

# Wood Panel Defect Detection Based on Improved YOLOv8n

Rui Li, Shilu Zhong, and Xuemei Yang

Wood panel surface defect detection is critical to product quality. Traditional detection methods are time-consuming and subjective, and they can lead to economic waste, while deep learning image recognition techniques offer a new approach. However, the accuracy and convergence speed of existing defect detection techniques still require improvement. In this paper, an improved algorithm based on YOLOv8n was designed for accurate detection of wood panel defects. The C-ADown method was designed to replace traditional downsampling, while preserving high-frequency features. The combination of the Dilation-wise Residual Module and multi-scale dilation attention was employed to enhance the multiscale robustness of defect detection. A hybrid encoder was added to improve localization accuracy. The loss function was optimized to improve detection accuracy and convergence speed. Compared to the base YOLOv8 version, the improved model achieved a 6.1% increase in mAP, an 8% increase in recall, and a 3.6% increase in precision, significantly enhancing the model's detection capabilities. The GitHub link to the improved algorithm files is as follows: (<https://github.com/humblefactos1/YOLOV8-CDC/tree/main>.)

DOI: 10.15376/biores.20.2.2556-2573

**Keywords:** *Wood panel; Deep learning; YOLOv8n; C-ADown; Dilation-wise Residual; Multi-scale dilation attention; Loss function*

**Contact information:** *College of Furnishings and Industrial Design, Nanjing Forestry University, Nanjing 210037, China; \*Corresponding author: mgslirui0909@gmail.com*

## INTRODUCTION

The detection of surface defects in wood panels has always been an urgent problem for the wood processing industry. The traditional method of relying on manual visual inspection has many drawbacks, including low efficiency, low accuracy, and high subjectivity. In addition, due to the low degree of automation of the production line, it is impossible to realize real-time monitoring and feedback, resulting in a high rate of defective products, which brings huge economic losses to enterprises. Studies have shown that traditional manual defect detection results in approximately 25% of wood resources being wasted. A 1% reduction in raw material waste can decrease overall production costs by about 2% (Buehlmann and Thomas 2002). In addition, the repetitive labor of manual inspection easily leads to inspector fatigue, which affects the quality of inspection and reduces the mechanical properties, appearance and utilization of wood, resulting in a serious waste of wood resources (Cheng 2020). The quality of inspections is affected by inspector fatigue.

Rapid advances in computer and sensor technology have revolutionized the wood industry, and non-destructive testing techniques have emerged. Among them, acoustic, radiographic, and optical inspection methods were once the traditional mainstream means

(Wang *et al.* 2013). However, as deep learning technology has become increasingly sophisticated, it has become increasingly popular. With the maturity of deep learning technology, NDT methods based on image recognition have gradually become a research hotspot in academia and the industry due to its high efficiency and accuracy (Wang *et al.* 2024a). This method realizes automatic identification and classification of defects through deep learning of wood panel images, providing a new paradigm for wood quality inspection. Related studies have shown that the deep learning method shows great potential in wood panel defect detection, which is expected to significantly improve the productivity and product quality in the wood processing industry (Liu *et al.* 2023). Urbonas *et al.* (2019) used a Faster R-CNN-based target detection network to localize and classify surface defects in wood veneer, achieving an average accuracy of 80.6% using ResNet152 as a pre-trained model. Cheng (2023) proposed a copy-paste-based class coverage method to address imbalanced datasets. To tackle real-time performance and detection accuracy, a CBi2-YOLO model was developed for wood panel defect detection. To fulfill the requirement of calculating defect areas, Jia *et al.* (2023) proposed a quantitative recognition method based on YOLOv5, incorporating a dual-channel attention module to improve the model's ability to recognize specific wood defects. Additionally, a shallow weighted feature fusion network was introduced to fuse feature information from various layers extracted by the backbone network, reducing the loss of feature information for small wood defects. Jiang and Zhao (2024) proposed YOLOv7-ESS based on YOLOv7, which embeds a dual-channel attention module to improve the model's ability to recognize special defects in wood panels. A shallow weighted feature fusion network is introduced to fuse the feature information of each layer extracted by the backbone network to reduce the loss of feature information of small defects in wood panel. Yang *et al.* (2023) employed global and local adaptive thresholding algorithms to segment surface defects and extract image patches. By replacing the ReLU activation function with ReLU6 and introducing an inverted residual structure, the MobileNetv2 deep learning network was optimized for defect detection and classification. Wang *et al.* (2024b) constructed a Wood-Net network, which realizes the defect recognition in the process of wood preference with high accuracy. Wang *et al.* (2024c) introduced a two-way feature fusion network based on the YOLO-v8 algorithm and proposes a feature fusion network model that combines the attention mechanism and loss function optimization.

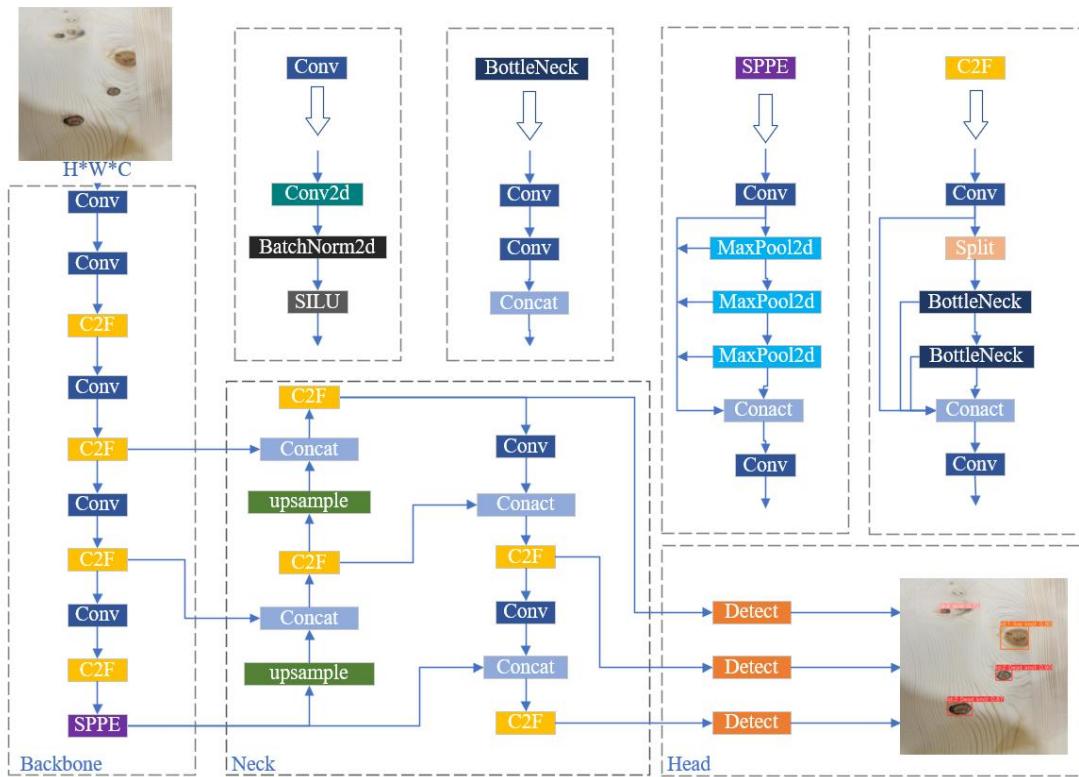
Aiming at the lack of detection and leakage caused by the complexity of defects and low recognition degree in the nondestructive testing of wood panel, this paper designs an improved detection model based on YOLOv8. The main improvement points are as follows:

1. The design of C-ADown instead of the traditional convolutional downsampling retains the main features of the wood panel defects while effectively reducing the size of the feature map and enhancing the model's ability to perceive the local features.
2. A dynamic weight residual (DWR) module combining the attention mechanism of multi-scale feature alignment (MSDA) is designed to replace the C2F and bottleneck modules in the original YOLOv8. This module can adaptively adjust the weights of different scale features to improve the detection accuracy of the model for multi-scale targets.

3. The Neck structure is improved by utilizing a hybrid encoder to convert multi-scale features into a series of image features through intra-scale feature interaction and cross-scale feature fusion.
4. The loss function is improved to speed up the convergence of the model and improve the detection accuracy.

## YOLOv8 Detection Algorithm

YOLOv8, released by Ultralytics in 2023, is the latest iteration of the YOLO series, building upon the significant speed and accuracy improvements achieved by YOLOv5. It consistently demonstrates state-of-the-art performance on various publicly available datasets and is considered an enhanced version of existing YOLO variants such as YOLOv5 and YOLOX (Varghese and Sambath 2024). However, the original YOLOv8 architecture exhibits limitations when tasked with detecting objects such as wood, which contain numerous small defects. These defects typically occupy a small portion of the image pixels and possess low feature resolution, hindering the capture of fine defect details in the deeper network layers. To address these challenges, this paper conducts a thorough investigation of the YOLOv8n model and proposes several enhancements to bolster its performance in detecting wood panel surface, thereby better aligning with the practical demands of wood defect detection.



**Fig. 1.** YOLOv8 structure

## EXPERIMENTAL

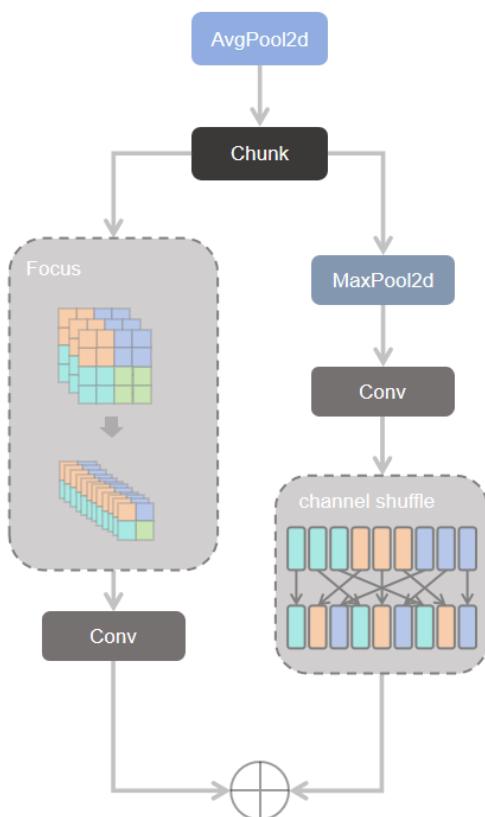
### Improved YOLOv8n Algorithm

#### *C-ADown downsampling module design*

Downsampling is a technique employed to expand the receptive field by reducing the feature map size, which allows the model to capture a broader range of contextual information within an image. Traditional downsampling methods for convolution operations often increase the number and size of convolution kernels, which results in a significant rise in both model parameters and computational complexity (Varghese and Sambath 2024). These conventional approaches, while effective in capturing hierarchical information, tend to be resource-intensive and can lead to overfitting, particularly when working with high-resolution images.

ADown, the downsampling method employed in YOLOv9, effectively preserves global image information through average pooling, aiding in the understanding of overall image structure and texture. Additionally, maximum pooling is used to capture local features such as edges and corner points, contributing to target localization.

In this paper, FOCUS slicing is introduced as a replacement for the parallel 3x3 convolution module in ADown downsampling. This module downsamples the feature map by slicing the image at the pixel level and converting spatial information into channel information, ensuring that original pixel information is not lost. By expanding the number of channels by a factor of four, the network can analyze the image from multiple perspectives, extracting richer features. An increased number of channels enhances the network's feature representation capabilities, allowing for better differentiation between various target types (Wang *et al.* 2024d).



**Fig. 2.** C-ADown downsampling structure

To prevent the network from over-relying on specific channels, channel shuffling is performed after maximum pooling of parallel data. This technique disrupts the channel order in the feature map, forcing the network to learn more complex feature representations and improving model generalization. This enhancement enables the model to adapt to different feature types, facilitating the extraction of complex features like wood grain and color and improving the perception of subtle defects.

The C-ADown module, therefore, not only minimizes information loss but also facilitates a smoother transition of the feature map across different scales. Its efficiency is evident in the reduction of computational complexity compared to traditional convolution-based downsampling methods, as it eliminates the need for large convolutional kernels while still capturing fine details. The enhanced channel manipulation capabilities allow the network to learn more intricate features, particularly in areas with complex textures or small defects. These improvements contribute significantly to the model's ability to accurately detect wood panel defects, making it particularly effective for fine-grained defect detection and regions with detailed structural patterns. The structure of the C-ADown module is illustrated in Fig. 2.

#### *C2F\_DWR (MSDA) design*

The C2F structure, a key architectural component introduced in YOLOv8, combines the strengths of the C3 and ELAN modules to enhance feature extraction, thereby improving model accuracy and robustness. However, the C2F architecture primarily focuses on local feature capture and may be less effective in processing global semantic information. This limitation can hinder the model's ability to differentiate between subtle category variations when dealing with complex wood defects, potentially affecting classification or detection accuracy.

To address this, a semantic segmentation module was introduced. This module classifies each pixel in the image, providing finer semantic information that aids in a deeper understanding of image content and improves the detection of tiny targets. Fusing the output of the semantic segmentation module with the C2F module's features enhances feature representation richness and enables the model to more effectively distinguish between different categories and instances.

In this study, the BottleNeck module within the C2F structure was improved by designing the C2F\_DWR structure and incorporating the MSDA multiscale null attention mechanism. This structure is capable of extracting multi-scale features, which are essential for detecting targets of varying sizes. The combination of DWR residuals and the MSDA attention mechanism allows the model to extract more discriminative features for defect detection and adaptively select features that are more relevant for defect classification, thus improving classification ability and robustness.

#### *DWRseg residual linkage*

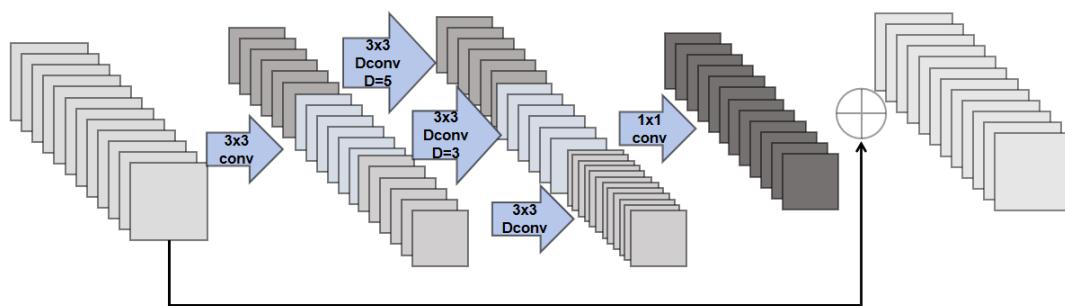
The Dilatation-wise Residual Module (DWR) is an efficient multi-scale context information extraction technique primarily used in the field of real-time semantic segmentation. The module's structure, as illustrated in Figure 3, employs a residual structure to efficiently extract multi-scale contextual information through a two-stage approach, fusing this information to generate a feature map with multi-scale receptive fields (Wei *et al.* 2022).

In the first stage, streamlined feature maps of varying sizes were generated through regional residualization to establish a foundation for semantic residualization in the second

stage. This process was achieved using a standard 3x3 convolution operation combined with a batch normalization (BN) layer and a ReLU activation layer. The 3x3 convolution operation was responsible for initial feature extraction.

Subsequently, in the second stage, semantic residualization, multirate depth-separable convolution was used for morphological filtering of regional features, ensuring that each channel feature utilized only one appropriate receptive field. In the first stage, depending on the desired receptive field size, the network selectively learned the appropriate streamlined regional feature map for efficient matching.

To accomplish this, the regional feature maps were first divided into different step groups, followed by the application of dilation depth convolution with varying rates to these groups. Different expansion rates and convolution capacities were designed for different network stages to fully leverage the varying feature map sizes produced at each stage. This design transformed the function of multirate depth-separable convolution from complex semantic information extraction to simple morphological filtering, thereby improving the efficiency of multi-scale contextual information extraction (Zhao *et al.* 2024).



**Fig. 3.** Dual-scale Wide Residual Architecture

#### MSDA multiscale void attention

In the field of wood panel defect detection, some defects are often obscured by the features of large targets due to their small size and inconspicuous features, leading to poor performance of target detection models such as YOLOv8 in recognizing small targets. To address this problem, a multi-scale dilation attention (MSDA) mechanism was incorporated into the C2F structure to improve the accuracy of the YOLOv8 model in detecting small targets of wood panel defects. The MSDA mechanism was designed to capture multi-scale features by configuring different dilation rates in different detection heads, effectively strengthening the model's ability to recognize the features of small targets and enhancing the accuracy of its localization.

The proposed MSDA was based on the Sliding Window Expansion Attention (SWDA) method. In this method, representative keys and values were selected for sparsity within a sliding window. Subsequently, these selected patches were weighted by performing the self-attention mechanism to obtain attention scores (Saito *et al.* 2019). The formula for this attention is as follows.

$$X = \text{SWDA}(Q, K, V, r) \quad (1)$$

$$\begin{aligned}
 x_{ij} &= \text{Attention}(q_{ij}, K_r, V_r), \\
 &= \text{Softmax}\left(\frac{q_{ij} K_r^T}{\sqrt{d_k}}\right) V_r, \quad 1 \leq i \leq w, 1 \leq j \leq h,
 \end{aligned} \tag{2}$$

The proposed MSDA was built upon SWDA by dividing the feature map into windows at multiple scales through value inflation. The sliding window expansion attention operation was then performed on these windows. Finally, the outputs from different windows were stitched together, and feature aggregation was carried out using a linear layer, as follows.

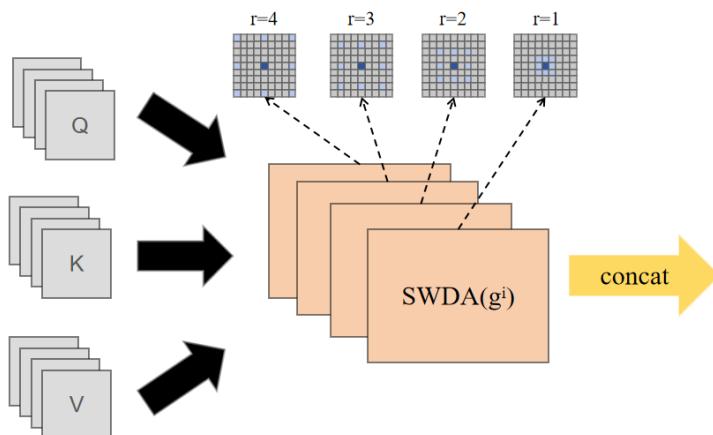
$$\begin{aligned}
 h_i &= \text{SWDA}(Q_i, K_i, V_i, r), \quad 1 \leq i \leq n, \\
 X &= \text{linear}(\text{Concat}[h_1, \dots, h_n])
 \end{aligned} \tag{3}$$

The features were passed to a linear layer for aggregation. Different expansion rates were configured for individual heads. This multi-scale feature aggregation effectively integrated semantic information at different scales within the supervised region, significantly reducing redundancy in the self-attention mechanism, while avoiding complex operations and additional computational overhead (Jiao *et al.* 2023).

Adaptive gating was incorporated into the MSDA null attention mechanism. The gating weights were calculated based on the similarity between features. Additionally, an initial weight was assigned to each type of defective feature by learning the defective label information in the dataset. This initial weight was continuously updated during the model training process, resulting in an adaptive weight that reflected the importance of different defective features (Nie *et al.* 2024). When performing attention calculations, the original attention weight was multiplied by the adaptive weight  $g_i$  to obtain the final attention weight. This allowed the model to focus more on defect features with higher importance, thereby improving detection accuracy. The formula is shown below.

$$x_{ij} = \text{Attention}(q_{ij}, K_r, V_r) g_i \tag{4}$$

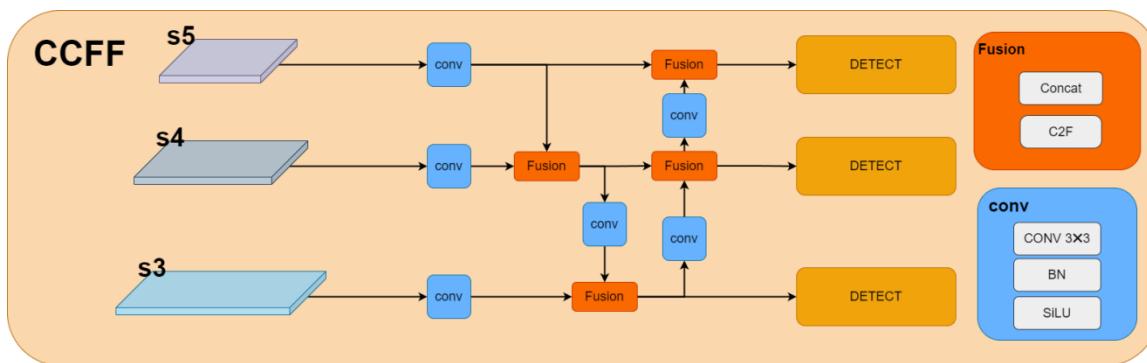
$$h_i = \text{SWDA}(Q_i, K_i, V_i, r, g), \quad 1 \leq i \leq n, \tag{5}$$



**Fig. 4.** Multi-Scale Dilated Attention Architecture

### Improvement of the neck structure

Multiple integration modules consisting of convolutional layers were embedded into integration paths, which is a common technique in deep learning models. The primary objective of this approach was to enhance the model's representational ability and overall performance by combining different feature maps. The CCFF (Cross-scale Context Fusion Framework) architecture was optimized based on the cross-scale integration module, which involved inserting multiple integration modules composed of convolutional layers into the integration path. These integration modules were designed to efficiently integrate the feature maps of two neighboring scales, generating a new feature map (Lv *et al.* 2024). Given the varying receptive fields of different convolutional layers, they were capable of capturing feature information at different scales. By strategically designing and arranging these integration modules, a more refined and comprehensive multi-scale feature integration was achieved. This fine-grained multi-scale feature integration mechanism enabled the model to more accurately capture the detailed and contextual information of the target object, leading to improved performance in various visual tasks. Its structure is shown in Fig. 4.



**Fig. 5.** Cross-scale Integration Module of CCFF

### Improving the loss function

IoU (Intersection over Union) was used as a target detection evaluation metric. It quantified the overlap between the predicted bounding box and the ground-truth box. The IoU loss function directly optimized the model by minimizing the difference between IoU and 1. However, this loss function suffered from a vanishing gradient problem when IoU approached 0 or 1, hindering network convergence. Moreover, IoU only considered the overlapping area, neglecting other geometric information such as the center point distance and aspect ratio.

YOLOv8 employed CIoU as the loss function for bounding box regression. CIoU not only considered the overlap area but also incorporated the centroid distance and aspect ratio, leading to more accurate and efficient bounding box regression (Zheng *et al.* 2021). By addressing the limitations of IoU, CIoU significantly improved the model's performance:

$$CIoU = IoU - \left( \frac{P^2}{C^2} \right) - aV \quad (6)$$

$$V = \frac{4}{\pi^2} \left( \arctan\left(\frac{w-gt}{h-gt}\right) - \arctan\left(\frac{w}{h}\right) \right)^2 \quad (7)$$

$$a = \frac{V}{1 - IoU + V} \quad (8)$$

In Eq. 6,  $p^2$  is the square of the Euclidean distance between the centroid of the predicted and real frames, whereas  $v$  is a measure of the consistency of the aspect ratio.

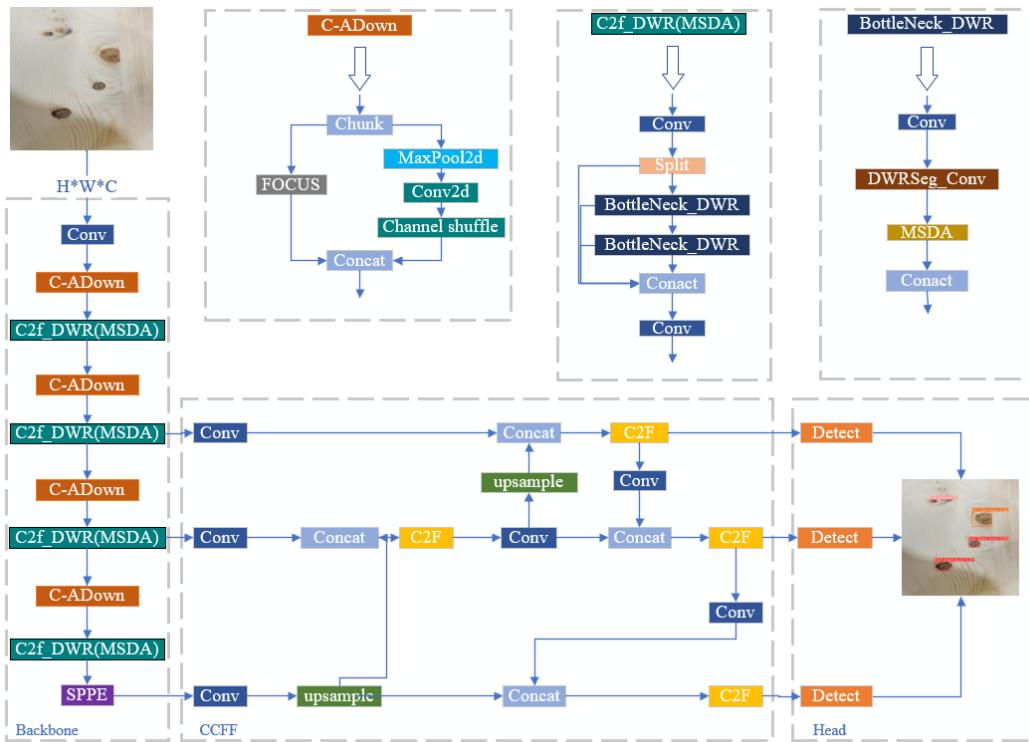
CIoU assumed that the aspect ratio of the target bounding box was a continuously varying value and used a trigonometric function to calculate its consistency. In wood defect detection, defects tended to occupy a small percentage of the area, leading to difficult samples. Conversely, defect-free areas were considered simple samples. Very small defects were challenging to accurately locate due to their size and inconspicuous features. The distribution of aspect ratios in the target bounding box might be more complex, and the CIoU trigonometric function might not accurately capture this complexity. Therefore, attention needed to be paid to bounding box regression for difficult samples.

The FocalerIoU introduced in this paper reconstructed the IoU loss through a linear interval mapping method, which could focus more on different types of samples. The weights of different samples in the loss function were adaptively adjusted according to the number of different defect types and the difficulty of detection (Zhang and Zhang 2024). This approach effectively addressed the challenges posed by difficult samples in wood defect detection.

$$IoU^{focaler} = \begin{cases} 0, & IoU < d \\ \frac{IoU - d}{u - d}, & d \leq IoU \leq u \\ 1, & IoU > u \end{cases} \quad (9)$$

where  $IoU^{focaler}$  is the reconstructed Focaler-IoU,  $IoU$  is the original IoU value, and  $[d, u] \in [0, 1]$ . By adjusting the values of  $d$  and  $u$ , the  $IoU^{focaler}$  can be made to focus on different regression samples. Its loss is defined as follows:

$$L_{focaler-IoU} = 1 - IoU^{focaler} \quad (10)$$



**Fig. 6.** YOLOv8-CDC structure diagram

The Focaler-IoU loss was combined with the CIoU loss to form the Focaler-CIoU loss. This new loss function incorporated a focusing mechanism based on the CIoU loss, assigning greater weight to hard samples. By doing so, the model was encouraged to pay more attention to these difficult samples, thereby improving the detection accuracy. The Focaler-CIoU loss formula is as follows:

$$L_{focaler-CIoU} = L_{Ciou} + IoU - IoU^{Focaler} \quad (11)$$

In the wood surface defect detection dataset used in this paper, a portion of the defect targets belong to small-sized targets, accounting for approximately one-third of the total targets in the dataset. Therefore, considering the training perspective of the dataset, this paper incorporated Focaler-CIoU as part of the primary loss function to guide the model's bounding box regression task. During the actual training process, the network calculates the CIoU loss and Focaler-CIoU loss for each detection box and dynamically combines them to form the final Focaler-CIoU loss. The improved model is named YOLOv8n-CDC, and its structure is shown in Fig . 6.

### Wood Panel Defects Database

This experiment utilized a wood panel dataset uploaded by individual users of roboflow. As shown in Fig. 7, the dataset contained four common types of wood surface defects: dead knot, live knot, crack, and scar. A total of 950 defect images were included. To expand the dataset, data augmentation techniques such as flipping, scaling, and cropping were applied. The augmented dataset was then re-labeled, resulting in a training set of 2117 images, a validation set of 248 images, and a test set of 248 images. The content of the dataset is presented in Fig.7. The number of wood panel defects is summarized in Table.1.



**Fig. 7.** Wood Panel Defects database

**Table 1.** Results of Ablation Experiments

Defect category	Quantities
Dead knot	4065
Live knot	3134
Scar	3495
Crack	3278

### Environment Configuration

The experiments were conducted on an NVIDIA GeForce RTX3070ti with 8GB of video memory, using the PyTorch framework. The experimental parameters included 300 epochs, a batch size of 16, and an image size of 640.

Precision, Recall, Average Precision, and mean Average Precision (mAP) were used as evaluation metrics. mAP, which measures the model's performance across various target categories and confidence thresholds, was the primary metric. Precision, also known as the detection rate, evaluates a classification model's performance by measuring the proportion of true positives among predicted positives. Recall measures the proportion of correctly detected targets to all true targets [10]. The formula for the metric is as follows, where TP represents true positives, FP represents false positives, and FN represents false negatives.

$$mAP = \frac{\sum_{d=1}^c AP(d)}{c} \quad (12)$$

$$precision = \frac{TP}{TP + FP} \quad (13)$$

$$recall = \frac{TP}{TP + FN} \quad (14)$$

### Ablation experiment

To assess the specific impact of each module on performance, the YOLOv8 base version was used as the baseline model. Eight sets of experiments were designed to conduct ablation analysis of the modules, ensuring consistent experimental conditions. The improved modules were gradually introduced to construct variant models, and their corresponding performance metrics were obtained. Table 2 compares the performance data of these models, analyzing the specific impact of removed or replaced modules on overall performance.

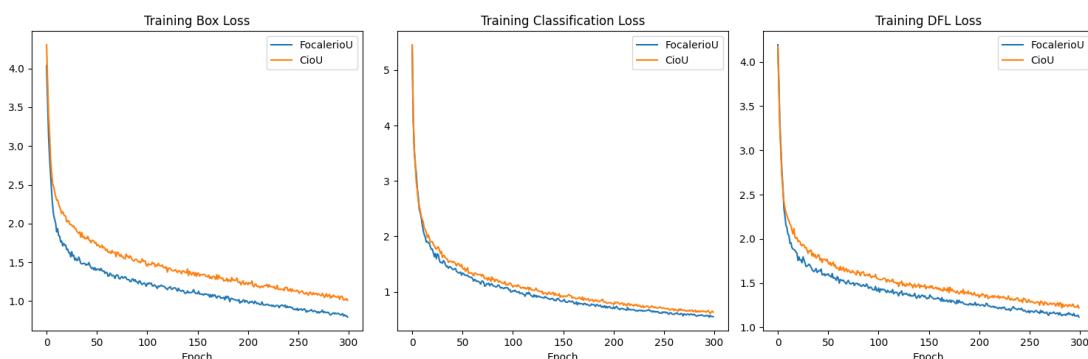
**Table 2.** Results of Ablation Experiments

C-ADown	M-C2F	CCF F	Focaler -CIoU	mAP50 (%)	mAP50 -95 (%)	Recall (%)	Precision (%)	GFLOPs (G)
				78.5	49	67	78.2	8.1
✓				79.3	48.9	73.7	75	<b>7.2</b>
	✓			81.4	49.5	71.6	79.3	9.1
		✓		79.8	48.7	68.8	81.2	7.6
			✓	80.8	48.8	74.1	79.2	8.1
✓	✓			82.9	49.1	72.5	78.5	8.5
✓	✓	✓		83.7	<b>50.3</b>	76.7	78.9	8.0
✓	✓	✓	✓	<b>84.8</b>	50.1	<b>77.5</b>	<b>84.7</b>	8.0

The second set of experiments significantly reduced computational effort by adjusting the downsampling strategy and replacing the traditional convolutional operation with C-ADown, while preserving feature information. This resulted in a 0.8% improvement in mean average precision (mAP) and a significant increase in recall of about 6.7%. The third set of experiments improved the C2F structure by introducing the DWR residual join and MSDA attention mechanism, enhancing the ability to capture multi-scale features, and contributing significantly to the improvement of precision and mAP. This resulted in a 4.4% improvement in mAP. The fourth group of experiments improved the NECK structure by incorporating the CCFF, which effectively fused features at different scales, enhancing feature fusion while reducing computational effort. The fifth group of experiments improved the loss function, accelerating model convergence speed without increasing computational cost. Ultimately, compared to the base version of YOLOv8, the improved model achieved a 6.3% improvement in mAP, a 10.5% improvement in recall, and a 6.5% improvement in precision. These improvements effectively enhanced the model's accuracy, demonstrating a good synergy between the improved modules.

#### *Loss function comparison*

When comparing the loss functions of the base and improved models, as shown in Fig. 8, the CIoU loss function had limitations in its performance due to its insufficient consideration of the balance of sample difficulty. By introducing the Focaler-IoU loss function on top of CIoU, the model was able to focus more on processing difficult samples. This improvement significantly accelerated the convergence of the model and reduced the overall loss value, thereby enhancing the overall performance.

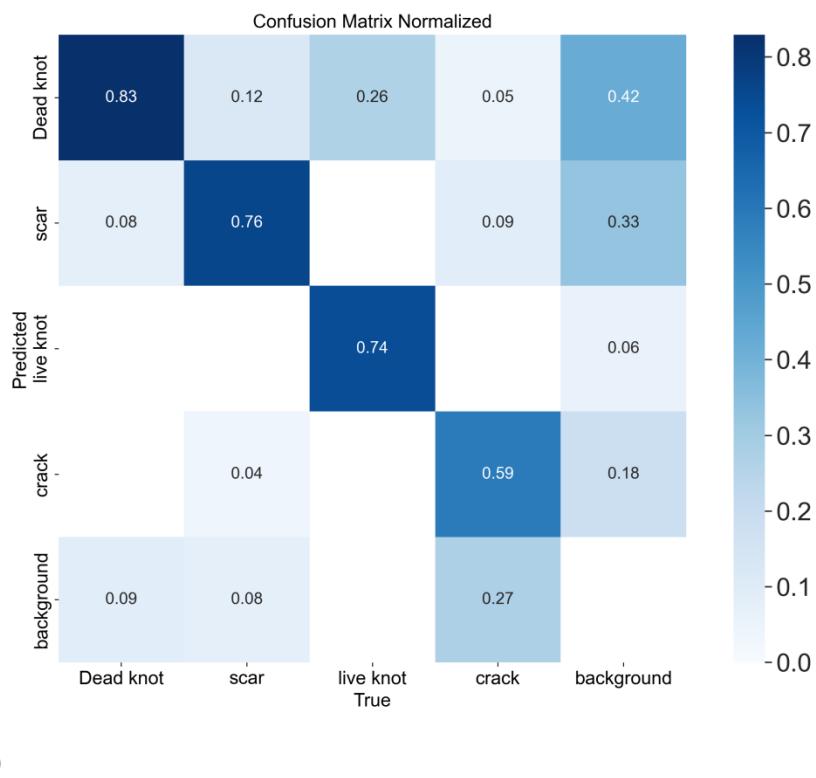


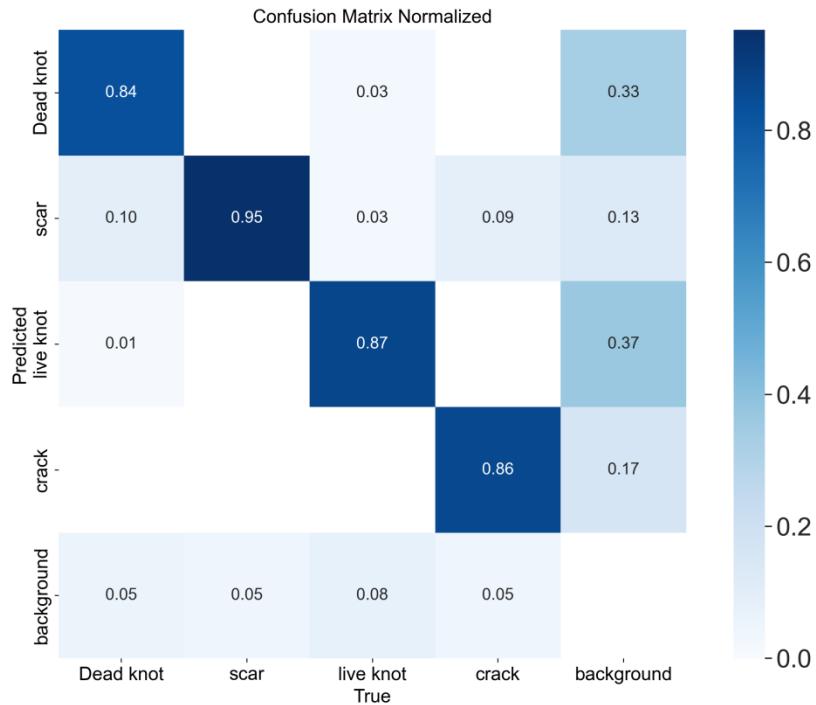
**Fig. 8.** Comparison of training loss functions*Comparison of different defects*

A confusion matrix was generated by comparing the actual categories of the validation set with the predicted categories of the model, as shown in Fig. 9. This figure compares the confusion matrix plots of the different algorithms. The true-negative squares on the diagonal represent the categories correctly predicted by the model. The other squares represent misdetections and false detections. The values on the diagonal of Fig. 8(b) were consistently higher than those on the diagonal of Fig. 8(a), especially for live knots and crack defects. This indicates a more significant improvement of the improved algorithm over the original algorithm. A comparison of the results is shown in Figs. 10 and 11.

*Comparison of different algorithms*

To demonstrate the superiority of the proposed algorithm in wood panel defect detection, it was compared with other mainstream algorithms. The comparison results are presented in Table 3. The proposed algorithm outperformed YOLOv3-tiny by 15.4% in accuracy and 6.2% in mAP. Compared to YOLOv5, it achieved a 10% increase in accuracy and a 6.8% increase in mAP, with a slight increase in the number of parameters. When compared to the larger YOLOv8s, the proposed algorithm exhibited a 4.5% increase in accuracy and a 1.9% increase in mAP. While the proposed algorithm demonstrated comparable performance to some of the improved algorithms, it surpassed them in the detection of dead knots, dry scars, and live knots. However, its performance in crack detection was slightly lower than that of wood-net. In summary, the proposed algorithm exhibited strong performance compared to both mainstream YOLO algorithms and other improved algorithms. It demonstrated superior accuracy in detecting scar and live knot defects Table 3 Comparison of different algorithms.



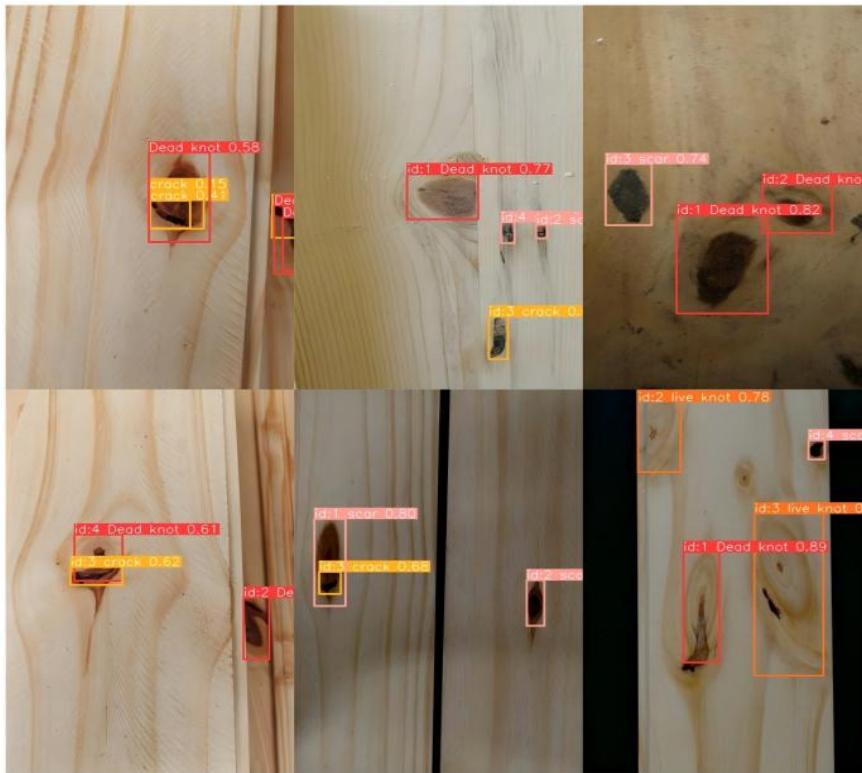


(b)

**Fig. 9.** (a) Confusion matrix for the YOLOv8 model (b) Confusion matrix for the improved model



**Fig. 10.** YOLOv8 defect detection results



**Fig. 11.** Improved algorithm defect detection results

**Table 3.** Comparison of Different Algorithms

Arithmetic	AP (%)				mAP50 (%)	mAP50 -95/%	Precision (%)	Recall (%)	GFLOPs (G)	FPS
	Dead knot	Scar	Live knot	Crack						
SSD	79.5	78.2	76	70.5	73.6	44.5	67.2	74.5	92	42.6
faster-RCNN	82.6	81.5	79.1	77.1	79.2	48.9	72.5	75.2	108	14.3
YOLOv3-tiny	85.7	79.6	75.2	73.8	78.6	47.5	69.3	73.8	19	<b>85.2</b>
YOLOv5	84.3	75.9	77.1	74.5	78	47.4	74.7	70.3	<b>7.1</b>	54.3
YOLOv6	84.5	90	74.9	69.7	79.8	46.8	70.9	73.1	11.8	43.1
YOLOv7	85.7	79.6	75.2	73.8	78.6	48.6	69.3	73.8	19	44.5
YOLOv8n	86.5	84.1	71.7	71.8	78.5	49	78.2	67.2	8.1	81.3
YOLOv8s	86.7	81.2	83	79.8	82.7	<b>50.8</b>	80.2	71.7	28.4	69.7
YOLOv5_HS	85.9	88.3	82.7	71.4	82.1	49.8	78.9	76	23.3	60.8
Wood-net	83.9	89.5	83.9	<b>76.2</b>	83.8	50.6	78.6	<b>78.7</b>	16.4	21.3
YOLOv8n-CDC	<b>88.7</b>	<b>92.2</b>	<b>85.3</b>	75.9	<b>84.8</b>	50.1	<b>84.7</b>	77.5	8.0	58.4

#### Comparison of versatility

To verify the generalization of the improved model, comparative experiments were conducted on the Roboflow public dataset. While maintaining the training parameters of the model, comparative experiments were also performed on related domain public datasets. The results of these comparative experiments are presented in Tables 4 and 5.

*Dataset 1* (<https://universe.roboflow.com/rtech/wood-surface-defects>)

*Dataset 2* (<https://universe.roboflow.com/laila-hammad-cxqz8/wood-jxhd5>)

**Table 4.** Comparison of Generalizability across Dataset 1

arithmetic	Precision/%	Recall/%	mAP50/%
YOLOv8n	47.3	<b>45.8</b>	46.1
YOLOv8n-CDC	<b>53.3</b>	45	<b>48.2</b>

**Table 5.** Comparison of Generalizability across Dataset 2

Arithmetic	Precision/%	Recall/%	mAP50/%
YOLOv8n	33.7	48.9	41.8
YOLOv8n-CDC	<b>34.9</b>	<b>55.6</b>	<b>43.6</b>

## CONCLUSIONS

1. To address the challenges of defect detection in wood panels caused by leakage and misdetection, an improved YOLOv8-based model was proposed in this study. Extensive ablation experiments and comparative evaluations with various existing models were conducted to effectively mitigate these issues.
2. The proposed model replaces the traditional downsampling approach in the backbone network with C-ADown and integrates the DWR module, which combines the MSDA null attention mechanism to enhance the C2F and bottleneck modules in YOLOv8. This modification allows for improved accuracy in detecting features across multiple scales. Additionally, the NECK structure is enhanced by introducing the CCFF hybrid coding framework, which facilitates intra-scalar, inter-scalar, and cross-scalar feature interactions. The loss function is also refined to further optimize the model's performance.
3. Future research will aim to address the limitations of the detection head, particularly the risk of overfitting that arises from the relatively small dataset size, which may hinder the model's generalization ability, especially when applied to diverse or unseen data. Additionally, future work will explore more effective feature fusion techniques, with a focus on improving the model's ability to detect small and densely packed targets. Furthermore, model compression strategies, such as pruning and distillation, will be investigated to enable real-time deployment on embedded devices. However, the potential trade-offs in model accuracy and robustness will be carefully evaluated, as these compression techniques could compromise the overall performance.
4. A critical aspect of the model's future development will be its generalizability across various industrial domains. Although the current improvements provide promising results for wood panel defect detection, the scalability of the model to other materials and industries will require thorough validation. The model's robustness in handling data from different sources, variations in environmental conditions, and diverse defect types needs to be systematically assessed. This includes extending its application to industries such as construction, furniture manufacturing, and packaging, where similar challenges in defect detection and quality control exist. Further refinement of the model, including addressing its sensitivity to data variation and incorporating domain adaptation techniques, will be essential to ensure its broad applicability and reliable performance across diverse operational contexts.

**REFERENCES CITED**

Cheng, D. (2023). *Research and Application of Deep Learning Based Wood Surface Defect Detection*, Qilu Univ. Technology. DOI: 10.27278/d.cnki.gsdqc.2023.000196

Cheng, R. (2020). "The influence of market orientation on the development of wood processing industry and countermeasures," *China Forest Products Industry* 57(11), 88-89+92. DOI: 10.19531/j.issn1001-5299.202011022

Buehlmann, U., and Thomas, R. E. (2002), "Impact of human error on lumber yield in rough mills," *Robotics and Computer-Integrated Manufacturing*, 18(3-4), 197-203. DOI: 10.1016/S0736-5845(02)00010-8

Jia, H., Xu, H., Wang, L., Zhang, J., Chu, X., and Tang, X. (2023). "Quantitative identification of surface defects in wood lumber based on improved YOLOv5," *J. Beijing Forestry University* 45(04), 147-155. DOI: 10.12171/j.1000-1522.20220419

Jiang, X., and Zhao, X. (2024). "Improved YOLOv7 algorithm for wood surface defect detection," *Computer Engineering and Applications* 60(07), 175-182. DOI: 10.3778/j.issn.1002-8331.2309-0185

Jiao, J., Y.-M. Tang, K.-Y. Lin, Y. Gao, A. J. Ma, Y. Wang, and W.-S. Zheng. (2023). "Dilateformer: Multi-scale dilated transformer for visual recognition," *IEEE Transactions on Multimedia* 25, 8906-8919. DOI: 10.1109/TMM.2023.3243616

Liu, Q., Yuan, Y., Xia, X., Si, L., and Duo, H. (2023). "Research advances in wood defect detection based on artificial intelligence," *World Forestry Research* 36(01), 66-71. DOI: 10.13348/j.cnki.sjlyyj.2022.0101.y

Lv, W., Y. Zhao, Q. Chang, K. Huang, G. Wang, and Y. Liu (2024). "RT-DETRv2: Improved baseline with bag-of-freebies for real-time detection transformer," *arXiv preprint arXiv:2407.17140*. DOI: 10.48550/arXiv.2407.17140

Nie, F., Li, M., Zhou, M., Dong, Y., Li, Z., and Li, L. (2024). "Multiscale dilated U-Net based multifocus image fusion algorithm," *Laser & Optoelectronics Progress* 61(14), 447-456. DOI: 10.3788/LOP232443

Saito, K., Y. Ushiku, T. Harada, and K. Saenko (2019). "Strong-weak distribution alignment for adaptive object detection," in: *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 6956-6965. DOI: 10.48550/arXiv.1812.04798

Urbonas, A., Raudonis, V., Maskeliūnas, R., and Damaševičius R. (2019). "Automated identification of wood veneer surface defects using faster region-based convolutional neural network with data augmentation and transfer learning," *Applied Sciences* 9(22), article 4898. DOI: 10.3390/app9224898

Varghese, R., and Sambath, M. (2024). "YOLOv8: A novel object detection algorithm with enhanced performance and robustness," in: *2024 International Conference on Advances in Data Engineering and Intelligent Computing Systems (ADICS)*. DOI: 10.1109/ADICS58448.2024.10533619

Wang, Z., Jiang, Y., Yan, F., Sun, Y., Zhang, Y., and Zhang, L. (2024a). "Research on wood defect detection model Wood-Net based on YOLOv7," *Journal of Forestry Engineering* 9(01), 132-140. DOI: 10.13360 /j.issn.2096-1359.202305016

Wang, M., Xiang, X., Cui, W., Yuan, M., and Duo, H. (2024b). "Research progress and prospect of intelligent detection of wood defects based on deep learning," *China Forest Products Industry* 61(03), 38-44. DOI: 10.19531/j.issn1001-5299.202403006

Wang, C., Yeh, I., and Liao, H. (2024c). "YOLOv9: Learning what you want to learn

using programmable gradient information,” *arXiv preprint arXiv:2402.13616*. DOI: 10.48550/arXiv.2402.13616

Wang, J., Chen, X., and Xia, D. (2013). “Research progresses on elasticity modulus nondestructive examination of wood and glulam structures,” *J. Central South Univ. of Forestry & Technol.* 33(11), 149-153. DOI: 10.14067/j.cnki.1673-923x.2013.11.005

Wei, H., X. Liu, S. Xu, Z. Dai, Y. Dai, and X. Xu. (2022). “DWRSeg: Rethinking efficient acquisition of multi-scale contextual information for real-time semantic segmentation,” *arXiv preprint arXiv:2212.01173*. DOI: 10.48550/arXiv.2212.01173

Yang, F., Yang, B., and Li, R. (2023). “Surface defect detection technology of wood-based panel based on image segmentation and deep learning,” *Journal of Zhejiang A&F University* 41(01), 176-182. DOI: 10.11833/j.issn.2095-0756.20230280

Zhang, H., and Zhang, S. (2024). “Focaler-IoU: More focused intersection over union loss,” *arXiv preprint arXiv:2401.10525*. DOI: 10.48550/arXiv.2401.10525

Zhao, W., Zhao, Z., and Kong, J. (2024). “Remote sensing image object detection based on inverted residual self-attention mechanism,” *CAAI Transactions on Intelligent Systems* 1-10. DOI: 10.11992/tis.202312001

Zheng, Z., Wang, P., Ren, D., Liu, W., Ye, R., Hu, Q., and Zuo, W. (2021). “Enhancing geometric factors in model learning and inference for object detection and instance segmentation,” *IEEE Transactions on Cybernetics* 52(8), 8574-8586. DOI: 10.1109/TCYB.2021.3095305

Article submitted: October 30, 2024; Peer review completed: November 23, 2024;  
Revised version received and accepted: January 29, 2025; Published: February 10, 2025.  
DOI: 10.15376/biores.20.2.2556-2573