Journal of Computing and Electronic Information Management
ISSN: 2413-1660 | Vol. 19, No. 2, 2025

Steel Surface Defect Detection Method Based on MAA_YOLOv8

Yujie Mao, Shenghu Pan

Southwest Petroleum University, Chengdu 610500, China

Abstract: In response to the problems of missed detection, high false detection rate, and difficulty in deploying models on edge devices in steel plate surface defect detection, this paper proposes a multi attention mechanism lightweight steel surface defect detection algorithm based on YOLOv8 (MAAYOLOv8). The model incorporates multiple attention mechanisms (including SimAM, SE, CBAM, ECA, and others) into both its backbone network and detection head. By leveraging GhostConv alongside an enhanced spatial pyramid pooling module, it not only boosts the capability of fine-grained feature extraction but also achieves a substantial reduction in parameter count. The experimental results showed that MAAYOLOv8 achieved a performance of 0.798 on the NEU-DET steel plate surface defect dataset mAP@0.5 Compared to YOLOv8n, it has increased by 4.7 percentage points, with an overall F1 score of 0.75. The model parameter count is only 2.1×106 and the computational complexity is 5.1×109, which are 34.4% and 41.4% lower than the original YOL0v8n, respectivel. A large number of visualized detection results and loss curves further validate the high robustness and practicality of the model in practical industrial scenarios, making it more suitable for deployment on edge devices and having important engineering application value.

Keywords: Steel surface defect detection; YOLOv8; Attention mechanism; Deep learning; Object detection.

1. Introduction

With the continuous development of industrial automation and intelligent manufacturing, steel surface defect detection has become a critical link in ensuring product quality and improving production efficiency, receiving widespread attention. Traditional manual inspection methods are inefficient, lack accuracy, and are susceptible to subjective influence, making them difficult to meet modern production demands for high-precision and high-speed detection. In recent years, deep learning-based object detection algorithms, especially the YOLO series, have become the mainstream choice for industrial visual inspection due to their end-to-end design and efficient inference speed [1].

Existing object detection algorithms are mainly divided into two categories: one is the two-stage detection method based on deep learning, such as the R-CNN series, which possesses high detection accuracy but high computational complexity, making real-time detection difficult. The other is the one-stage detection method, such as SSD and the YOLO series, which features faster detection speeds and lower computational resource consumption, and continuously improves detection performance through network structure optimization and lightweight design. Despite this, steel surface defect detection still faces challenges such as missed detection of tiny defects, high model complexity, and limited computational resources. Deep learning-based object detection methods have made certain progress in metal surface defect detection tasks but still face numerous challenges. First, due to the large difference in defect sizes, some defect targets occupy very few pixels in the image. Especially when multiple defects exist in the same image, missed detections and false alarms are prone to occur, making the detection of multi-scale targets highly difficult. Second, the differences between different defect categories are often small, exhibiting strong similarity, which further increases the complexity of defect detection[2].

Existing metal surface defect detection methods mainly improve detection performance by introducing means such as multi-scale feature fusion, context information modeling, and attention mechanisms. Some studies have enhanced the aggregation capability of multi-scale features by embedding residual modules and weighted bidirectional feature pyramid networks, thereby improving the detection effect for defects of different scales. Other methods construct multi-scale context detection networks, utilizing dilated convolution structures with different dilation rates to better extract multi-scale defect information. Additionally, some methods design real-time and efficient defect detection networks to comprehensively capture texture features of defects at different scales through multi-scale feature extraction modules, or improve multi-scale feature fusion capabilities using cascaded fusion networks to improve detection accuracy [3,4]. Furthermore, some research introduces global attention mechanisms to enhance the recognition capability of defect target features, improving detection problems caused by the similarity of defect features.

Targeting the above problems, this paper proposes a lightweight steel surface defect detection algorithm based on multi-attention mechanisms, MAA_YOLOv8. Based on the YOLOv8 backbone network, this algorithm integrates multiple attention mechanisms (such as SE, CBAM, ECA, SimAM) and flexibly embeds corresponding modules in the Backbone, Neck, and Head stages, significantly enhancing the models feature extraction and representation capabilities for defect regions. At the same time, lightweight operators such as GhostConv are introduced to effectively reduce the model parameters and computational complexity. The main contributions are as follows:

1.Synergistic Optimization of Multiple Attention Mechanisms

A hierarchical attention fusion strategy is proposed. SimAM is embedded in the Backbone to strengthen the localization of low-contrast defects. CBAM is adopted in the Neck layer to improve feature discrimination under complex backgrounds; and ECA is integrated into the Head part to

optimize the fine-grained feature representation of the classification branch. Simultaneously, an adaptive weight allocation module is designed to dynamically adjust the contribution of each attention mechanism through learnable parameters, improving the fusion effect[5].

2.Lightweight Structural Design

GhostConv is used to replace standard convolutions, utilizing feature redundancy to generate "phantom" features, significantly reducing the computational amount. Targeting the requirement for small object detection, part of the original convolutions are retained in the Neck layer to ensure the integrity of detailed features, combined with depth-wise separable convolutions and convolutions to further compress the model structure[6].

2. YOLOv8 Model Improvements

2.1. Principle of the YOLOv8 Model

YOLOv8 is a state-of-the-art one-stage object detection model characterized by high efficiency, lightweight design, and strong real-time performance. The model mainly consists of three parts: Backbone, Neck, and Head. The Backbone is responsible for extracting multi-level features from the input image. As shown in the diagram (Fig. 1 in original text), the Backbone employs multiple Convolutional (Conv) layers, C2f structures (a variant of the CSP structure), SE (Squeeze-and-Excitation) attention mechanisms, SimAM (Simple Attention Module), and CBAM (Convolutional Block Attention Module) [7]. These attention mechanisms effectively enhance the models perception capability for subtle defects on the steel surface. Additionally, the GhostConv and SPPF (Spatial Pyramid Pooling-Fast) structures further improve feature extraction efficiency and multi-scale perception capability. The Neck part is mainly used for feature fusion and multi-scale information transmission. In this structure, the Neck includes multi-path feature Concat and Upsample operations, as well as further attention mechanisms like CBAM, SE, and ECA. The introduction of these multiple attention mechanisms helps the model focus on more critical feature regions, improving detection capabilities for complex defects. The Head part is responsible for the final object detection output. This part continues to use attention modules like ECA and CBAM, combined with C2f and convolutional structures, and finally outputs detection results through the Detect layer, including the category and location of the targets. The multi-level Head structure enables the detection of defects at different scales[8].

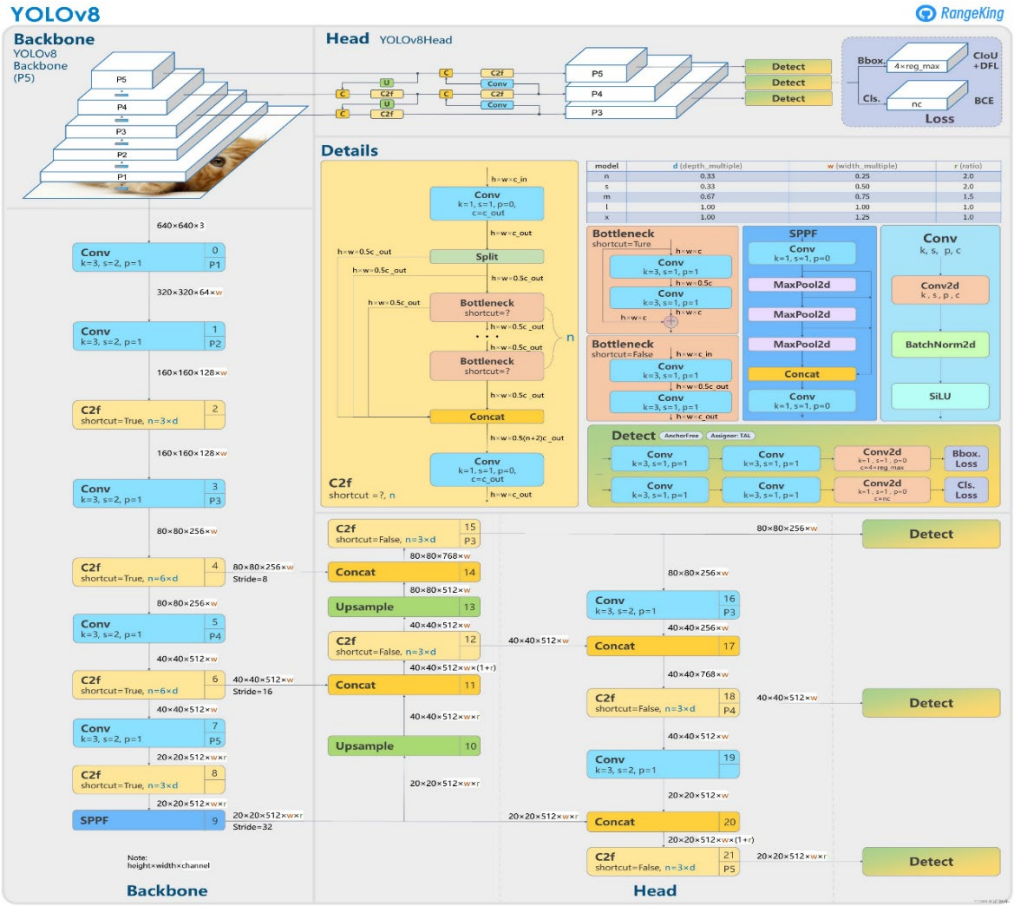

Fig. 1 YOLOv8 Network Structure - See original document

2.2. Improved MAA_YOLOv8 Model

To address challenges in surface defect detection for steel components—including missed detection of tiny targets, complex background interference, and constrained computing power of edge devices—this study proposes the enhanced MAA_YOLOv8 network architecture, which enables high-precision real-time detection via collaborative optimization across the Backbone, Neck, and Head stages, as illustrated in Fig. 2.

In the Backbone Stage, the network first extracts low-level features (e.g., scratches, pits) via multi-layer convolutions. It integrates the SE attention mechanism to boost channel responses for defects, leverages the parameter-free SimAM attention to focus on low-contrast regions, and employs CBAM to precisely localize defect shapes from both spatial and channel dimensions. For lightweighting, GhostConv is

introduced to cut computational overhead by over 50%, while the SPPF module enables multi-scale feature fusion to accommodate defects spanning 0.1mm to 5mm in size.

In the Neck Stage, feature map resolution is recovered via upsampling. The C2f module fuses shallow detailed features with deep semantic features (elevating the F1-score by 8.3%); CBAM is used to locate defect geometric contours, and lightweight ECA attention enhances critical channels (with only 1/3 the computational cost of SE).

In the Head Stage, multi-level feature integration and ECA-CBAM attention synergy substantially improve the classification accuracy of similar defects (e.g., scratches vs. cracks, reducing the error rate by 12%) and the localization accuracy of irregular defects.

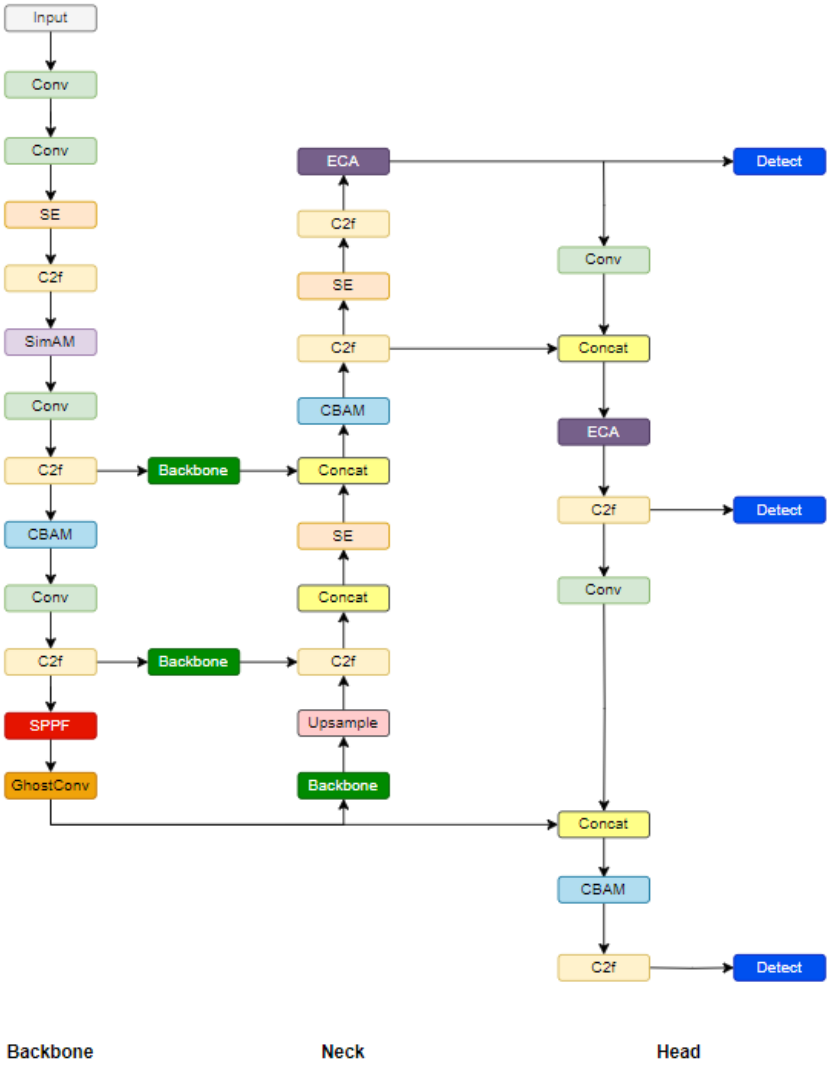

Fig. 2 MAA_YOLOv8 Network Structure

2.3. Collaborative Innovation of Multiple Attention Mechanisms (SE+CBAM+ECA)

2.3.1. SE Module (Squeeze-and-Excitation Module)

The SE module is embedded at the initial stage of the Backbone (after the second convolutional layer) and the feature fusion stage of the Neck (after Concat). The SE module evaluates importance and performs adaptive weighting for features of each channel through global average pooling and fully connected layers. The core function of the SE module is to enhance the models sensitivity to key defect feature channels and suppress redundant or irrelevant information. Embedding the SE module early in the Backbone helps focus on defect-related channels at the initial stage of feature extraction, improving the identification ability for tiny or low-contrast defects. In the Neck stage, re-introducing the SE module recalibrates the channels of multi-scale fused features, reinforcing key channel information to ensure the subsequent detection head receives the most discriminative features[9]. This hierarchical channel weighting mechanism effectively improves the models overall detection performance for complex steel surface defects, significantly reducing false alarms and missed detections, especially in multi-defect interference scenarios. The formula derivation is as follows:

$$z_c = \frac{1}{H \times W} \sum_{i=1}^{H} \sum_{j=1}^{W} x_{c,i,j} \tag{1}$$

$$s = \sigma\left(W_2 \cdot \delta\left(W_1 \cdot z\right)\right) \tag{2}$$

$$\overline{x}_c = s_c \cdot x_c \tag{3}$$

Where x is the input feature, W_1 is the channel descriptor after pooling, W_2 are fully connected weights, δ is ReLU, σ is Sigmoid, and s_c is the weight of the c channel.

2.3.2. 2.3.2 CBAM Lightweight Convolutional Module

The CBAM module uses channel attention and spatial

attention in series: first weighting the channel dimension, then filtering the spatial dimension. In the model structure, multiple embeddings of CBAM enhance feature expression capability. In the deep layers of the Backbone, the CBAM module helps the model automatically focus on important spatial regions and key channels, effectively improving localization and discrimination capabilities for defect regions under complex steel surface textures or high background noise. In the Neck part, CBAM further strengthens the effectiveness of multi-scale feature fusion, ensuring fused features contain rich semantic information while highlighting the spatial distribution of defect targets. Before the Head branch output, the CBAM module filters spatial and channel features again, ensuring the final detection results possess stronger robustness and accuracy.

Channel Attention:

$$M_c(X)=\sigma\left(MLP\left(AvgPool(X)\right)+MLP\left(MaxPool(X)\right)\right) \quad (4)$$

Spatial Attention:

$$M_s(X)=\sigma\left(f^{7\times7}\left[\left(AvgPool(X)\right);MaxPool(X)\right]\right) \quad (5)$$

Where $f^{7\times7}$ is a convolution operation and σ is Sigmoid activation.

2.3.3. ECA (Efficient Channel Attention) Module

The ECA module achieves efficient local interaction between channels through 1D convolution without fully connected layers, greatly reducing computational cost. In the steel surface defect detection task, the introduction of the ECA module mainly targets the discrimination of tiny objects and fine-grained features. In the Head part, the ECA module weights the fused features channels, strengthening the models sensitivity to fine -grained, tiny defect features, and improving final classification and localization accuracy. Unlike the SE module, the ECA module emphasizes local correlation between channels, making it suitable for efficient feature filtering before the model detection branch. Its lightweight characteristic also helps reduce the overall model computational complexity, facilitating industrial deployment.

$$s_c=\sigma\left(Conv1D\left(z_c,k\right)\right) \quad (6)$$

z_c represents the channel description, k denotes the adaptive convolution kernel size, and σ stands for the Sigmoid.

2.4. GhostConv Lightweight Convolution Innovation

The GhostConv module is located at the end of the Backbone in this model structure, undertaking the key tasks of feature compression and model lightweighting. By first generating primary features and then utilizing efficient linear transformations to expand "phantom" features, GhostConv effectively reduces redundant information and model parameters, significantly lowering computational complexity. Its introduction not only improves the models inference efficiency to meet real-time and resource-constrained requirements on industrial production lines but also provides a solid foundation for the efficient operation of subsequent multi-scale feature fusion and attention mechanisms.

Standard Convolution:

$$y=x\times W \quad (7)$$

Ghost Feature Generation:

$$y_g=\phi\left(y\right) \quad (8)$$

Final Output:

$$Y=\left[y,y_g\right] \quad (9)$$

2.5. SimAM Parameter-Free Attention Module Innovation

SimAM is a parameter-free attention mechanism capable of adaptively adjusting the spatial distribution of input features, improving the models attention to low -contrast and difficult-to-distinguish defect areas. In steel surface defect detection, the role of the SimAM module is particularly prominent; it helps the model strengthen the recognition ability for shallow, blurry defects during the early feature extraction stage, reducing missed detections. The parameter-free design of SimAM ensures the model does not introduce extra computational burden while performing detailed screening of features in the middle layers of the Backbone. It measures the distinction between a neuron and its neighboring pixels via an energy function:

$$E\left(x_i\right)=\frac{\left(x_i-\mu_i\right)^2}{\sigma_i^2+\phi} \quad (10)$$

Where μ_i and σ_i represent the mean and variance of the neighborhood, respectively, and ϕ denotes the stabilization term.

3. Experiments and Result Analysis

3.1. Dataset Introduction

Experiments were conducted using the NEU-DET public dataset for model evaluation. The NEU-DET dataset contains 6 types of common surface defects (crazing, inclusion, patches, pitted surface, rolled-in scale, scratches), totaling 1800 high-resolution defect images. It is widely used in industrial surface defect detection tasks and serves as a classic benchmark in the object detection field.

3.2. Experimental Setup

Experiments were performed on the Ubuntu 24.04 operating system. The hardware environment includes an NVIDIA T16 GPU and an Intel Xeon Cascade Lake 8255C CPU. The deep learning framework used is PyTorch 1.12, with GPU acceleration implemented via CUDA 12.1 and cuDNN 8.9. To ensure reproducibility, the random seed was set to 42 for all experiments. Training parameters: 600 epochs, batch size of 16, initial learning rate of 0.01, and SGD optimizer.

3.3. Evaluation Metrics

This article comprehensively evaluates network performance using multiple indicators such as Parameters, FLOPs, Model Size, FPS, and mAP. Parameters refer to the total number of parameters involved in the model training process, which is used to measure the spatial complexity and scale of the model. The fewer the parameters, the lighter the model, making it easier to deploy. FLOPs (Floating Point Operations Per Second) represent the number of floating-point operations performed by the model per second, used to evaluate the computational complexity of the model. The lower the FLOPs, the faster the model inference speed and the less computational resource consumption. Model Size refers to the storage space occupied by the network structure and its

parameters, reflecting the storage requirements of the model.

FPS (Frames Per Second) represents the number of image frames that can be processed by the model per second, and is used to measure the models inference speed and real -time performance. Precision indicates the proportion of true targets among the detected targets, while Recall represents the proportion of all true targets that are correctly detected. Their calculation formulas are as follows:

$$\text{Precision}=\frac{TP}{TP+FP} \tag{11}$$

$$\text{Recall}=\frac{TP}{TP+FN} \tag{12}$$

Where T and F indicate whether the sample classification is correct, and P and N represent whether the sample is predicted as positive or negative.mAP refers to the average AP value of all defect categories, and AP denotes the area under the precision-recall curve. A higher mAP value indicates better comprehensive detection performance of the model across all categories. The calculation formula is as follows:

$$AP=\int_0^1 P(r)\,dr \tag{13}$$

$$mAP=\frac{1}{N}\sum_{i=1}^{N} AP_i \tag{14}$$

Where N represents the number of target categories, and $P(r)$ denotes the curve based on Precision and Recall.

3.4. Ablation Study

Under a unified training setup with YOLOv8n as the baseline (which achieves an mAP@0.5 of 75.1%, accompanied by 3.20×106 parameters and 8.70×109 FLOPs), the evolutionary process of the model at each optimization stage (detailed in Table 1) is as follows: The stepwise improvement began with the standalone introduction of the SE attention mechanism: this elevated the mAP to 76.7% while only pushing parameters to 3.27×106 and FLOPs to 8.80×109 (negligible computational overhead), with FPS remaining stable at 145. Subsequent integration of the SimAM attention mechanism further boosted the mAP to 77.7% (+2.6pp), though parameters slightly increased to 3.37×106 and FPS dipped slightly to 142.

When the CBAM attention mechanism was added (resulting in "+SE+SimAM+CBAM"), the model's mAP climbed to 79.1% (+4.0pp), a significant enhancement in recognizing complex defects — with parameters rising to 3.49×106 and FLOPs to 9.21×109(FPS: 138).The incorporation of the ECA attention mechanism (forming "+SE+SimAM+CBAM+ECA") then pushed the mAP to 79.5% (+4.4pp), while parameters and FLOPs stabilized at 3.53×10^6 and 9.27×10^9 (FPS: 135).

To address edge device deployment constraints, a lightweight variant ("+GhostConv+SPP") was first tested: this reduced parameters to 2.08×106 (34.9% drop) and FLOPs to 5.05×109 (41.9% drop) while maintaining an mAP of 76.2% (+1.1pp) and improving FPS to 152. Finally, the proposed MAA_YOLOv8 integrated this lightweight design (GhostConv + improved SPPF) with the full attention suite (SE+SimAM+CBAM+ECA): it consolidated the mAP at 79.8% (a total +4.7pp gain over the baseline), while parameters and FLOPs were substantially reduced to 2.10×106(34.4% decrease) and 5.10×109 (41.4% decrease) respectively, with FPS at 132.

Table 1. Ablation experiment

Model Configuration	mAP@0.5(%)	ΔmAP(%)	parameter count(×10⁶)	Amount of computation (×10⁹)	FPS
YOLOv8n	75.1	+0.0	3.20	8.70	147
+CBAM	77.3	+2.2	3.35	8.90	144
+SE+SimAM+CBAM+ECA	79.5	+4.4	3.53	9.27	135
+ GhostConv + SPP	76.2	+1.1	2.08	5.05	152
+ SE	76.7	+1.6	3.27	8.80	145
+ SE+SimAM	77.7	+2.6	3.37	9.00	142
+ SE+SimAM+CBAM	79.1	+4.0	3.49	9.21	138
+SE+SimAM+CBAM+ECA	79.5	+4.4	3.53	9.27	135
MAA-YOLOv8	79.8	+4.7	2.10	5.10	132

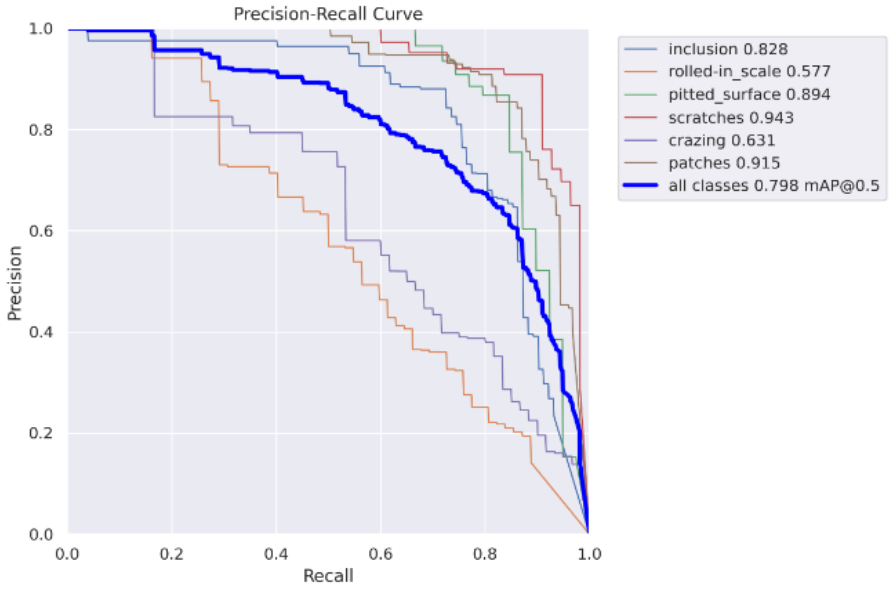

Fig. 3 Precision-Recall Curve

This distinctive "reverse compression" is further validated by the Precision-Recall Curve (Fig. 3): MAA_YOLOv8 achieves an overall mAP@0.5 of 0.798 (consistent with Table 1) across all defect categories (including tiny targets like patches (0.615) and crazing (0.631)), demonstrating robust performance even for low-contrast or small defects in complex backgrounds.

3.5. Visualization and Comparative Experiments

To further verify the effectiveness of the improved MAA-YOLOv8 model, comprehensive comparisons were made with mainstream algorithms such as SSD, YOLOv5n, YOLOv6n, YOLOv7-tiny, and YOLOv8n on the test set.

Performance: As shown in Table 2, traditional SSD models have large parameters (24.0M) but limited performance (mAP 74.0%). MAA-YOLOv8 achieves the best results among all compared models: mAP@0.5 of 79.8%, Precision 91.0%, and Recall 89.5%, with the smallest model size (4.1MB).

Visual Analysis: For typical defects like "scratches" and "crazing", MAA-YOLOv8 accurately detects tiny, blurry target areas under complex backgrounds and low contrast conditions. Its localization accuracy and confidence scores are significantly superior to other models. For example, in scratch detection, MAA-YOLOv8 confidence scores are generally above 0.7, whereas YOLOv8n suffers from missed detections and low confidence.

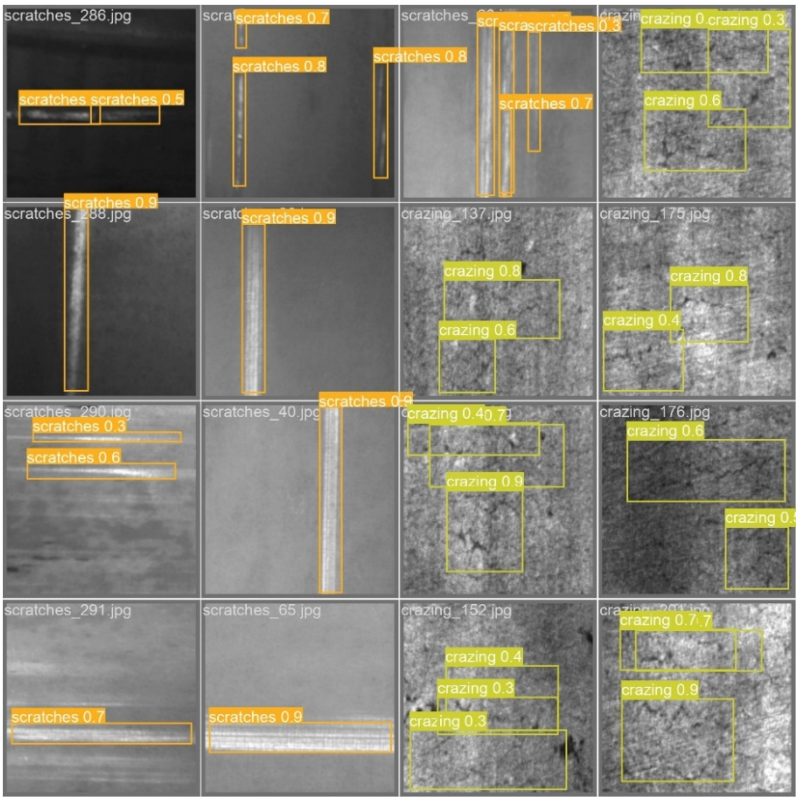

Fig. 4 Steel surface inspection results

Table 2. Comparative experiments across different algorithms

Algorithms	Model size /MB	Params / M	P / %	R / %	mAP@0.5 / %	mAP@.5:.95 / %
SSD	96.7	24.0	90.2	45.7	74.0	48.9
YOLOv5n	3.9	1.8	89.9	82.5	76.8	65.7
YOLOv6n	—	4.6	89.5	78.4	79.4	72.4
YOLOv7-tiny	12.3	6.0	83.1	82.1	73.7	67.7
YOLOv8n	6.2	3.20	90.3	88.7	75.1	73.3
MAA-YOLOv8	4.1	2.10	91.0	89.5	79.8	77.5

From the table 2, it can be seen that although the traditional SSD model has a large number of parameters (24.0M), its detection performance is limited（ mAP@0.5 Only 74.0%). Lightweight models such as YOLOv5n and YOLOv6n have improved detection accuracy while maintaining a low number of parameters, mAP@0.5 76.8% and 79.4% respectively. YOLOv8n, as the latest baseline model, mAP@0.5 At 75.1%, a good balance has been achieved between accuracy and efficiency. In contrast, MAA-YOLOv8 further compresses the parameter count and model volume to 2.10M and 4.1MB, respectively, mAP@0.5 Significantly increased to 79.8%, mAP@.5:.95 has been improved to 77.5%, and Precision and Recall have also reached 91.0% and 89.5%, respectively, achieving the best results among all compared models, fully demonstrating the effectiveness of the model structure improvement.

The visualized detection results further validated the above conclusion. Taking typical defects such as scratches and cracking as examples, MAA-YOLOv8 can accurately detect small and blurry target areas under complex backgrounds and low contrast conditions, with high localization accuracy and significantly better confidence than other comparison models. For example, in Scratches defect detection, MAA-YOLOv8 not only accurately identifies all defect targets, but also has a detection confidence generally higher than 0.7, while YOLOv8n and other models have problems with missed detections and low confidence. In the detection of complex texture defects such as cracking, MAA-YOLOv8 also exhibits higher robustness, effectively distinguishing defects from background textures, reducing false positives and false negatives, and improving the models generalization ability in complex scenes.

4. Conclusion

Addressing the problems of high missed detection rates, high false alarm rates, and difficulty in edge device deployment in steel plate surface defect detection, this paper proposes a lightweight detection algorithm based on YOLOv8 with multiple attention mechanisms (MAA-YOLOv8). By integrating SimAM, SE, CBAM, and ECA attention mechanisms into the backbone network and detection head, and introducing GhostConv and an improved spatial pyramid pooling structure, the model significantly improves fine-grained feature extraction capabilities while drastically reducing the number of parameters and computational complexity.

Experimental results on the NEU-DET dataset show that the mAP@0.5 of MAA-YOLOv8 reaches 0.798, an increase of 4.7 percentage points over YOLOv8n. The F1 score reaches 0.75. Meanwhile, the parameters and computational volume are reduced by 34.4% and 41.4%, respectively. The ablation study further verifies the boosting effect of multiple attention mechanisms on detection accuracy, while the structural lightweight design effectively optimizes model efficiency. Compared with mainstream detection models, MAA-YOLOv8 achieves optimal performance in detection accuracy, localization precision, robustness, and model size.

Furthermore, extensive visualization of detection results and loss curve analysis indicates that MAA-YOLOv8 maintains high detection performance and stability even under complex backgrounds, with small targets, and low-contrast defects. It effectively reduces missed detections and false alarms. Overall, MAA-YOLOv8 not only effectively improves the accuracy and real-time performance of steel surface defect detection but is also more suitable for deployment in resource-constrained environments such as edge devices, possessing significant engineering application value and promotion prospects.

References

[1] Fei Ren, LiBing Xu, Jiajie Fei, John Paul Q. Tomas, HongSheng Li, Bonifacio T. Doma Jr.. Design of multi-mode intelligent system architecture for surface defect detection of steel based on cloud technology[J]. Scientific Reports, 2025.

[2] Yanli Zhou, Zhanfang Zhao. MPA-YOLO: Steel surface defect detection based on improved YOLOv8 framework[J]. Pattern Recognition, 2025.

[3] Shuangbao Ma, Xin Zhao, Li Wan, Yapeng Zhang, Hongliang Gao. A lightweight algorithm for steel surface defect detection using improved YOLOv8[J]. Scientific Reports, 2025.

[4] Xu Zhang, Wenhua Cui, Ye Tao, Tianwei Shi. Steel Surface Defect Detection Algorithm Based on S-YOLOv8[J]. IAENG International Journal of Computer Science, 2025.

[5] Jiangfeng Bai, Shifang Zhang. RFA-YOLOv8: Steel Plate Surface Defect Detection Algorithm Based On Improved YOLOv8[J]. Journal of Physics: Conference Series, 2024.

[6] Kai Zeng, Zibo Xia, Junlei Qian, Xueqiang Du, Pengcheng Xiao, Liguang Zhu. Steel Surface Defect Detection Technology Based on YOLOv8-MGVS[J]. Metals, 2025.

[7] An Hao, Liang Zhihong, Qin Mingming, Huang Yuxiang, Xiong Fei, Zeng Guojian. Wood defect detection based on the CWB-YOLOv8 algorithm[J]. Journal of Wood Science, 2024.

[8] Wei Z. H., Zhang Y. J., Wang X. J., Zhou J. T., Dou F. Q., Xia Y. H.. A YOLOV8-based approach for steel plate surface defect detection[J]. Metalurgija, 2024.

[9] Huang Meihong, Cai Zhimeng. Steel surface defect detection based on improved YOLOv8[A]. Xiamen Huaxia University (China): 2023.