Specifically, WIoU adjusts the loss function in the following way:

$$w_{i'} = \alpha \cdot w_i \cdot (1 - \mathrm{IoU}(b_i, g_i))^{\beta}. \tag{4}$$

This mechanism can dynamically adjust the degree of attention of the loss function to predicted boxes of different qualities, thereby optimizing the model more effectively. The calculation formula of WIoU is as follows:

$$\mathrm{WIoU} = \sum_{i=1}^{n} w_{i'} \cdot \mathrm{IoU}(b_i, g_i) \Big/ \sum_{i=1}^{n} w_{i'}. \tag{5}$$

WIoU effectively solves the limitations of traditional IoU loss in bounding box regression tasks by introducing weighted mechanism and dynamic focusing mechanism, and improves the performance and robustness of the model in object detection tasks. At this point, we have obtained the CW-YOLOv8s model proposed in this paper, which combines the CBAM attention mechanism and the WIoU loss function. The CBAM attention mechanism enhances the model's ability to focus on key features, while the WIoU loss optimizes the accuracy of bounding box regression.

III. Experiments and Results

A. Datasets

TABLE I. Hyper-parameter settings

Hyper-parameters	Values
Learning rate	0.04
Momentum	0.95
Weight decay	0.0005
Epoch	200
Batch size	64
Decay type	Cosine annealing

In order to verify the effectiveness of the CW-YOLOv8s model, we use the NEU-DET and GC10-DET datasets for experiments. The NEU-DET dataset covers six typical steel surface defects, namely cracks (cr), inclusions (in), patches (pa), pitted surfaces (ps), rolled scales (rs) and scratches (sc), with 300 images for each defect and an image size of 200×200 pixels. The training set and test set of this dataset are divided in an 8:2 ratio, that is, 1440 images are used for training and 360 images are used for testing. The GC10-DET dataset contains ten defect types, including punching (pu), welding seam (wl), crescent seam (cg), water spot (ws), oil spot (os), silk spot (ss), inclusion (in), rolling pit (rp), crease (cr) and waist fold (wf), a total of 2294 images, with an image size of 2048×1000 pixels. The training set and test set of this dataset are divided in a ratio of 9:1, that is, 2064 images are used for training and 230 images are used for testing.

B. Experimental Setup

The experiments reported in this study were performed on a system equipped with Windows 10, featuring an Intel Core i7-11700F CPU (operating at 2.50 GHz) and an NVIDIA GeForce RTX 3060 GPU, which has 12 GB of dedicated video memory. The development environment utilized PyCharm 2020, Python 3.6, and PyTorch 1.7, with GPU acceleration enabled through CUDA 11.2.0 and cuDNN 8.2.1. The training parameters for the model are outlined in Table I. Specifically, the learning rate was set to 0.04, the momentum value was 0.95, and the weight decay was 0.0005. The training was run for a total of 200 epochs, using a batch size of 64. The learning rate decay followed a cosine annealing schedule to optimize convergence.

C. Evaluation Metrics

To evaluate the performance of the proposed CW-YOLOv8s model, this study employs several widely-used metrics: Precision (P), Recall (R), Average Precision (AP), and mean Average Precision (mAP). The specific formulas for these metrics are provided below:

$$\begin{aligned} P &= \frac{TP}{TP + FP},\ R = \frac{TP}{TP + FN}, \\ AP &= \int_0^1 P(R)\mathrm{d}R,\ mAP = \frac{1}{C}\sum_{c \in C} AP(c), \end{aligned} \tag{6}$$

where TP signifies the count of positive samples correctly identified, FP indicates the number of positive samples misclassified as positives, and FN corresponds to the number of negative samples incorrectly labeled. Precision reflects the proportion of true positive predictions among all positive predictions, whereas recall measures the fraction of actual positives that were correctly detected. AP acts as a comprehensive measure to assess the detection performance for individual categories, with higher values suggesting superior performance for a specific category. mAP offers an aggregate evaluation of the model's overall performance across all categories. Model complexity is evaluated based on the parameter count and computational requirements. Typically, models with fewer parameters achieve faster detection speeds, often measured in frames per second (FPS), which indicates the number of frames processed within one second.

TABLE II. PERFORMANCE OF ALL MODELS ON THE NEU-DET DATASET

Models	AP/%						mAP	FPS
	cr	in	pa	ps	rs	sc		
YOLOv3	42.6	68	84.2	79.3	68.6	89.9	72.1	20
YOLOv5L	42.5	81.5	89.8	84.3	69.2	91.5	76.5	**98**
YOLOX [5]	56	85.7	94.6	87.2	57.5	**96.6**	79.6	46
CNNM [4]	58.6	84.1	91.3	81.6	68	94.3	79.7	31
Faster-RCNN [3]	43.2	67.2	84.8	79.5	67.6	90.2	72.1	21
EfficientDet [9]	45.6	62.5	83.5	84.4	71.2	73.1	70.1	15
DCC-CenterNet [7]	46.1	**90.1**	85.1	83	76.8	96.3	79.6	69
DEA-RetinaNet [8]	61.2	82.6	93.5	**95.8**	70.5	74.8	79.7	15
RDD-YOLO [6]	52.9	85.8	94.3	86.6	70.7	96.5	81.1	59
CW-YOLOv8s	**61.4**	87.3	**95.0**	90.4	**77.6**	85.9	**82.9**	71

TABLE III. PERFORMANCE OF ALL MODELS ON THE GC10-DET DATASET

Models	AP/%										mAP	FPS
	pu	wl	cg	ws	os	ss	in	rp	cr	wf		
YOLOv2	72.5	32.8	81.9	47.6	40.3	47.3	9.6	1.8	19.2	80.5	43.4	51
YOLOv5L	94.4	**97.2**	94.3	83.6	65.2	67.4	37.4	28.7	56.3	69.5	69.4	**100**
YOLOX [5]	**99.4**	94.1	97.7	86.5	67.8	69.6	36.8	48.8	58.2	86.5	74.5	45
EDDN [10]	90.1	88.5	84.6	55.7	62.1	65.2	25.8	36.2	52.0	91.8	65.2	30
Faster-RCNN [3]	89.8	55.3	87.1	60.0	65.4	57.8	19.5	36.5	**73.7**	81.7	62.7	24
RetinaNet [11]	76.5	94.5	98.6	84.3	60.2	66.5	34.2	48.8	37.5	89.0	69.0	45
VFNet-Res50 [12]	94.8	94.5	**99.1**	83.2	66.5	68.4	**39.2**	**55.1**	58.5	**93.0**	75.2	46
RDD-YOLO [6]	96.5	95.5	98.7	87.1	69.2	70.0	36.1	48.9	58.8	89.5	75.0	58
CW-YOLOv8s	93.2	96.7	97.8	**88.5**	**69.6**	**70.2**	38.3	49.3	60.2	90.5	**75.4**	61

D. Model Comparison Analysis

To evaluate the performance of CW-YOLOv8s, we compare with several widely used methods. Note that some of our baseline results are from [6], and the rest are obtained on our device. The detection results of CW-YOLOv8s compared with other baselines on NEU-DET and GC10-DET are shown in Tables II and III.

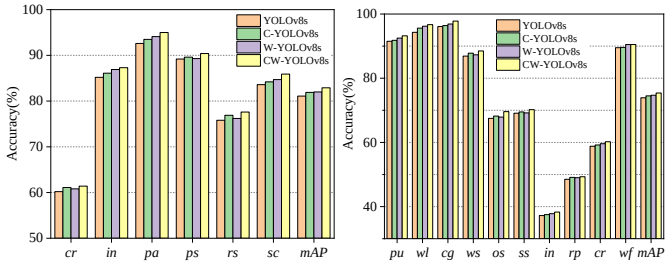
As shown in Table II, CW-YOLOv8s stands out with the highest overall performance. It achieves the highest AP for cr (61.4%) and pa (95.0%), and the highest mAP of 82.9%. Additionally, CW-YOLOv8s demonstrates a strong balance between accuracy and speed, with an inference speed of 71 FPS. This makes it highly suitable for real-time applications requiring both high accuracy and efficiency. Other models also show notable performance in specific areas. For example, DEA-RetinaNet achieves the highest AP for in (93.5%) and ps (95.8%), but its inference speed is relatively slow at 15 FPS. YOLOX and CNNM also perform well in several categories, with YOLOX achieving a high AP for sc (96.6%) and CNNM for pa (94.3%), but both models have lower inference speeds compared to CW-YOLOv8s. Traditional models like YOLOv3 and Faster-RCNN exhibit lower AP values and slower inference speeds, with YOLOv3 achieving an mAP of 72.1% at 20 FPS and Faster-RCNN achieving a similar mAP of 72.1% at 21 FPS. EfficientDet and DCC-CenterNet show promising results in terms of accuracy but are limited by their lower FPS, indicating a trade-off between accuracy and speed.

As shown in Table III, CW-YOLOv8s demonstrates a well-balanced performance, achieving a high mAP of 75.4% with an inference speed of 61 FPS. It shows strong AP values across multiple defect categories, such as pu (93.2%) and rp (60.2%), indicating its robustness in detecting a variety of defects. This combination of accuracy and speed makes CW-YOLOv8s a practical choice for real-time defect detection applications. Other models also exhibit notable strengths in specific areas. For example, YOLOX achieves the highest AP for pu (99.4%) and rp (58.2%), while VFNet-Res50 shows high AP values for wl (99.1%) and rp (55.1%). However, both models have lower inference speeds compared to CW-YOLOv8s, with YOLOX at 45 FPS and VFNet-Res50 at 46 FPS. Traditional models like YOLOv2 and Faster-RCNN generally have lower AP values and slower inference speeds. YOLOv2 achieves an mAP of 43.4% at 51 FPS, while Faster-RCNN reaches an mAP of 62.7% at 24 FPS. These models may struggle to meet the demands of real-time applications due to their slower processing speeds. RDD-YOLO and RetinaNet show promising results in terms of accuracy, with RDD-YOLO achieving a high mAP of 75.0% and RetinaNet reaching 69.0% mAP. However, both models have lower inference speeds, with RDD-YOLO at 58 FPS and RetinaNet at 45 FPS, indicating a trade-off between accuracy and speed.

Overall, CW-YOLOv8s stands out as a strong candidate for practical deployment in steel surface defect detection, offering a balance of high detection accuracy and efficient inference speed. Its performance across multiple defect categories and its relatively high FPS make it suitable for real-time applications where both precision and speed are critical.

E. Ablation Study

In order to verify the effectiveness of the designed module, we conducted a series of ablation experiments. Specifically, we used the initial YOLOv8s as a single baseline, and then added the CBAM attention mechanism to it as the C-YOLOv8s model. We add the WIoU loss to the YOLOv8s as the W-YOLOv8s model. The last one is the CW-YOLOv8s model proposed in this paper. Table III shows the experimental results.



(a) Results on NEU-DET (b) Results on GC10-DET

Figure 3. Experimental results of ablation study.

As shown in Fig. 3(a), the initial YOLOv8s model was used as a baseline, achieving an mAP of 81.1%. When the CBAM attention mechanism was added to YOLOv8s to form the C-YOLOv8s model, the mAP improved to 81.9%, with noticeable gains in the AP of several defect categories such as cr and pa. Adding the WIoU loss to YOLOv8s resulted in the W-YOLOv8s model, which further increased the mAP to 82.0%, showing improvements in the AP of in and ps. Finally, the proposed CW-YOLOv8s model, which integrates both the CBAM attention mechanism and the WIoU loss, achieved the highest mAP of 82.9%. This model demonstrated the best performance across all defect categories, with the highest AP for cr (61.4%), pa (95.0%), and ps (90.4%).

As shown in Fig. 3(b), taking the initial YOLOv8s model as the baseline, its mAP is 73.9%. On this basis, the CBAM attention mechanism is added to form the C-YOLOv8s model, and the mAP is increased to 74.5%, indicating that the CBAM module can enhance the model's ability to extract defect features. Furthermore, the WIoU loss is added to YOLOv8s to obtain the W-YOLOv8s model, and the mAP is increased to 74.7%, indicating that the WIoU loss plays a positive role in optimizing bounding box regression. Finally, the CW-YOLOv8s model proposed in this paper combines CBAM and WIoU, achieving the highest mAP of 75.4%, and achieving the best or near-best AP values on multiple defect categories, such as pu (93.2%), cg (97.8%), and rp (60.2%), which also proves the superiority of the CW-YOLOv8s model in improving detection accuracy. These results in highlight the significant contributions of both the CBAM attention mechanism and the WIoU loss in enhancing the detection performance of the YOLOv8s model.

IV. Conclusions

In view of the shortcomings of the YOLOv8s model in the steel surface defect detection task, this paper proposes an improved detection model, CW-YOLOv8s. By introducing the CBAM attention mechanism and combining channel and spatial attention, the model can more effectively extract defect features and significantly improve the detection ability of complex backgrounds and small target defects. At the same time, CW-YOLOv8s adopts the Wise-IoU (WIoU) loss function, optimizes the bounding box regression through a dynamic non-monotonic focusing mechanism, reduces the interference of low-quality samples on model training, and further improves the detection accuracy and robustness. Experimental results show that CW-YOLOv8s performs well on multiple datasets, verifying its effectiveness and superiority in the steel surface defect detection task.

Although CW-YOLOv8s performs well in steel surface defect detection, the inference speed may be affected to a certain extent when processing high-resolution images. Future work will focus on introducing more fine-grained feature extraction modules and multi-scale feature fusion mechanisms to improve small target detection capabilities, while exploring model compression techniques such as pruning and quantization to optimize inference speed [16].

References

[1] A. Sangaiah, F. Yu, Y. Lin, W. Shen ,and A. Sharma, "UAV T-YOLO-Rice: An Enhanced Tiny Yolo Networks for Rice Leaves Diseases Detection in Paddy Agronomy," *IEEE Transactions on Network Science and Engineering*, vol. 11, no. 6, pp. 5201-5216, 2024.

[2] W. Zhou, C. Cai, C. Li, H. Xu, and H. Shi, "AD-YOLO: A Real-Time YOLO Network With Swin Transformer and Attention Mechanism for Airport Scene Detection," IEEE Transactions on Instrumentation and Measurement, vol. 73, no. 5036112, pp. 1-12, 2024.

[3] S. Ren, K. He, R. Girshick, and J. Sun, "Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 39, no. 6, pp. 1137-1149, 2017.

[4] J. Xing and M. Jia, "A convolutional neural network-based method for workpiece surface defect detection," *Measurement*, vol. *176*, p. 109185, 2021.

[5] Z. Ge *et al*., "Yolox: Exceeding yolo series in 2021," *arXiv preprint arXiv:2107.08430*, 2021.

[6] C. Zhao, X. Shu, X. Yan, X. Zuo, and F. Zhu, "RDD-YOLO: A modified YOLO for detection of steel surface defects," *Measurement*, vol. 214, p. 112776, 2023.

[7] R. Tian and M. Jia, "DCC-CenterNet: A rapid detection method for steel surface defects," *Measurement*, vol. 187, p. 110211, 2022.

[8] X. Cheng and J. Yu, "RetinaNet with difference channel attention and adaptively spatial feature fusion for steel surface defect detection," *IEEE Transactions on Instrumentation and Measurement*, vol. 70, pp. 1-11, 2021.

[9] M. Tan, R. Pang, and Q. Le, "Efficientdet: Scalable and efficient object detection," *In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2020, pp. 10781-10790.

[10] X. Lv, F. Duan, J. Jiang, et al., "Deep metallic surface defect detection: The new benchmark and detection network," *Sensors*, vol. 20, no. 6, pp. 1562, 2020.

[11] T. Lin, P. Goyal, R. Girshick et al., "Focal loss for dense object detection," *In Proceedings of the IEEE International Conference on Computer Vision*, 2017, pp. 2980-2988.

[12] H. Zhang, Y. Wang, F. Dayoub et al., "Varifocalnet: An iou-aware dense object detector," *In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2021, pp. 8514-8523.

[13] S. Woo, J. Park, J. Lee, and I. Kweon, "Cbam: Convolutional block attention module," *In Proceedings of the European conference on computer vision*, pp. 3-19, 2018.

[14] 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.

[15] X. Liu, Y. Chu, Y. Hu, and N. Zhao, "Enhancing Intelligent Road Target Monitoring: A Novel BGS-YOLO Approach Based on the YOLOv8 Algorithm," *IEEE Open Journal of Intelligent Transportation Systems*, vol. 5, pp. 509-519, 2024.

[16] C. Schaefer, P. Taheri, M. Horeni, and S. Joshi, "The Hardware Impact of Quantization and Pruning for Weights in Spiking Neural Networks,"" *IEEE Transactions on Circuits and Systems II: Express Briefs*, vol. 70, no. 5, pp. 1789-1793, 2023.