

A Fuel Rod Weld Defect Detection Method Based on Unbalanced Convolutional Feature Extraction

Authors: Huang Fan

Date: 2025-04-28T18:35:54+00:00

Abstract

Nuclear fuel rods are subjected to extreme service environments during reactor operation, and the precision inspection of cladding tube-to-end plug welds constitutes a critical safeguard for nuclear safety. To address challenges including the scarcity of high-quality training datasets arising from low real-defect occurrence probability and high costs of X-ray imaging equipment, along with the inability of existing models to satisfy real-time detection requirements, this paper proposes an intelligent defect detection algorithm based on unbalanced convolutional feature extraction. By incorporating a distance attention mechanism into the boundary regression loss function and constructing a backbone network for feature extraction based on unbalanced convolution, we obtain the detection model YOLOv8n-WIOU-Fasternet, which achieves high detection accuracy and favorable real-time performance while resolving issues of limited accuracy and poor stability caused by feature imbalance in defects. Experimental results indicate that the proposed model significantly outperforms manual feature extraction-based defect detection methods and the original YOLOv8 model, with both missed detection and false detection rates below 5%, demonstrating excellent balance between detection accuracy and speed.

Full Text

A Defect Detection Method for Fuel Rod Welds Based on Imbalanced Convolution Feature Extraction

Fan Huang¹, Bo Xiang¹, Ping Li¹, Yue Liu²

¹ CNNC Jianzhong Nuclear Fuel Co., Ltd, Yibin 644000, China

² Harbin Institute of Technology, Harbin 150001, China

Abstract

Nuclear fuel rods face extreme service environments during reactor operation, and precise inspection of the cladding tube and end plug welds is an important safeguard measure for nuclear safety. Due to the low probability of real defects, high cost of X-ray imaging equipment, lack of high-quality training datasets, and difficulty of existing models to meet real-time detection requirements, this paper designs an intelligent defect detection algorithm based on imbalanced convolution feature extraction. By introducing a distance attention mechanism into the boundary regression loss function and constructing a feature extraction backbone network based on imbalanced convolution, we obtain a detection model YOLOv8n-WIOU-Fasternet with high detection accuracy and good real-time performance, solving the problems of limited accuracy and poor stability caused by imbalanced defect features. Experimental results show that this model's detection accuracy is significantly better than artificial feature extraction methods and the original YOLOv8 model, with both missed detection rate and false detection rate below 5%. The model demonstrates excellent capability in balancing detection accuracy and speed.

Keywords: fuel rods; weld seams; defect detection; imbalanced convolution feature extraction; convolutional neural network

1. Introduction

Nuclear fuel rods face extreme service environments during reactor operation, continuously enduring high-temperature, high-pressure conditions, repeated scouring by coolant flowing at speeds exceeding ten meters per second, coupled with chemical corrosion from fission products, complex mechanical loads, and high-temperature steam erosion. As the first barrier of the nuclear reactor safety system, the fuel element cladding forms annular welds with end plugs that are prone to multiple types of defects such as tungsten inclusion, porosity, and incomplete penetration during the welding process, becoming a major risk source for radioactive material leakage [1]. According to HAF60409 nuclear safety regulations and ASME BPVC-III nuclear equipment construction standards, fuel rod manufacturing must implement full-process quality control. 100% non-destructive testing of fuel rod welds is listed as a mandatory quality control node. X-ray digital radiography (DR) is an important detection method for welds [2]; however, it still relies on manual experience for defect interpretation, which has subjective constraints in detection efficiency and consistency of evaluation standards.

Current methods for component defect detection in X-ray DR images are mainly divided into traditional artificial feature extraction methods and deep learning methods [3]. Traditional artificial feature extraction methods extract features through explicitly designed or selected operators, typically offering fast speed, small size, and strong interpretability of detection results. However, these methods have limited detection effectiveness in low-contrast DR image scenarios,

making it difficult to obtain high-level semantic information and highly discriminative features for defects, resulting in a lower upper limit for defect detection accuracy and higher missed detection and false detection rates. Deep learning network models can automatically learn features with less reliance on prior knowledge, but they require large datasets and have high time costs during training. In recent years, with the vigorous development of deep learning, its application in defect detection has been continuously explored, and it has gained increasing favor from researchers due to its high detection accuracy and good stability. Currently, mainstream neural network models for defect detection include convolutional neural networks, autoencoder networks, and recurrent neural networks.

Internationally, in 2018, Ferguson M et al. added a semantic segmentation module to Faster R-CNN based on Mask R-CNN, and defined 15 types of prior boxes with various aspect ratios according to the scale distribution of actual defects in the dataset. They transferred the knowledge learned by the constructed model on ImageNet and COCO datasets to the internal defect localization task in precision component X-ray images, achieving component internal defect localization with a mean average precision (mAP) of 0.957 [4]. In 2021, Mery D's team addressed the time-consuming and labor-intensive annotation problem in precision component internal defect localization tasks based on X-ray images. They used a small number of non-defective aluminum alloy wheel casting X-ray images superimposed with simulated defects while automatically obtaining the ground truth bounding boxes to construct a dataset, and trained eight deep models (YOLO series, RetinaNet, and EfficientDet) on this dataset [5]. Among them, YOLOv5s achieved good results on their constructed test set: average precision of 0.90, F1 score of 0.91, and FPS of 90. In 2023, Lal R et al. proposed a lightweight model that minimized computational requirements through depthwise separable convolution and global average convolution, demonstrating advantages in both accuracy and inference time compared to advanced pre-trained CNN models such as MobileNet, Inception, and ResNet. The customized model with 590K parameters using depthwise separable convolution outperformed pre-trained architectures like ResNet and Vision Transformers in accuracy (81.87%) and easily surpassed ResNet and Inception in inference time (12 ms) [6]. This work demonstrated that it is possible to build customized networks with efficient architectures and faster inference times without relying on pre-trained architectures.

Domestically, in 2019, Zhou Zhou et al. proposed a two-stage defect localization model for precision components. This model used VGG16 as the backbone network, applied atrous convolution to the features output by the backbone network to increase the receptive field, fused multi-scale features to obtain complete information, and used position-sensitive convolutional and pooling layers to achieve object classification and bounding box regression within proposals. Experimental results showed that this model outperformed Faster R-CNN and SSD on their task, with an average intersection over union of 80.59% between localization results and manual annotations [7]. In 2020, Guo Ruiqi et al. im-

proved the Unet model by introducing dropout layers to reduce overfitting risk and using multi-stride convolution instead of max pooling for downsampling to reduce information loss, thereby enhancing the network's defect segmentation and localization performance [8]. In the same year, Yu H et al. proposed an adaptive depth and receptive field selection network for semantic segmentation of defects in precision component X-rays [9]. First, they designed ResNet18 with ADSM (Adaptive Depth Selection Mechanism) to extract and adaptively aggregate different depth features, which helps distinguish similar defects. Then, they proposed ARFB (Adaptive Receptive Field Block) to select the optimal receptive field in a data-driven manner to adapt to defect scale variations. Finally, they established a casting defect segmentation dataset SRIF-CDS to train and evaluate the method. Experiments on SRIF-CDS showed that the method achieved 0.86 mIoU (mean intersection over union) and 0.92 mAcc (mean accuracy). In 2021, Du W et al. used DetNet as the backbone network to compensate for position information lost during downsampling, constructed a feature pyramid with enhanced bottom-up paths based on PANet (Path Aggregation Network) to narrow feature differences between feature maps and corresponding proposal boxes, and adopted soft-NMS (soft Non-Maximum Suppression) to retain more proposals with higher confidence. Through these improvements, the mAP value increased by 0.118 compared to the baseline network architecture [10].

However, current deep learning methods suffer from a lack of high-quality training datasets due to the low probability of real defects and high cost of X-ray imaging equipment, and the real-time performance of models often fails to meet actual industrial production demands. Since YOLO series single-stage detection models offer good real-time performance while ensuring high detection accuracy, they are widely applied in various actual industrial production scenarios. This paper redesigns the loss function and backbone feature extraction network based on the YOLOv8 model to make it more suitable for low-contrast defect detection scenarios, introducing a novel partial convolution concept to design a new intelligent 2D defect detection algorithm.

2. Methodology

2.1 Regression Loss Function Design for the Intelligent Detection Algorithm

(1) Boundary Regression Loss In the original YOLOv8 model, CIOU loss is used as the boundary regression loss function, which primarily combines factors of center point distance, overlap degree, and aspect ratio between the predicted box and target box based on their ratio. The specific formula is as follows:

$$\text{CIOU} = \text{IoU} - \frac{\rho^2(b, b^{gt})}{c^2} - \alpha v$$

where

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

and

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

where IoU is the intersection over union between the defect prediction box and ground truth box, G represents the defect ground truth box, D represents the defect prediction box; d represents the distance between the center points of the predicted box and target box; b represents the prediction box; b^{gt} represents the target box; c represents the diagonal distance of the minimum enclosing rectangle of the predicted box and target box; α is a weight parameter, and v is a correction factor that considers the shape and orientation of the target box to further optimize the loss function, where w^{gt} , h^{gt} and w , h represent the width and height of the target box and prediction box, respectively.

Although this CIOU loss function comprehensively considers multiple factors including position, shape, and orientation of the predicted box and target box, and its sensitivity to localization accuracy is improved compared to traditional single IoU value loss functions, some key parameters in CIOU, such as the weight parameter α of the correction factor, have a significant impact on loss function convergence and require careful tuning to obtain better model performance. Furthermore, the aspect ratio correction factor cannot fully reflect the shape and orientation of the predicted box and target box. When the predicted box and target box have the same aspect ratio but different actual width and height values, the penalty term of this correction factor is always zero, hindering further model optimization. Since training data inevitably contains both low-quality and high-quality anchor boxes, according to the CIOU calculation method, the correction factor increases penalties for low-quality anchor boxes while enhancing the competitiveness of high-quality anchor boxes, weakening the influence of most normal-quality anchor boxes on the model and thereby reducing model generalization performance.

To address the above issues, this paper introduces a boundary regression loss function based on a dual-layer distance attention mechanism, WIOU, which makes model convergence focus more on normal-quality anchor boxes through a layer of intersection over union calculation and a layer of distance outlieriness calculation:

$$\text{WIOU} = \text{IoU} \times \exp\left(\frac{\text{IoU} - \text{IoU}^*}{\text{IoU}^*}\right)$$

where $\frac{\text{IoU}-\text{IoU}^*}{\text{IoU}^*}$ reflects the outlieriness of two anchor boxes, significantly amplifying the loss function for normal-quality anchor boxes; IoU reflects the overlap degree between two anchor boxes, significantly reducing the competitiveness of high-quality anchor boxes. WIOU focuses on the distance between the center points of two anchor boxes and does not directly introduce geometric metrics such as aspect ratio into the loss function calculation, reducing harmful gradients generated by low-quality anchor boxes.

(2) Distribution Focal Loss Distribution focal loss, also known as DFL loss, further enhances the model's focus capability on labels based on boundary regression loss. The DFL loss function primarily uses a cross-entropy function to optimize the probability distribution of the two positions closest to the target label, enabling the model to focus more quickly and accurately on the vicinity of the target label. Through this approach, the distribution learned by the theory is located near the target floating-point coordinates, and weights for the two nearest integer coordinates are obtained through linear interpolation. The DFL loss calculation formula is as follows:

$$\text{DFL} = - \sum_{i=0}^{n-1} y_i \log(S_i) + (1 - y_i) \log(1 - S_i)$$

where y represents the predicted value; y_{i+1} represents the adjacent predicted value; S_i represents the estimated probability that the label is class $y_i=i$. In simple terms, it transforms a regression problem into a classification problem and optimizes it using a cross-entropy function.

In summary, this paper's model uses a combination of WIOU and DFL as the regression loss function, where DFL is responsible for optimizing the position of each side of the predicted box, and WIOU optimizes the overall position, shape, and orientation of the optimized predicted box.

2.2 Backbone Feature Extraction Network Design for the Intelligent Detection Algorithm

Considering the real-time detection requirements in fuel rod production, this paper adopts the network structure of the single-stage detection model YOLOv8 as the foundation for defect detection. Its backbone uses the CSPDarknet structure, which effectively reduces computational load through differential feature maps to ensure model lightweighting. Additionally, compared with other YOLO series models, YOLOv8 incorporates ELAN ideas to optimize the C3 module into a C2f module, achieving further lightweighting while ensuring the model obtains richer gradient flow information, which is more conducive to maintaining detection accuracy. However, despite these improvements, due to the limited performance of the CSPDarknet structure itself in small target detection, the detection accuracy of the YOLOv8 model still falls short of expected targets.

To address these issues, we further optimize and redesign the backbone feature extraction network of the detection model.

2.2.1 Research on Novel Partial Convolution Algorithms When improving detection accuracy of YOLO series models, various attention mechanism modules are introduced to enhance the model's focus capability on target foreground features. However, the introduction of attention mechanism modules not only requires specialized hardware support but also leads to more complex network structures and increased network latency, resulting in substantial increases in computational costs. To solve this problem while improving accuracy and minimizing computational costs and model runtime, most current methods attempt to reduce the number of floating-point operations. However, network latency depends not only on the number of floating-point operations but also on memory access, as shown in the following formula. Simply reducing the total number of floating-point operations may have the side effect of increasing memory access, ultimately failing to achieve true low network latency and computational cost savings:

$$\text{Latency} = \frac{\text{FLOPs}}{\text{FLOPS}}$$

where FLOPs is the total number of floating-point operations; FLOPS is the number of floating-point operations per second, which depends on the memory access of each module in the model network and represents the effective computational speed of the model.

Depthwise separable convolution (DWConv) is a typical method combined with attention mechanism modules to reduce the total number of floating-point operations. It performs convolution operations by sliding c filters of size $k \times k$ across each input channel and calculating only one output channel value, as shown in Figure 1 [Figure 1: see original paper]. Analysis shows that the total floating-point computation of regular convolution is $h \times w \times k^2 \times c$, while depthwise separable convolution's total floating-point computation is $h \times w \times k^2 \times c$, effectively reducing FLOPs. However, since depthwise separable convolution calculates only one output channel value, it cannot be used alone to replace regular convolution, as this would cause severe accuracy degradation. It is typically used in conjunction with pointwise convolution (PWConv), which increases the channel width to c' (where $c' > c$). The memory access of a depthwise separable convolution module becomes $h \times w \times c' \times c$, while the memory access of a regular convolution is $h \times w \times c$. Since $h \times w \times c' \times c > h \times w \times c$, the memory access of depthwise separable convolution modules is greater than that of regular convolution. The reduction in total floating-point operations brought by depthwise separable convolution cannot be effectively converted into reduced network latency due to increased memory access, failing to effectively shorten model runtime and save computational costs.

To address the problems with depthwise separable convolution, this paper introduces a novel partial convolution PConv that can simultaneously reduce both total floating-point operations and memory access. The novel partial convolution PConv performs regular convolution operations on partial input channels while keeping the remaining channels unchanged, as shown in Figure 2 [Figure 2: see original paper]. Let the number of channels performing regular convolution operations be p_c , the convolution kernel size be $k \times k$, and the width and height of the input feature map be h and w , respectively. The total floating-point computation of a novel partial convolution module is:

$$\text{FLOPs}_{\text{PConv}} = h \times w \times k^2 \times p_c$$

In practice, $p_c = c/4$ is used, making the total floating-point computation of a novel partial convolution module only $1/4$ of that of a regular convolution module. Additionally, its memory access is:

$$\text{Memory Access}_{\text{PConv}} = h \times w \times p_c + h \times w \times c$$

which is only $h \times w \times c + h \times w \times p_c$. Since $c - p_c$ channels do not participate in convolution operations, they consume no memory access.

Since some channels in the above partial convolution are not convolved and their information is not utilized, a pointwise convolution module (PWConv) is connected after the novel partial convolution, as shown in Figure 3 [Figure 3: see original paper]. After the PConv with p_c channels, a 1×1 PWConv is connected. The combined total floating-point computation of the PConv+PWConv module is:

$$\text{FLOPs}_{\text{PConv+PWConv}} = h \times w \times k^2 \times p_c + h \times w \times c \times c'$$

Since the effect of the PConv+PWConv module can be equivalent to a single-step T-shaped convolution, as shown in Figure 3 [Figure 3: see original paper], the total floating-point computation of a single-step T-shaped convolution is $h \times w \times k^2 \times c + h \times w \times c \times c' > h \times w \times k^2 \times p_c + h \times w \times c \times c'$. Therefore, the decoupled two-step implementation of PConv+PWConv fully utilizes the redundancy between p_c filters and reduces total floating-point operations FLOPs more than its equivalent single-step T-shaped convolution.

2.2.2 Backbone Feature Extraction Network Design Based on Novel Partial Convolution

Based on the above novel partial convolution concept, we design a four-layer backbone feature extraction network structure that matches YOLOv8-head, as shown in Figure 4 [Figure 4: see original paper]. Each layer contains several fasternet blocks. In the first layer, an embedding module with a stride of 4 and kernel size of 4×4 is applied before the fasternet block. In the second, third, and fourth layers, a downsampling module with a stride of 2 and kernel size of 3×3 is applied before the fasternet block.

is used before each fasternet block for spatial downsampling and channel expansion. Since the last two layers consume less memory access and have greater floating-point computation redundancy, the number of fasternet modules in the last two layers is greater than in the first two layers to undertake more computational tasks.

Each fasternet module consists of a novel partial convolution PConv connected to two pointwise convolutions PWConv, forming an inverted residual structure, as shown in Figure 5 [Figure 5: see original paper]. In each fasternet module, partial convolution is performed on designated channels, pointwise convolution is performed on remaining channels, and BN layers and activation functions are placed only after the middle pointwise convolution to ensure feature diversity and reduce computational complexity. The advantage of BN layers is that they can be merged into adjacent convolutional layers through structural reparameterization to improve computational speed. For the activation function, ReLU is used as the activation function.

The above backbone feature extraction network based on novel partial convolution only completes the extraction of input image feature information. A neck layer is still needed for feature dimensionality reduction and adjustment, and a head layer is needed as a bounding box regressor and classifier to process features and produce final outputs.

The PANet structure effectively utilizes information from different levels of feature maps through top-down and bottom-up bidirectional path aggregation mechanisms, making the model more adaptable to different scale targets and scenarios. PANet contains two path aggregation modules: the top-down PA module and the bottom-up PAM module. PA starts from lower-resolution feature maps and passes information upward layer by layer through upsampling, while PAM starts from higher-resolution feature maps and passes information downward layer by layer through downsampling. These two path aggregation modules cooperate to enable PANet to better integrate multi-scale feature information. Therefore, this paper adopts the PANet structure as the neck layer. The backbone network's output ultimately undergoes $32\times$ downsampling after an SPPF layer. This $32\times$ downsampling result, along with previous $16\times$ and $8\times$ downsampling results, is input into the PANet structure. After upsampling and channel fusion operations, three outputs are obtained and fed into the subsequent head layer to calculate the loss function and output results.

The Decoupled-head structure adds an additional branch on top of the backbone network features to separate feature extraction and pixel-level prediction, making the network more flexible in processing semantic information at different scales and further improving target detection accuracy. Meanwhile, the Decoupled-head structure can fuse multi-scale feature information by adding branches at different levels of the backbone network, improving the entire network's detection capability for multi-scale targets. Treating pixel-level prediction as an independent task can better preserve detail and edge information, improving detection accuracy. Additionally, compared with other head struc-

tures, the Decoupled-head structure can be more conveniently extended and modified according to needs to adapt to different detection tasks and dataset requirements. In this paper, three output feature information at different scales from the PANet layer are input into the Decoupled-head structure. After decoupled separation, feature fusion, and independent prediction branch processing, the loss function and prediction results are obtained.

Using the four-layer backbone feature extraction network structure based on novel partial convolution as the model's backbone feature extraction network, inputting its three different scale output features into the PANet-neck structure for dimensionality reduction and fusion adjustment, and finally inputting the three different scale features processed by the neck into the Decoupled-head for three types of loss function calculations and result prediction achieves accurate defect detection. The four-layer fasternet backbone network, PANet, and Decoupled-head together form a novel network structure YOLOv8-WIOU-Fasternet, which demonstrates outstanding performance in low-contrast small defect target detection.

2.3 Overall Algorithm Flow

This paper redesigns the loss function and backbone feature extraction network based on the YOLOv8 model to make it more suitable for low-contrast small target detection scenarios while achieving effective control of time costs. The structure of the imbalanced depthwise separable convolutional neural network (YOLOv8-WIOU-fasternet model) is shown in Figure 6 [Figure 6: see original paper], mainly including the loss function, backbone feature extraction network, neck layer for feature dimensionality reduction and adjustment, and head layer as bounding box regressor and classifier.

3. Experiments

3.1 Data Acquisition

All experimental data in this study are self-collected digital DR images of 500 fuel rods, not a public dataset. X-ray irradiation parameters were adjusted with a focal length of 1000 mm, tube voltage of 220 kV, and image display window width and level adjusted to approximately 30000 and 35000, respectively. Each DR projection simultaneously collected data from multiple fuel rod end plugs, as shown in Figure 7 [Figure 7: see original paper].

During sample preparation, a batch of abnormal defect samples including porosity, gas swelling, incomplete penetration, tungsten inclusion, and blocked holes were prepared by adjusting welding process parameters (Figure 7). After acquiring the original X-ray images of fuel rod end plugs shown in Figure 7, the fuel rod end plugs were automatically segmented into individual end plugs, and the area near the end plug weld was selected as the region of interest. Defects were manually annotated within the region of interest.

3.2 Precision Performance Analysis of Defect Detection Models

Based on the training parameters, the boundary regression loss and distribution focal loss function curves are shown in Figure 9 [Figure 9: see original paper]. The figure shows that the loss function of the YOLOv8-WIOU-Fasternet detection model designed in this paper converges more smoothly and faster, with particularly outstanding performance in the boundary regression loss function. The boundary regression loss function ultimately converges to approximately 0.5, while the distribution focal loss function converges to approximately 0.8.

(1) Missed Detection Rate The missed detection rate is related to the IoU threshold used to determine whether a sample is correctly detected. With unchanged prediction results, a larger IoU threshold leads to a smaller proportion of correctly predicted samples. Considering factors such as defect size and contrast, this IoU threshold was set to 0.25 in this experiment. Additionally, in deep learning-based detection models, the missed detection rate is also related to the confidence threshold, which filters candidate boxes with confidence scores above the threshold as prediction results. According to the definition of missed detection rate:

$$R_{\text{miss}} = \frac{FN}{TP + FN}$$

where p_R is called recall rate, characterizing the proportion of real defects detected among all real defects. Based on the recall rate output by the model, the curve of missed detection rate varying with model confidence threshold can be calculated, as shown in Figure 10 [Figure 10: see original paper]. The missed detection rate of the YOLOv8-WIOU-Fasternet model is significantly lower than other models.

(2) False Detection Rate The false detection rate is related to the IoU threshold used to determine whether a sample is correctly detected, with the IoU threshold set to the same 0.25 as the missed detection rate. Additionally, in deep learning-based detection models, the false detection rate is also related to the confidence threshold. According to the definition of false detection rate:

$$R_{\text{false}} = \frac{FP}{TP + FP}$$

where p_P is called precision rate, characterizing the proportion of real defects among all detected defects. Based on the precision rate output by the model, the curve of false detection rate varying with model confidence threshold can be calculated, as shown in Figure 11 [Figure 11: see original paper]. The false detection rate of the YOLOv8-WIOU-Fasternet model is significantly lower than other models.

Based on the relationship between missed detection rate, false detection rate, and confidence threshold, the relationship curve between false detection rate and missed detection rate can be obtained, as shown in Figure 12 [Figure 12: see original paper]. As the confidence threshold increases, the missed detection rate increases while the false detection rate decreases. Therefore, to measure the comprehensive precision performance of the model under different confidence thresholds, the F1 score and AP value are commonly used in this field to reflect the model's optimal comprehensive precision performance and average comprehensive precision performance.

(3) Optimal Comprehensive Precision Performance Metric: F1 Score

The F1 score is a metric for measuring the optimal precision performance of a model under different confidence thresholds, defined as the harmonic mean of precision p_P and recall p_R :

$$F1 = 2 \times \frac{p_P \times p_R}{p_P + p_R}$$

The value ranges in $[0,1]$, with larger values indicating better comprehensive precision performance. The maximum F1 score reflects the model's optimal precision performance.

Table 1 shows the missed detection and false detection rates at maximum F1 score for several models. The table demonstrates that the comprehensive precision performance of the YOLOv8n-WIOU-Fasternet model proposed in this paper is significantly better than other models, with false detection and missed detection rates of 3.1% and 3.5%, respectively, at optimal comprehensive precision performance.

Table 1: Maximum F1 Score of Several Models

Model	Confidence Threshold	False Detection Rate	Missed Detection Rate
YOLOv8n	-	10.2%	10.8%
YOLOv8n-WIOU	-	-	-
YOLOv8n-WIOU-Fasternet	-	3.1%	3.5%

(4) Average Comprehensive Precision Performance Metric: AP

To further measure the comprehensive precision performance of the model under different confidence thresholds, the average comprehensive precision performance metric AP is introduced, defined as the area enclosed by the false detection rate-missed detection rate curve and the lines $x=1$, $y=1$. This value depends only on

the IoU threshold. The curve of AP value varying with IoU threshold is shown in Figure 13 [Figure 13: see original paper]. Higher AP values indicate better average comprehensive precision performance across different confidence thresholds. The figure shows that the average comprehensive precision performance of the YOLOv8n-WIOU-Fasternet model designed in this paper is significantly better than other models.

3.3 Speed Performance Analysis of Defect Detection Models

In addition to the precision metrics above, this paper tested the detection speed of the models. The experimental platform uniformly used NVIDIA GeForce RTX 3060, with results shown in Table 2. The implementation results show that the designed YOLOv8n-WIOU-Fasternet model achieves significant precision improvement with relatively small computational cost increase compared to the original YOLOv8 model. Meanwhile, compared with traditional models with self-attention mechanism modules, the computational cost is substantially reduced, better achieving a balance between detection precision performance and speed performance.

Table 2: Detection Speed Comparison of Different Models

Model	FLOPs	Parameters	Single Image Detection Time/ms
YOLOv8n	-	-	-
YOLOv8n-WIOU	-	-	-
YOLOv8n-WIOU-Fasternet	-	-	-
YOLOv8-Swintransformer	11.4G	79.8G	>20 min per training epoch

4. Conclusion

This paper designs an intelligent 2D defect detection algorithm based on novel partial convolution, obtaining a detection model YOLOv8n-WIOU-Fasternet with high detection accuracy and good real-time performance by optimizing the boundary regression loss function and redesigning the backbone feature extraction network based on novel partial convolution. Experiments demonstrate that this model's detection accuracy is significantly better than artificial feature extraction methods and the original YOLOv8 model, with both missed detection rate and false detection rate below 5%. The model exhibits excellent capability in balancing detection accuracy and speed.

References

1. 杨钦轲, 王华才, 程焕林, 汤琪, 宋武林, 王玮, 郭一帆: 压水堆核电站完整和破损燃料棒硬度和杨氏模量研究. 原子能科学技术 2024, 58(1):157-165.
2. Naddaf-Sh M-M, Naddaf-Sh S, Zargarzadeh H, Zahiri SM, Dalton M, Elpers G, Kashani AR: Defect detection and classification in welding using

- deep learning and digital radiography. In: Fault diagnosis and prognosis techniques for complex engineering systems. Elsevier; 2021.
3. 武博, 材料加工工程: 航空钛合金铸件内部缺陷自动识别关键技术研究. 华中科技大学.
 4. Ferguson M, Ak R, Lee Y-TT, Law KH: Detection and segmentation of manufacturing defects with convolutional neural networks and transfer learning. Smart and sustainable manufacturing systems 2018, 2(1):137-164.
 5. Mery D: Aluminum casting inspection using deep object detection methods and simulated ellipsoidal defects. Machine Vision and Applications 2021, 32(3):72.
 6. Lal R, Bolla BK, Sabeesh E: Efficient neural net approaches in metal casting defect detection. Procedia Computer Science 2023, 218:1958-1967.
 7. 周洲等人针对精密构件中的缺陷定位问题, 提出了一种双阶段缺陷定位模型. (2019)
 8. 郭瑞琦, 王明泉, 张俊生, 张曼, 张馨心: 基于 U-Net 卷积神经网络的轮毂缺陷分割. 自动化与仪表 2020 , 35(4):5.
 9. Yu H, Li X, Song K, Shang E, Liu H, Yan Y: Adaptive depth and receptive field selection network for defect semantic segmentation on castings X-rays. NDT & E International 2020, 116:102345.
 10. Du W, Shen H, Fu J, Zhang G, Shi X, He Q: Automated detection of defects with low semantic information in X-ray images based on deep learning. Journal of Intelligent Manufacturing 2021, 32:141-156.

Note: Figure translations are in progress. See original paper for figures.

Source: ChinaXiv — Machine translation. Verify with original.