Manuscript
Click here to access/download;Manuscript;11.15.docx
Click here to view linked References
This is the Pre-Published Version.
This version of the article has been accepted for publication, after peer review (when applicable) and is subject to Springer Nature's AM terms of use (https://www.springernature.com/gp/open-research/policies/accepted-manuscript-terms), but is not the Version of Record and does not reflect post-acceptance improvements, or any corrections. The Version of Record is available online at: http://dx.doi.org/10.1007/s10845-024-02550-9.

Real-time detection of blade surface defects based on the improved RT-DETR

Dongbo Wu[1,5#], Renkang Wu[2,#], Hui Wang[3*], Zhijiang Cheng[4], Suet To[1**]

1. State Key Laboratory of Ultra-precision Machining Technology, Department of Industrial and Systems Engineering, The Hong Kong Polytechnic University, HongKong, China;
2. College of Electrical Engineering, Xinjiang University, Urumqi 830017, China;
3. Research Institute of Aero-Engine, Beihang University, Beijing 102206, China;
4. School of Intelligence Science and Technology, Xinjiang University, Urumqi 830017, China;
5. Institute for Aero Engine, Tsinghua University, Beijing 100086, China;
#These authors contributed equally to this work.
** Corresponding author. E-mail address: wanghuisx@buaa.edu.cn, sandy.to@polyu.edu.hk.*

Abstract: During the CNC machining, the blades exhibit various surface defects, including diverse morphologies and dimensions. Deep learning-based intelligent detection algorithms for the blade production line aim to improve computational efficiency and accuracy while minimizing model dimensions. This study proposes an enhanced blade detection method predicated upon a real-time detection transformer (RT-DETR) to detect blade surface defects precisely and efficiently during the blade production line. A dataset of blade surface defects in the blade machining process is first constructed, focusing on four surface defect types: gash, scratch, bruise, and pockmark. Secondly, the backbone network segment is substituted with an improved and more lightweight ResNet18 to optimize defect detection efficiency. The original feature fusion approach in RT-DETR is replaced by a Hierarchical Scale-based Feature Pyramid Network (HS-FPN) to enhance the model's capability of detecting blade surface defects across various scales. The inner-GIoU loss function is employed in RT-DETR to expedite model convergence and improve the accuracy of detecting minor surface defects. The results illustrate that the approach developed in this study raises the detection accuracy (mAP@0.5) by 3.5% and reduces the computational time required for detecting a single blade by 1.16 seconds compared to the traditional RT-DETR. This algorithm exhibits a relatively faster detection speed and higher accuracy in the automated real-time detection of blade surface defects.

Keywords: Blade surface defects; RT-DETR; Real-time detection; Detection speed; Detection accuracy.

1 Introduction

The compressor blade is a crucial functional component in a jet engine. It further compresses the airflow into the inner culvert, increasing pressure and temperature to meet the combustion chamber's requirements. The most crucial aspect of the compressor blade is its intricate profile torsion and slender blade thickness, which effectively mitigate power loss caused by airflow.

The research on blade manufacturing technology is quite comprehensive, encompassing adaptive CNC machining processing technology (Jiang et al., 2024; Wu et al., 2020a), machine tool and fixture design (Wu et al., 2020b), multi-energy field-assisted processing technology (Zhao et al.,2024a; Zhao et al.,2024b), material removal technology for difficult-to-process materials (Geng et al., 2022), machining deformation (Wang et al., 2021), machining error (Wu et al., 2021) and surface integrity control technology (Wu et al., 2023). These advancements aim to enhance the accuracy of blade processing and improve surface integrity. The complex structural composition and challenging working conditions necessitate the detection of blades during machining to assess any blade surface defects, which will avoid blade damage during operation (Wu et al., 2022).

Blade surface defects can be detected and identified through some nondestructive testing techniques, such as ultrasonic (Lee et al., 2022), eddy current (Li et al., 2016a), alternating magnetic fields (Li et al.,2016b), and infrared thermographic detection (Ciampa et al., 2018). The recent development of industrial cameras in conjunction with artificial intelligence methods for automatically acquiring and identifying blade surface defects has proven an effective approach. Researchers have proposed several methods for automatically obtaining and intelligent detection of blade surface defects (Li et al., 2022a; Shen et al., 2019; Zhang et al., 2024). Wang et al. (2024) developed a real-time X-ray imaging technique for automatically detecting surface defects on blades. Zhang et al. (2020) employed the YOLOv3 to identify blade defects effectively. Cheng et al. (2022) advanced this approach by enhancing the YOLOv4 model, incorporating migration learning tailored to the specific features of the blade camber dataset, along with various data augmentation techniques. The outcomes demonstrated that it is superior to the original YOLOv4 model in terms of evaluation indexes such as precision and recall rate. He et al. (2021) conducted a study that focused on real-time blade detection using the SW-YOLO model, improving detection effectiveness through algorithm enhancements. Zhang et al. (2021) proposed an improved methodology utilizing R-CNN cascade mask deep learning,

specifically tailored to detect edge failures in turbine blades. Hui et al. (2021) introduced an advanced YOLOv4 model for blade surface defect detection. This model could better distinguish between foreground and background. Meanwhile, utilizing bilinear interpolation (BI) to improve multi-scale feature fusion ability significantly boosts detection capabilities. According to Faster R-CNN, Yixuan et al. (2023) proposed a comprehensive set of detection methods. They developed a blade defect detection platform to acquire and identify blade surface defects.

These studies have contributed to the advancement of blade surface defect detection. However, further enhancements in efficiency and accuracy are necessary. In particular, the defect detection system utilized in the blade production line must possess adequate precision and efficiency while enabling real-time detection of minor defects. This is crucial as prolonged detection time and low efficiency of multiple processes can hinder the implementation of the entire production line. This paper presents a strategy for identifying defects on blade surfaces through an improved RT-DETR approach. The paper's organization is as follows: Section 2 outlines the modified RT-DETR, Section 3 details the experiment condition and procedures, Section 4 discusses the experiment's results, and Section 5 concludes.

2 An improved RT-DETR for the blade surface defect detection

2.1 The improved RT-DETR model

The RT-DETR is a real-time end-to-end object detector with higher speed and accuracy (Zhao et al., 2024c). It does not necessitate threshold filtering and non-maximum suppression (NMS). This algorithm adopts HGNet-v2 and an efficient hybrid encoder to facilitate the extraction and fusion of multi-scale features. Meanwhile, the IoU-aware query selection enhances the objective query initialization. These improvements allow RT-DETR to reduce computational costs and diminish parameter count.

The surface defects of jet engine blades exhibit significant scale variations and a high proportion of minor defects, leading to challenges in achieving accurate detection. Three improvements to the traditional RT-DETR allow for efficient and precise defect detection. Fig. 1 is the topology architecture of the improved network. Firstly, an improved ResNet18 is utilized instead of the original backbone network, aiming to decrease model parameters, increase detection speed, and minimize redundant

feature computation. Secondly, HS-FPN is utilized for feature selection and fusion, facilitating the model's multi-scale defect detection ability. Thirdly, the Inner-GIoU replaces the original GIoU as the loss function, realizing faster convergence and improving the model's minor defect detection ability.

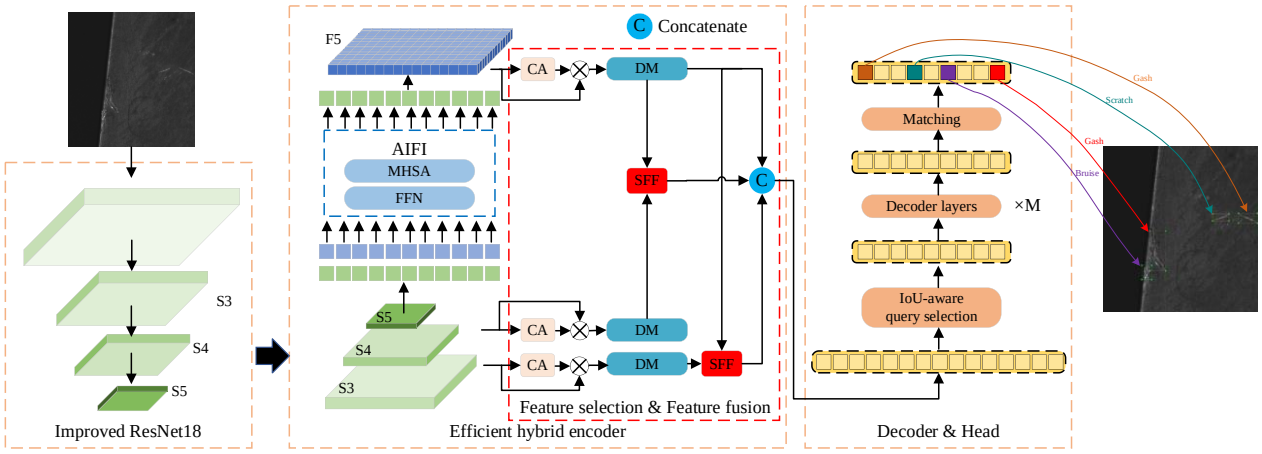

Fig. 1 Topology diagram of the proposed model.

The improved ResNet18 extracts the input image features to obtain feature maps S3, S4, and S5. Subsequently, S5 is converted into vectors and processed by lightweight Attention-based Intra-scale Feature Interaction (AIFI) (Zhao et al., 2024c) to generate F5. AIFI contains the feed-forward neural network (FFN) and multi-head self-attention (MHSA). F5 is fused with feature maps S3 and S4 through the HS-FPN. The IoU-aware query selection selects image features. Finally, after optimizing the object query multiple times, the decoder incorporates auxiliary prediction heads to generate the detection bounding box and confidence score (Liu et al., 2023).

2.2 An improved ResNet18

Models with larger parameters are often employed to achieve higher detection accuracy. However, obtaining sufficient high-quality typical defects for jet engine blade surface samples is challenging. Using larger models may pose a risk of overfitting, and the increased computational load will affect the detection's real-time performance. RT-DETR utilizes HGNet-v2 as its backbone network, which stacks excessive convolutional operations. This leads to high similarity among different channels during feature extraction, resulting in excessive redundant features and wastage of computational resources.

Partial convolution (Pconv) (Liu et al., 2018) is used for the convolution operation to reduce redundant features. It selects certain input channels for extracting spatial features, with the others left

unprocessed. The retained and convolved channels are concatenated as the input to get the output feature map, reducing the feature extraction's computational complexity. This study applies Pconv to the feature maps of 1/4 of the channels, while the remaining 3/4 remain unchanged. The backbone network employs a more lightweight ResNet18 (He et al., 2016) to enhance the network detection speed. The improved ResNet18 replaces the original convolution operations in the residual blocks with the FasterNet Block, which utilizes Pconv as the primary operator. Fig. 2 illustrates the structure of the improved ResNet18, FasterNet Block, and Pconv. FasterNet Block includes a Pconv layer, a pair of 1×1 Conv layers, batch normalization (BN), and the activation function ReLU. This topology structure decreases the parameter load of the model without compromising detection accuracy (Chen et al., 2023).

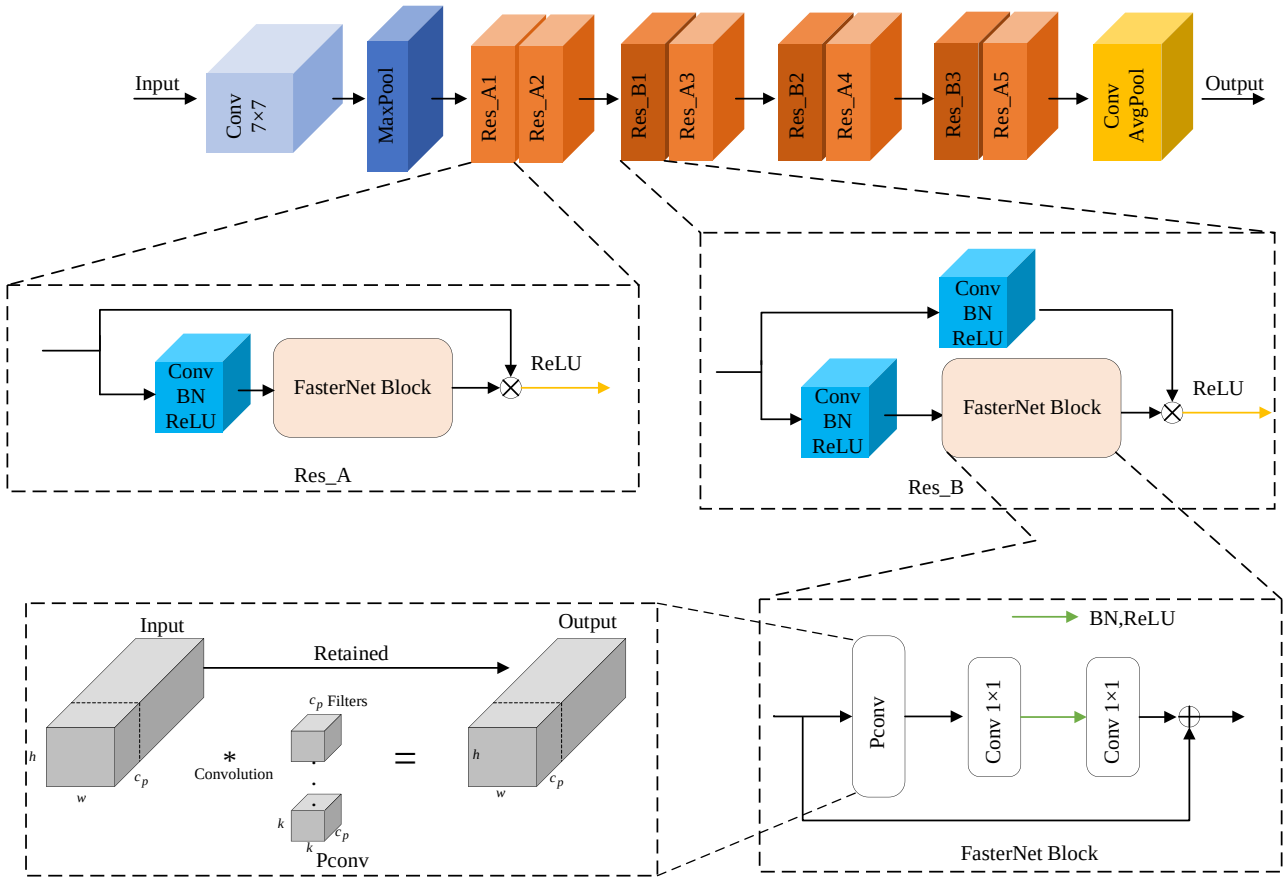

Fig. 2 The structural diagram of the improved ResNet18, FasterNet Block, and Partial convolution (Pconv) schematic illustrations.

2.3 A Hierarchical Scale-based Feature Pyramid Network

The dataset of blade surface defects presents a challenge due to its inherent multi-scale issues,

which impede the model's accurate identification. The same defect may exhibit varying sizes, while different defects demonstrate similar morphology under diverse imaging conditions. The improved ResNet18 extracts multi-scale features containing abundant semantic information at higher levels, yet the localization of targets remains relatively coarse. Conversely, low-level features could provide accurate localization information while presenting limited semantic context.

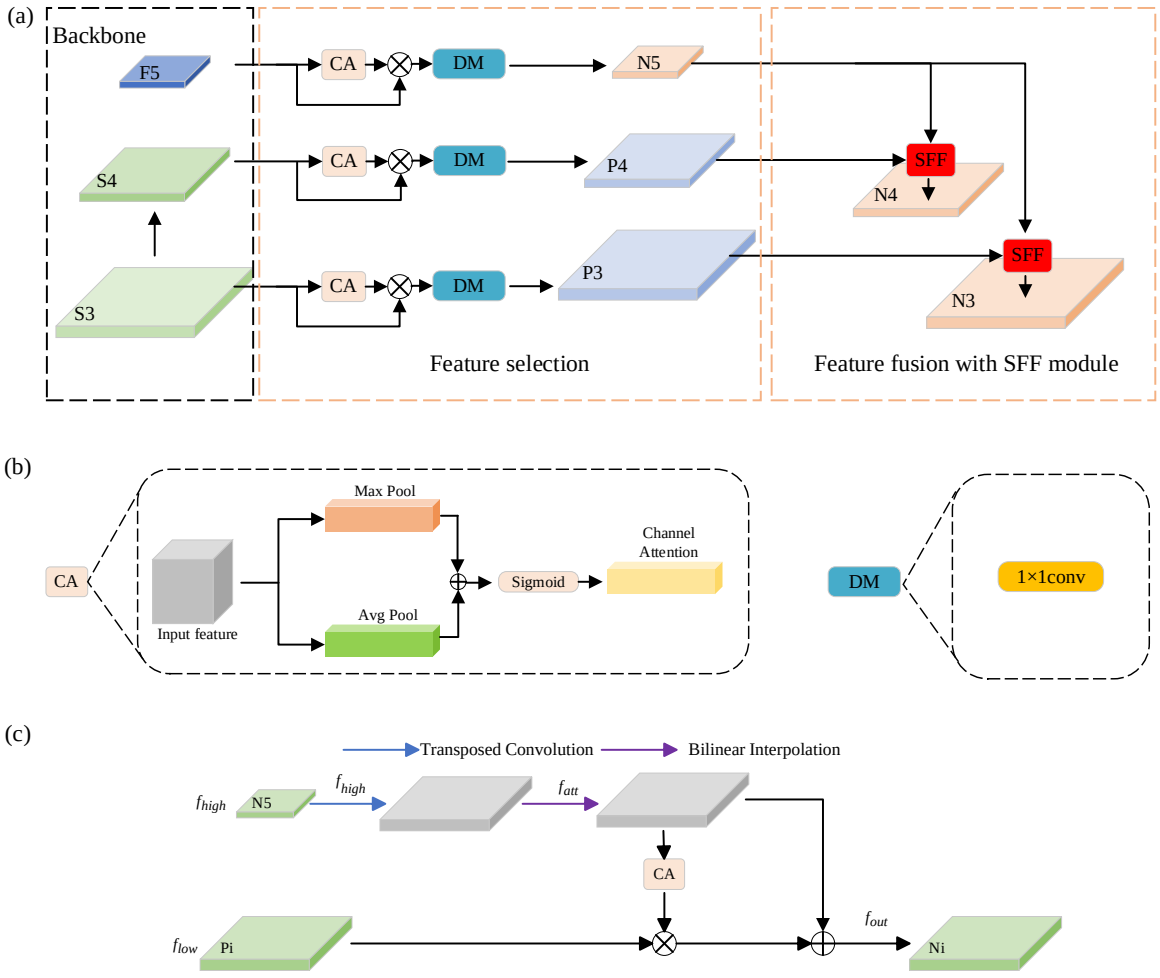

Fig. 3 The framework of the HS-FPN: (a) the HS-FPN topology architecture, (b) the channel attention (CA) mechanism and the dimensional matching (DM) module, and (c) the select feature fusion (SFF) module.

The typical approach to enhancing the semantic information within each layer is integrating pixel-level up-sampled high-level with low-level features. The traditional RT-DETR feature fusion module lacks feature selection and performs pixel addition across multiple feature layers. HS-FPN (Chen et al., 2024) improves the model's capacity for capturing defect feature information through multi-scale and selective feature fusion. Fig. 3(a) illustrates the topology structure of HS-FPN, which consists of a feature selection and a fusion module. The improved ResNet18 initially processes the input data to

generate multi-scale feature maps. Subsequently, the feature selection module selects and feeds them into the select feature fusion (SFF) module. This process produces semantically enriched features that improve the learning and understanding of the texture information about the blade surface defects.

Suppose that the feature map's scale input to the channel attention (CA) mechanism (Fig. 3(b)) is $f_{in} \in R^{C\times H\times W}$, which C represents the channel number, H and W indicate the height and width of the feature map, respectively. Global max pooling and average pooling extract information from the input feature maps. In this case, global max pooling extracts the most relevant information from each channel. In contrast, global average pooling enables uniform data collection. The combination of the generated features is then passed through the Sigmoid activation function to obtain the weight for each channel, $f_{CA} \in R^{C\times 1\times 1}$. The weight information is multiplied by the corresponding scaled feature map to generate the filtered feature map. Finally, the dimensional matching (DM) module (1×1 Conv) reduces the channels of the filtered feature map to 256.

The outputs from P3, P4, and N5 are put into the feature fusion module, as depicted in Fig. 3(c). Assume that the outputs of N5 are high-level features with a scale of $f_{high} \in R^{C\times H\times W}$, while outputs of P3 and P4 are low-level features with a scale of $f_{low} \in R^{C\times H_1\times W_1}$. The high-level features are up-sampled by applying a transposed convolution (T-Conv) with a stride of two and a 3×3 convolutional kernel, generating feature maps with increased size $f_{high} \in R^{C\times 2H\times 2W}$. Subsequently, to ensure dimensional consistency with the outputs of P3 and P4, the high-level features are either up-sampled or down-sampled utilizing BI, resulting in $f_{att} \in R^{C\times H_1\times W_1}$.

The high-level features are then transformed using the CA mechanism to generate corresponding attention weights, which are subsequently utilized for filtering the low-level features. As a result, features of identical dimensions are obtained. They are ultimately integrated to enhance the model's ability to represent features, resulting in a scale of $f_{out} \in R^{C\times H_1\times W_1}$. The SFF process is calculated as follows.

$$f_{att} = \mathrm{BI}\left(\mathrm{T\text{-}Conv}\left(f_{high}\right)\right) \tag{1}$$

$$f_{out} = f_{low} \times \mathrm{CA}\left(f_{att}\right) + f_{att} \tag{2}$$

2.4 Loss function optimization

The loss function usually utilized is intersection over union (IoU). It measures the overlap between the detection and ground truth (*gt*) boxes by calculating the proportion of their intersection to the union. However, if the distance between two bounding boxes is excessively large, the IoU will have a zero value. Traditional RT-DETR utilizes the GIoU loss function (Rezatofighi et al., 2019) instead of IoU to overcome this problem. The GIoU method employs the smallest enclosing rectangle to accurately measure the distance, effectively addressing gradient backpropagation failure when no intersection exists between the two boxes, as shown in Fig. 4. GIoU exhibits a faster convergence rate, enhancing model performance. Equation (3) illustrates how GIoU is calculated.

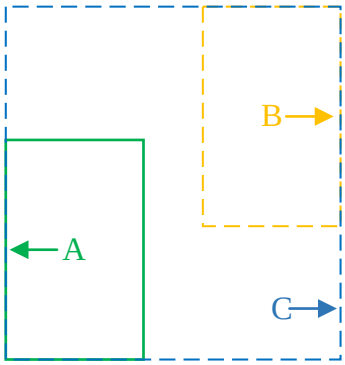

Fig. 4 Two non-intersecting rectangular boxes.

$$GIoU = IoU - \frac{Area(C) - Area(A \cup B)}{Area(C)}, GIoU \in (-1,1] \tag{3}$$

where C is the smallest enclosing rectangle of bounding boxes A and B. The GIoU is inversely proportional to the distance between A and B when they do not overlap.

In blade surface defect detection, the GIoU tends to become IoU when the detection and *gt* boxes exhibit a containment phenomenon. If two boxes intersect, the GIoU will converge slowly, horizontally, and vertically. The inner-IoU loss (Zhang et al., 2023) leverages auxiliary bounding boxes to overcome the limitations. Larger auxiliary boxes accelerate regression for samples with low IoU, while smaller auxiliary boxes facilitate the regression for samples with high IoU, resulting in quicker convergence. Based on this, Inner-IoU imports a scaling factor 'ratio' to adjust the dimensions of auxiliary bounding boxes, facilitating more efficient and rapid regression across various datasets and detectors. Fig. 5 demonstrates the computation process of Inner-IoU.

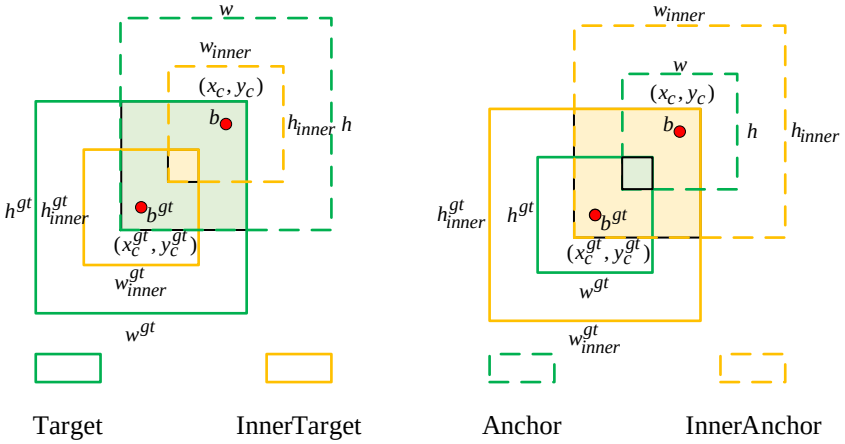

Fig. 5 The Inner-IoU computation process.

Equations (4) to (11) present the calculations for employing Inner-IoU within the GIoU loss function.

$$b_l^{gt} = x_c^{gt} - \frac{w^{gt} * ratio}{2}, b_r^{gt} = x_c^{gt} + \frac{w^{gt} * ratio}{2} \tag{4}$$

$$b_t^{gt} = y_c^{gt} - \frac{h^{gt} * ratio}{2}, b_b^{gt} = y_c^{gt} + \frac{h^{gt} * ratio}{2} \tag{5}$$

$$b_l = x_c - \frac{w * ratio}{2}, b_r = x_c + \frac{w * ratio}{2} \tag{6}$$

$$b_t = y_c - \frac{h * ratio}{2}, b_b = y_c + \frac{h * ratio}{2} \tag{7}$$

$$inter = \left(\min\left(b_r^{gt}, b_r\right) - \max\left(b_l^{gt}, b_l\right)\right) * \left(\min\left(b_b^{gt}, b_b\right) - \max\left(b_t^{gt}, b_t\right)\right) \tag{8}$$

$$union = \left(w^{gt} * h^{gt}\right) * \left(ratio\right)^2 + \left(w * h\right) * \left(ratio\right)^2 - inter \tag{9}$$

$$IoU^{inner} = \frac{inter}{union} \tag{10}$$

$$L_{Inner-GIoU} = L_{GIoU} + IoU - IoU^{inner} \tag{11}$$

Inner-GIoU focuses on accurately representing the bounding boxes, enabling a more precise assessment of their overlapping region. It provides greater accuracy and detail and emphasizes the effectiveness of detecting the target's central point. Utilizing ‘ratio’ to implement the scaling function of the auxiliary box allows Inner-GIoU to adjust dynamically for varying tasks and targets, improving the detection accuracy and efficiency of the algorithm.

3 Experiment condition and method

3.1 Blade surface defect detection process

Fig. 6 illustrates the detection process for surface defects on the blade. Initially, a dataset focused on surface defects in jet engine blades is constructed, and the improved RT-DETR is trained to obtain the model. Subsequently, the trained model undergoes evaluation using various metrics to get the training results. In the defect detection phase, the industrial computer controls the motion controller and guides the camera in capturing images of the blade's surface for detection. The captured images are then segmented into 80 images of 916×960 pixels. Subsequently, these segmented images are fed into the trained and improved RT-DETR model. The resulting outputs are stitched together according to the previous segmentation rules, achieving precise defect detection and localization on the complete blade image.

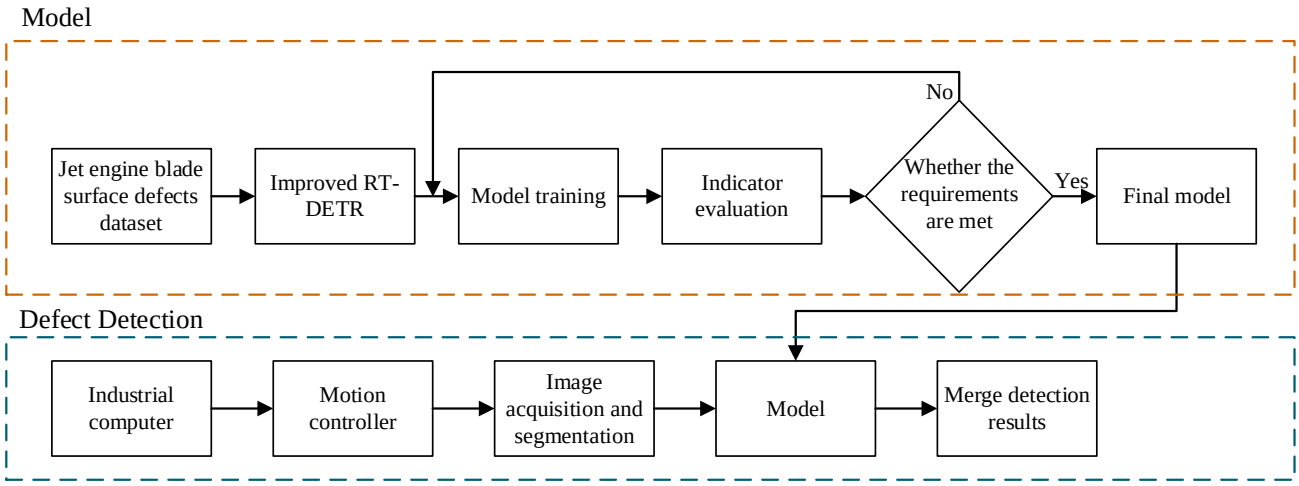

Fig. 6 Blade surface defect detection process.

3.2 Experiment setup

The experimental equipment utilized in dataset construction is a four-degree-of-freedom blade surface defect-specific data acquisition device developed by our research team, as shown in Fig. 7. This platform comprises components such as motion rails in the X, Y, and Z directions, a motion controller, an industrial camera, an annular light source, a workpiece rotation table, and an industrial computer. In the detection process, the industrial computer commands the motion controller through a serial port. Then, the controller controls the control motors of each coordinate axis and the workpiece

rotation table via the controller area network bus. The operating speed of each motion axis rail is 10 mm/s. Finally, the industrial camera captures clear pictures and transmits real-time blade surface data to the computer via the Gigabit Ethernet protocol. Table 1 presents the configuration of the algorithm environment.

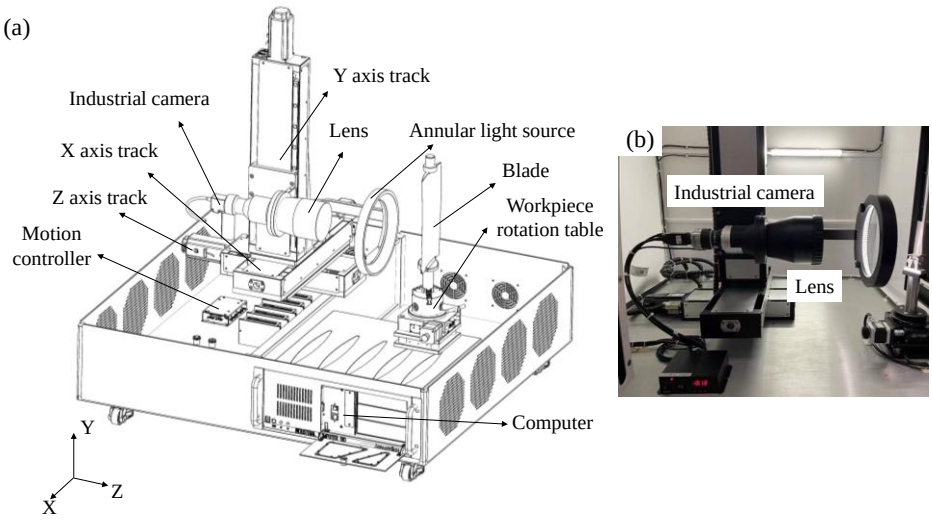

Fig. 7 Data acquisition equipment (Yixuan et al., 2023), (a) the complete model of the equipment, and (b) the built-up physical equipment.

Table 1 Experimental environment configuration and corresponding value.

Environmental configuration	Value
Operation system	Ubuntu 20.04
CPU	Intel(R) Xeon(R) Platinum 8352V @ 2.10GHz
GPU	NVIDIA GeForce RTX 4090
Graphics memory	24 GB
Deep learning framework	PyTorch 2.0.0
Parallel computing platform	CUDA 11.8
Input image size	640×640
Epoch	150
Batch size	32

3.3 Experimental dataset construction

We collected 72 images of jet engine blades, each with a resolution of 7344×16488. Each complete blade image was cropped into 80 surface defect images with a resolution of 918×960.

Images containing defect information were selected, invalid data were filtered out, and the images were enhanced. After careful detection, 5,461 images were selected for experimentation. The blade surface defects dataset includes four types: gash, scratch, bruise, and pockmark. Multiple defects may exist on a single image simultaneously.

Fig. 8 demonstrates that gashes and scratches are relatively prominent, whereas bruises and pockmarks appear comparatively smaller. Regarding pixel (px) area, the pockmark represents only 0.0005% of the entire field of the blade image, while a larger gash covers only 0.2286%. Fig. 9 provides an overview of the type and quantity of labels. A total of 4,856 gashes, 4,005 scratches, 1,217 bruises, and 3,668 pockmarks are recorded. The dataset was partitioned into a training, a validation, and a test in the ratio of 8:1:1. Specifically, the training, validation, and test sets had 4357, 558, and 546 images. The data annotation process was carried out using LabelImg.

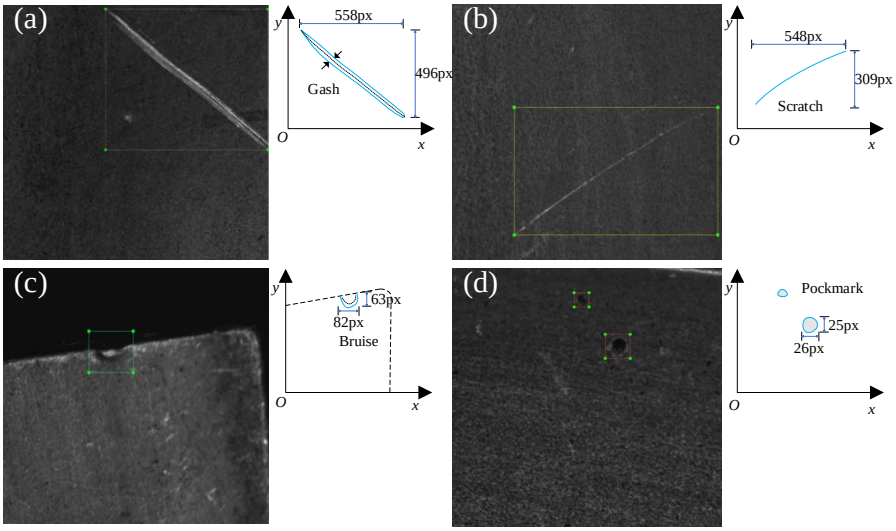

Fig. 8 Four typical defects of blade surface: (a) Gash, (b) Scratch, (c) Bruise, (d) Pockmark.

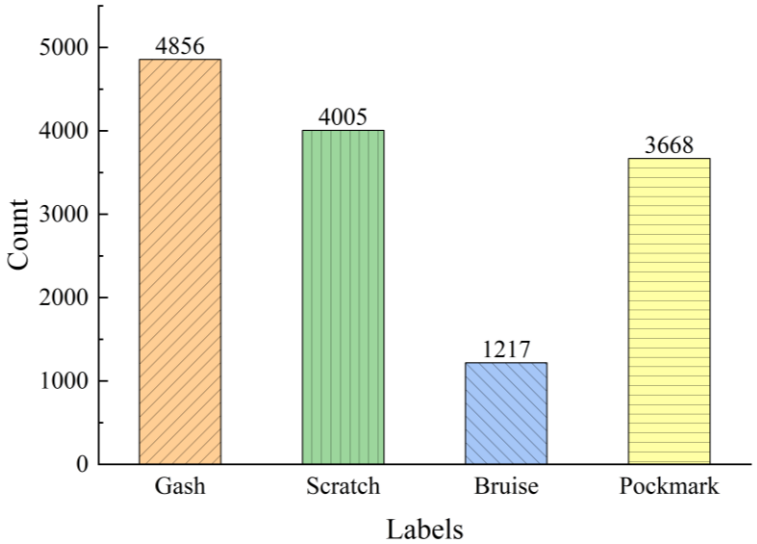

Fig. 9 Type and quantity of four labels.

3.4 Model assessment metrics

The metrics listed below are employed to assess the model performance.

(1) Precision

Precision is the percentage of positive samples predicted by the model that are correctly predicted. Its expression is shown in Equation (12).

$$P = \frac{TP}{TP + FP} \times 100\% \tag{12}$$

(2) Recall

Recall represents the ratio of correctly predicted positive samples out of positive samples, as formulated in Equation (13).

$$R = \frac{TP}{TP + FN} \times 100\% \tag{13}$$

(3) Mean average precision

Average precision (AP) refers to the average precision values obtained across all recall rate values. Mean average precision (mAP) is calculated as the mean AP across all classes, as Equation (14) defines.

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

4 Results and discussion

4.1 Backbone network optimization experiment

The proposed backbone network's effectiveness was validated through a comparative experiment involving HGNet-v2, ResNet18, and the improved ResNet18. During these experiments, mAP@0.5 denotes mAP at an IoU threshold of 0.5. Meanwhile, mAP@0.95 refers to the mAP calculated among multiple IoU thresholds from 0.5 to 0.95 in steps of 0.05. "Params" denotes the overall parameter count. Floating-point operations (FLOPs) quantify the algorithm's computational workload by measuring the floating-point operations needed for execution.

Table 2 Comparison experiment of various backbone networks.

Model	mAP@0.5	mAP@0.95	Params	FLOPs

	(%)	(%)	(MB)	(G)
RT-DETR_HGNet-v2	68.3	37.1	125.8	108.3
RT-DETR_ResNet18	67.2	36.4	75.8	57.0
RT-DETR_ Improved ResNet18	67.5	36.7	64.8	49.5

Table 2 demonstrates that using ResNet18 instead of HGNet-v2 as the backbone network in RT-DETR decreases mAP@0.5 by 1.1%, accompanied by a reduction of 39.7% in Params and a decrease of 47.4% in FLOPs. By adopting the improved ResNet18 instead of the original ResNet18, the average precision remains largely unchanged, with a 14.5% decrease in Params and a 13.2% decrease in FLOPs. Compared to the significant reduction in Params and FLOPs, average precision is only a small decrease. Utilizing the improved ResNet18 as the backbone network could decrease the model's parameters and FLOPs, thereby reducing hardware dependence and enhancing the detection speed.

4.2 Experimental comparison of loss functions

To validate the impact of Inner-GIoU, multiple sets of experiments were designed using RT-DETR_Improved ResNet18 as a baseline model. These experiments involved optimizing the ratio parameter for the Inner-GIoU loss function and conducting a comparative analysis of different loss functions. The objective of ratio parameter optimization experiments was to identify the optimal ratio setting that balances precision and recall to enhance overall detection performance. Five distinct ratios were tested: 0.6, 0.8, 1.0, 1.2, and 1.4. Then, five distinct loss functions, namely GIoU, CIoU (Zheng et al., 2020), DIoU (Zheng et al., 2020), MPDIoU (Ma et al., 2023), and Inner-GIoU, were selected for a comprehensive evaluation under controlled experimental conditions.

Table 3 Experiments on the optimization of the ratio parameter for Inner-GIoU.

Ratio	Precision (%)	Recall (%)	mAP@0.5 (%)	mAP@0.95 (%)
Ratio=0.6	72.3	62.9	68.0	35.7
Ratio=0.8	67.5	67.3	67.1	36.1
Ratio=1.0	70.7	65.2	67.5	36.7
Ratio=1.2	70.2	65.5	68.3	36.3
Ratio=1.4	72.1	65.7	68.7	36.5

Table 4 Comparison experiments with different loss functions.

Loss	Precision (%)	Recall (%)	mAP@0.5 (%)	mAP@0.95 (%)
GIoU	70.7	65.2	67.5	36.7
CIoU	69.6	64.4	67.7	36.7
DIoU	71.2	63.1	67.7	36.4
MPDIoU	72.5	62.9	67.4	36.4
Inner-GIoU	72.1	65.7	68.7	36.5

Table 3 illustrates the variation patterns of the Inner-GIoU loss function when using different ratios. When the ratio is set to 1.0, it degenerates into the GIoU loss function, achieving a mAP@0.5 of 67.5%. The optimal ratio is 1.4, obtaining a mAP@0.5 of 68.7%. These results indicate that a larger auxiliary bounding box improves the model's minor defect detection accuracy.

Table 4 demonstrates the comparison experiment results. The Inner-GIoU loss function exhibits the highest recall and mAP@0.5 values. At the same time, there is no significant decrease in the precision and mAP@0.95. Its mAP@0.5 values and recall surpass that of the GIoU loss function by 1.2% and 0.5%, respectively. The Inner-GIoU loss function enhances recall and mAP@0.5 by 1.3% and 1.0%, respectively, in contrast to the CIoU loss function. Furthermore, the Inner-GIoU demonstrates a 1.0% improvement in mAP@0.5 and a 2.6% increase in recall compared to the DIoU loss function. Finally, compared to the MPDIoU, the Inner-GIoU loss function improves recall and mAP@0.5 by 2.8% and 1.3%, respectively.

Therefore, the Inner-GIoU loss function combined with RT-DETR_Improved ResNet18 effectively enhances detection accuracy, validating the feasibility of the proposed method.

4.3 Model training experiment

RT-DETR_ResNet18 was used as the baseline model for the experiments. Various enhancement techniques were integrated into the model, and training was conducted accordingly. The model with the best test result was selected and outputted in each set of experiments. Table 5 illustrates the ablation experiment results, where the experiment numbers correspond to the following descriptions.

1# represents the RT-DETR model using ResNet18 as the backbone network.

2# represents the RT-DETR model using the improved ResNet18 as the backbone network.

3# represents the RT-DETR model combining the improved ResNet18 with the HS-FPN network.

4# represents the RT-DETR model combining the improved ResNet18, the HS-FPN network, and the Inner-GIoU loss function.

Table 5 The ablation experiment results.

Number	Improved ResNet18	HS-FPN	Inner-GIoU	Precision (%)	Recall (%)	mAP@0.5 (%)	mAP@0.95 (%)
1#				69.9	68.2	67.2	36.4
2#	√			70.7	65.2	67.5	36.7
3#	√	√		71.4	68.4	71.0	37.9
4#	√	√	√	73.9	68.5	71.8	39.1

Compared to Experiment 1#, which uses the original ResNet18 feature extraction network, the mAP@0.5 of Experiment 2#, which employs the improved ResNet18 feature extraction network, increases by 0.3%. The mAP@0.5 of Experiment 3#, which incorporates the HS-FPN feature selection and fusion network into Experiment 2#, increases by 3.8% compared to Experiment 1#. Compared to Experiment 1#, the mAP@0.5 in Experiment 4# rises by 4.6% when the Inner-GIoU loss function is added.

Fig. 10 illustrates the precision, recall, and mAP@0.5 curves achieved through training the proposed model and RT-DETR_ResNet18 for 150 epochs with 20,550 iterations. The proposed method's precision, recall, and mAP@0.5 exceed the RT-DETR_ResNet18's. Fig. 11 illustrates the proposed model's detection results, proving that this method can accurately detect minor defects on the blade surface.

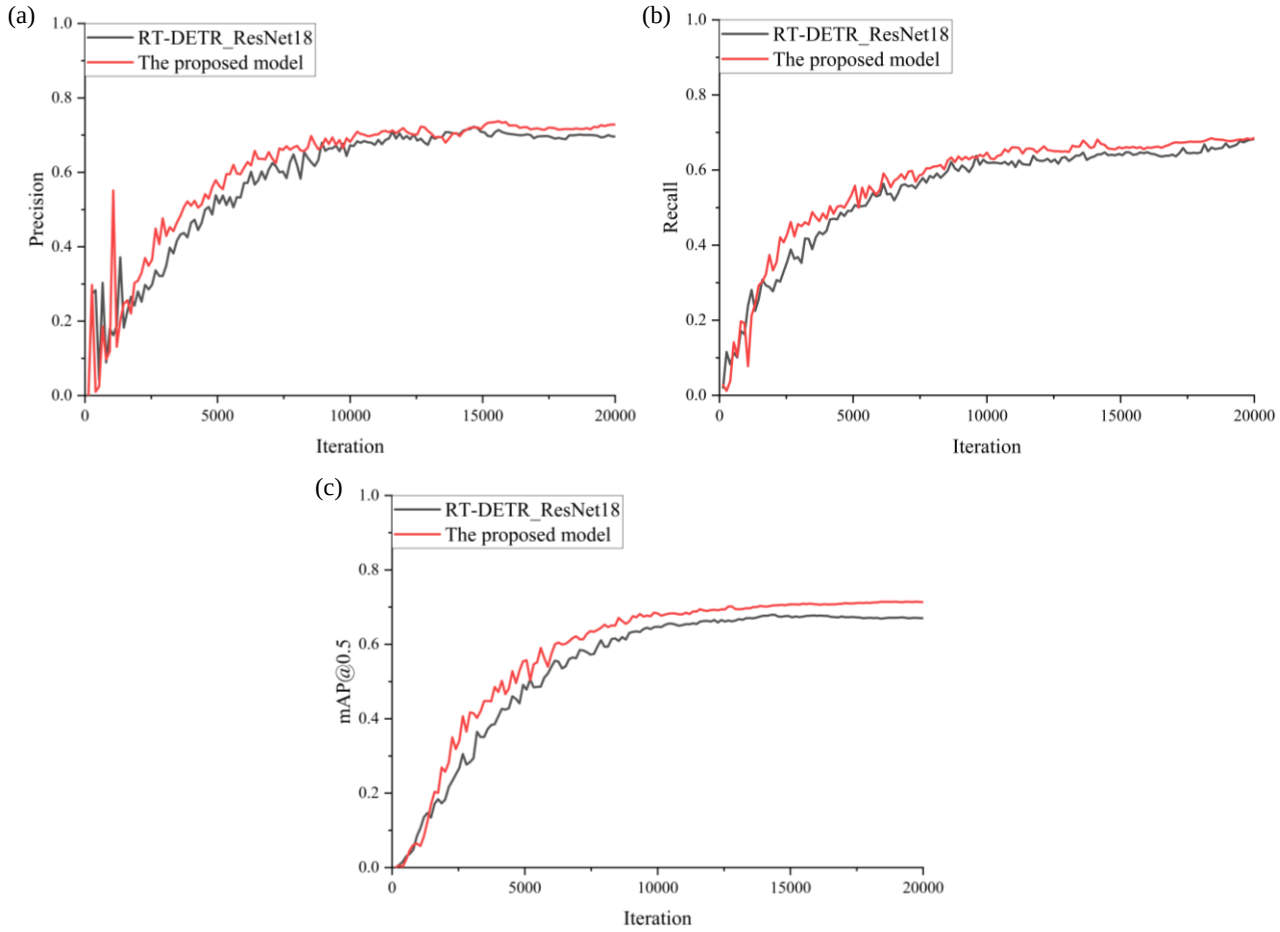

Fig. 10 The curves of (a) Precision, (b) Recall, and (c) mAP@0.5.

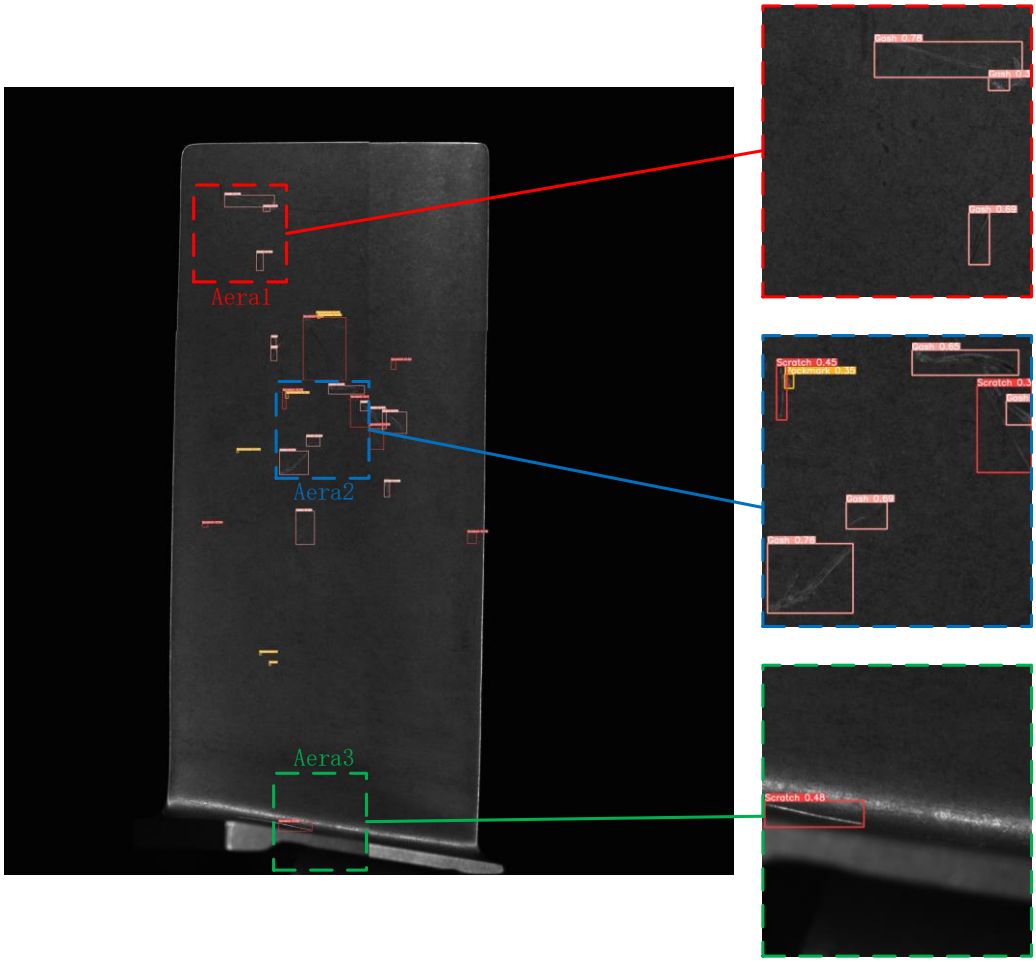

Fig. 11 Detection results of the proposed model.

4.4 Model comparison experiment

To showcase the performers of the proposed model, a comparison with the other object detection models was conducted, comprising Faster R-CNN (Ren et al., 2016), SSD (Liu et al., 2016), DETR (Carion et al., 2020), YOLOv3 (Redmon et al., 2018), YOLOv5 (Li et al., 2022b), YOLOv7 (Wang et al., 2023), and RT-DETR. The input image dimensions were set to 640×640 for 150 epochs. The results from the comparative experiments are summarized in Table 6. The detection mAP@0.5 rate of the model in this paper is 71.8%, and the detection time for a single blade is 2.4 seconds (time to process 80 blade images of 916×960). Compared to Faster R-CNN, SSD, DETR, YOLOv3, YOLOv5, YOLOv7, and the original RT-DETR, the proposed model improves mAP@0.5 by 18.3%, 14.5%, 8.9%, 5.7%, 6.0%, 5.3%, and 3.5%, respectively. Additionally, the detection time for a single blade reduces by 18.25 seconds, 2.07 seconds, 28.1 seconds, 3.56 seconds, 1.96 seconds, 2.45 seconds, and 1.16 seconds, respectively. These results demonstrate that the proposed model offers improved

detection performance and speed.

Table 6 Comparison experiments of the proposed model with different network structures.

Model	mAP@0.5 (%)	Detection time (s)
Faster R-CNN	53.5	20.65
SSD	57.3	4.47
DETR	62.9	30.50
YOLOv3	66.1	5.96
YOLOv5	65.8	4.36
YOLOv7	66.5	4.85
RT-DETR	68.3	3.56
The proposed model	71.8	2.40

5 Conclusions

This study proposes an improved blade detection method based on RT-DETR that can precisely and efficiently detect blade surface defects during the blade production line.

1) To enhance the model's capability of detecting blade surface defects across various scales, the backbone network segment is substituted with an improved and more lightweight ResNet18. Simultaneously, the HS-FPN replaces the original feature fusion approach in RT-DETR.

2) The inner-GIoU loss function replaces RT-DETR’s original loss function to expedite model convergence and enhance the detection accuracy of minor target surface defects.

3) Compared to RT-DETR, the approach proposed in this study enhances mAP@0.5 by 3.5% across four defect types: gash, scratch, bruise, and pockmark. It also reduces the computational time required for detecting a single blade by 1.16 seconds.

The improved RT-DETR enhances minor defect detection accuracy and speed on the blade surface by incorporating refined feature extraction, fusion, and loss function methods. Subsequent research will focus on optimized feature fusion and loss function to improve the algorithm's effectiveness.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal

relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request.

Acknowledgments

The research is financially supported by the National Natural Science Foundation of China (grant number 52305482). Thanks to Huilin Li for revising this paper.

Reference

1. Carion, N., Massa, F., Synnaeve, G., Usunier, N., Kirillov, A., & Zagoruyko, S. (2020). End-to-end object detection with transformers. In *European conference on computer vision* (pp. 213-229). Cham: Springer International Publishing. https://doi.org/10.1007/978-3-030-58452-8_13
2. Chen W, Zhong X, Zhang J, et al. (2022). Object Detection of Aeroengine Blade Boss Based on Yolov4. *Computer Simulation*, *39*(7), 5. https://doi.org/10.3969/j.issn.1006-9348.2022.07.004
3. Chen, J., Kao, S. H., He, H., Zhuo, W., Wen, S., Lee, C. H., & Chan, S. H. G. (2023). Run, don't walk: chasing higher FLOPS for faster neural networks. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition* (pp. 12021-12031). https://doi.org/10.48550/arXiv.2303.03667
4. Chen, Y., Zhang, C., Chen, B., Huang, Y., Sun, Y., Wang, C., Fu, X., Dai, Y., Qin, F., Peng, Y., & Gao, Y. (2024). Accurate leukocyte detection based on deformable-DETR and multi-level feature fusion for aiding diagnosis of blood diseases. *Computers in Biology and Medicine*, *170*, 107917. https://doi.org/10.1016/j.compbiomed.2024.107917
5. Ciampa, F., Mahmoodi, P., Pinto, F., & Meo, M. (2018). Recent advances in active infrared thermography for non-destructive testing of aerospace components. *Sensors*, *18*(2), 609. https://doi.org/10.3390/s18020609
6. Farhadi, A., & Redmon, J. (2018). Yolov3: An incremental improvement. In *Computer vision and pattern recognition* (Vol. 1804, pp. 1-6). Berlin/Heidelberg, Germany: Springer. https://doi.org/10.48550/arXiv.1804.02767
7. Geng, H., Wu, D., & Wang, H. (2022). Experimental and simulation study of material removal behavior in ultra-precision turning of magnesium aluminate spinel (MgAl2O4). *Journal of Manufacturing Processes*, *82*, 36-50. https://doi.org/10.1016/j.jmapro.2022.07.044

8. He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep residual learning for image recognition. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 770-778). https://doi.org/10.48550/arXiv.1512.03385
9. He, W., Li, C., Nie, X., Wei, X., Li, Y., Li, Y., & Luo, S. (2021). Recognition and detection of aero-engine blade damage based on Improved Cascade Mask R-CNN. *Applied optics*, *60*(17), 5124-5133. https://doi.org/10.1364/AO.423333
10. Hui, T., Xu, Y., & Jarhinbek, R. (2021). Detail texture detection based on Yolov4-tiny combined with attention mechanism and bicubic interpolation. *IET Image Processing*, *15*(12), 2736-2748. https://doi.org/10.1049/ipr2.12228
11. Jiang, Y., Liu, K., Huang, J., Zhao, D., Yang, W., & Wang, Y. (2024). Intelligent in-process enhancement technique for machining efficiency in CNC machine tools based on spindle power. *Mechanical Systems and Signal Processing*, *216*, 111495. https://doi.org/10.1016/j.ymssp.2024.111495
12. Lee, D., Yoon, S., Park, J., Eum, S., & Cho, H. (2022). Demonstration of model-assisted probability of detection framework for ultrasonic inspection of cracks in compressor blades. *NDT & E International*, *128*, 102618. https://doi.org/10.1016/j.ndteint.2022.102618
13. Li, Y., Yan, B., Li, W., & Li, D. (2016a). Thickness assessment of thermal barrier coatings of aeroengine blades via dual-frequency eddy current evaluation. *IEEE Magnetics Letters*, *7*, 1-5. https://doi.org/10.1109/LMAG.2016.2590465
14. Li, W., Xin, A. Y., Chen, G., Yin, X., Ge, J., Kong, Q., Zhang, Y., & Wu, Y. (2016b). Detection of cracks in metallic objects by arbitrary scanning direction using a double u-shaped orthogonal ACFM probe. *Materials Transactions*, *57*(5), 608-612. https://doi.org/10.2320/matertrans.M2015384
15. Li, X., Wang, W., Sun, L., Hu, B., Zhu, L., & Zhang, J. (2022a). Deep learning-based defects detection of certain aero-engine blades and vanes with DDSC-YOLOv5s. *Scientific Reports*, *12*(1), 13067. https://doi.org/10.1038/s41598-022-17340-7
16. Li, X., Wang, C., Ju, H., & Li, Z. (2022b). Surface defect detection model for aero-engine components based on improved YOLOv5. *Applied Sciences*, *12*(14), 7235. https://doi.org/10.3390/app12147235
17. Liu, G., Reda, F. A., Shih, K. J., Wang, T. C., Tao, A., & Catanzaro, B. (2018). Image inpainting for irregular holes using partial convolutions. In *Proceedings of the European conference on computer vision (ECCV)* (pp. 85-100). https://doi.org/10.48550/arXiv.1804.07723
18. Liu, W., Anguelov, D., Erhan, D., Szegedy, C., Reed, S., Fu, C. Y., & Berg, A. C. (2016). Ssd: Single shot

multibox detector. In *Computer Vision–ECCV 2016: 14th European Conference, Amsterdam, The Netherlands, October 11–14, 2016, Proceedings, Part I 14* (pp. 21-37). Springer International Publishing. https://doi.org/10.1007/978-3-319-46448-0_2

19. Liu, X., Rao, Z., Zhang, Y., & Zheng, Y. (2023). UAVs Images Based Real-Time Insulator Defect Detection with Transformer Deep Learning. In *2023 IEEE International Conference on Robotics and Biomimetics (ROBIO)* (pp. 1-6). IEEE. https://doi.org/10.1109/ROBIO58561.2023.10354816
20. Ma, S., & Xu, Y. (2023). Mpdiou: a loss for efficient and accurate bounding box regression. *arXiv preprint arXiv:2307.07662*. https://doi.org/10.48550/arXiv.2307.07662
21. Ren, S., He, K., Girshick, R., & Sun, J. (2016). Faster R-CNN: Towards real-time object detection with region proposal networks. *IEEE transactions on pattern analysis and machine intelligence*, *39*(6), 1137-1149. https://doi.org/10.1109/TPAMI.2016.2577031
22. Rezatofighi, H., Tsoi, N., Gwak, J., Sadeghian, A., Reid, I., & Savarese, S. (2019). Generalized intersection over union: A metric and a loss for bounding box regression. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition* (pp. 658-666). https://doi.org/10.1109/CVPR.2019.00075
23. Shen, Z., Wan, X., Ye, F., Guan, X., & Liu, S. (2019). Deep learning based framework for automatic damage detection in aircraft engine borescope inspection. In *2019 International Conference on Computing, Networking and Communications (ICNC)* (pp. 1005-1010). IEEE. https://doi.org/10.1109/ICCNC.2019.8685593
24. Wang D, Xiao H, Wu D. (2024). Automatic radiographic testing for aeroengine turbine blades based on deep learning. *Journal of Propulsion Technology*, 45(5),222-230. https://doi.org/10.13675/j.cnki.tjjs.2210024
25. Wang, C. Y., Bochkovskiy, A., & Liao, H. Y. M. (2023). YOLOv7: Trainable bag-of-freebies sets new state-of-the-art for real-time object detectors. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition* (pp. 7464-7475). https://doi.org/10.48550/arXiv.2207.02696
26. Wang, H., Zhang, K., Wu, D., Yu, T., Yu, J., & Liao, Y. (2021). Analysis and optimization of the machining fixture system stiffness for near-net-shaped aero-engine blade. *The International Journal of Advanced Manufacturing Technology*, *113*, 3509-3523. https://doi.org/10.1007/s00170-021-06730-2
27. Wu, D., Liang, J., Wang, H., & Yu, J. (2022). Investigative on the remanufacturing process of TC4 blade based on selective laser melting and CNC machining. *Journal of Materials Research and Technology*, *21*, 450-464. https://doi.org/10.1016/j.jmrt.2022.09.039
28. Wu, D., Wang, H., & Yu, J. (2021). Research on machining error transmission mechanism and compensation method for near-net-shaped jet engine blades CNC machining process. *The International Journal of Advanced*

Manufacturing Technology, *117*(9), 2755-2773. https://doi.org/10.1007/s00170-021-07818-5

29. Wu, D., Wang, H., Zhang, K., Zhao, B., & Lin, X. (2020a). Research on adaptive CNC machining arithmetic and process for near-net-shaped jet engine blade. *Journal of Intelligent Manufacturing*, *31*, 717-744. https://doi.org/10.1007/s10845-019-01474-z
30. Wu, D., Wang, H., Peng, J., Zhang, K., Yu, J., Zheng, X., & Chen, Y. (2020b). Machining fixture for adaptive CNC machining process of near-net-shaped jet engine blade. *Chinese Journal of Aeronautics*, *33*(4), 1311-1328. https://doi.org/10.1016/j.cja.2019.06.008
31. Wu, D., Zheng, Y., Wang, H., Lv, H., Liang, J., & Yu, J. (2023). Formation mechanism of nano-crystal on the blade surface produced by low-plasticity ultrasonic rolling strengthening process. *Journal of Manufacturing Processes*, *90*, 357-366. https://doi.org/10.1016/j.jmapro.2023.01.023
32. Yixuan, L., Dongbo, W., Jiawei, L., & Hui, W. (2023). Aeroengine blade surface defect detection system based on improved faster RCNN. *International Journal of Intelligent Systems*, *2023*(1), 1992415. https://doi.org/10.1155/2023/1992415
33. Zhang, D., Zeng, N., & Lin, L. (2020). Detection of blades damages in aero engine. In *2020 Chinese Automation Congress (CAC)* (pp. 6129-6134). IEEE. https://doi.org/10.1109/CAC51589.2020.9327202
34. Zhang, H., Chen, J., & Yang, D. (2021). Fibre misalignment and breakage in 3D printing of continuous carbon fibre reinforced thermoplastic composites. *Additive manufacturing*, *38*, 101775. https://doi.org/10.1016/j.addma.2020.101775
35. Zhang, H., Xu, C., & Zhang, S. (2023). Inner-IoU: more effective intersection over union loss with auxiliary bounding box. *arXiv preprint arXiv:2311.02877*. https://doi.org/10.48550/arXiv.2311.02877
36. Zhang, Y., Zuo, Z., Guo, W., Liang, Q., & Chen, H. (2024). Design Strategy of Diagonal Compressors in Compressed Air Energy Storage System. *Journal of Thermal Science*, *33*(3), 872-887. https://doi.org/10.1007/s11630-024-1899-y
37. Zhao, G., Zhao, B., Ding, W., Xin, L., Nian, Z., Peng, J., He, N., & Xu, J. (2024a). Nontraditional energy-assisted mechanical machining of difficult-to-cut materials and components in aerospace community: a comparative analysis. *International Journal of Extreme Manufacturing*, *6*(2), 022007. https://doi.org/10.1088/2631-7990/ad16d6
38. Zhao, B., Wang, Y., Peng, J., Wang, X., Ding, W., Lei, X., Wu, B., Zhang, M., Xu, J., Zhang, L.,& Das, R. (2024b). Overcoming challenges: advancements in cutting techniques for high strength-toughness alloys in aero-engines. *International Journal of Extreme Manufacturing*, *6*(6), 062012. https://doi.org/10.1088/2631-

7990/ad8117

39. Zhao, Y., Lv, W., Xu, S., Wei, J., Wang, G., Dang, Q., Liu, Y., & Chen, J. (2024c). Detrs beat yolos on real-time object detection. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition* (pp. 16965-16974). https://doi.org/10.48550/arXiv.2304.08069

40. Zheng, Z., Wang, P., Liu, W., Li, J., Ye, R., & Ren, D. (2020). Distance-IoU loss: Faster and better learning for bounding box regression. In *Proceedings of the AAAI conference on artificial intelligence* (Vol. 34, No. 07, pp. 12993-13000). https://doi.org/10.1609/aaai.v34i07.6999