

# A Novel Wood Surface Defect Detection Model Based on Improved YOLOv8

Wenmiao Dou and Jun You \*

To address the challenges posed by complex and variable backgrounds coupled with the small-target characteristics of wood surface defects such as knots and cracks, a novel wood surface defect detection model based on improved You Only Look Once version 8 (YOLOv8) is proposed. The model integrates a multi-head mixed self-attention mechanism into the backbone to improve the representation of fine-grained defect features. A learnable dynamic upsampling module replaces traditional nearest-neighbor interpolation to mitigate feature loss during resolution recovery. Additionally, a structural Re-parameterizable Block is adopted to enhance feature expressiveness during inference, and a small-object detection head is added to enhance the detection of small defects while minimizing both missed and incorrect detections. The experimental results demonstrate that the proposed model effectively enhances detection performance, increasing the mAP of the baseline model from 72.9% to 79.5%. Furthermore, the proposed model surpasses other YOLO variants in mAP across all defect categories. This improvement better meets the quality control requirements of wood processing and manufacturing, ensuring the quality of wood products.

DOI: 10.15376/biores.20.3.5709-5730

**Keywords:** Wood surface defect detection; You Only Look Once; Multi-head Mixed Self-Attention; Dynamic upsampling; Structural re-parameterizable block

**Contact information:** School of Electronic Engineering, Guilin Institute of Information Technology, 541101, Guilin, China; \*Corresponding author: giit\_youjun@126.com

## INTRODUCTION

Wood surface defects significantly affect the quality and efficiency of wood processing. These defects not only reduce the utilization of wood, causing resource waste, but they also have a considerable negative impact on the mechanical properties and functional value of wood products (Chen *et al.* 2023a; Li *et al.* 2024). Therefore, surface defect detection is a critical step in wood production and classification. Timely detection and removal of defective wood can effectively improve the quality of wood products, maximize the utilization of wood resources, and thus promote the sustainable development of the wood industry (Yi *et al.* 2024).

In recent years, wood processing has gradually shifted from manual labor to automation, mechanization, and intelligence. In this process, various non-destructive testing (NDT) technologies have been widely applied in wood surface defect detection, such as ultrasonic testing (Jiang *et al.* 2024; Wang *et al.* 2024; ), X-ray inspection (Stängle *et al.* 2015; Zhang 2017), infrared detection (López *et al.* 2014; Yu *et al.* 2019), and machine vision techniques (Hittawe *et al.* 2017; Ji *et al.* 2024a). These technologies provide practical solutions for wood defect detection, successfully facilitating the transition

from manual to automated and mechanized processes. For instance, Connors *et al.* (1983) designed an automated wood processing system based on computed tomography and optical scanning, capable of identifying and classifying eight common wood defects. Wyckhuysse and Maldague (2001) experimentally validated the feasibility of infrared thermography for wood surface defect detection. Sandak *et al.* (2020) developed a portable spectrometer covering visible and near-infrared light, which can directly detect defects such as knots, decay, and resin on the surface of logs in the forest. Although these NDT technologies offer high precision and efficiency, their adaptability and real-time capabilities in complex environments still face significant limitations.

With the rapid development of computer and artificial intelligence technologies, wood surface defect detection technology has evolved from traditional rule-based methods to data-driven deep learning approaches. Deep learning-based detection techniques, due to their ability to accurately identify a variety of wood surface defect types and adapt to different lighting conditions and wood species, have become a current research hotspot. For instance, Sun *et al.* (2022) proposed a multicriteria framework that integrates deep learning for comprehensive wood quality assessment, whereas Ji *et al.* (2024b) concentrated on knot detection in Chinese fir lumber using traditional vision methods. Özcan *et al.* (2024) applied deep learning for general anomaly detection on wood surfaces but without structural enhancements for small defect sensitivity. Zhu *et al.* (2024) introduced a multi-source data fusion network targeting fine-grained defect segmentation. Currently, deep learning-based defect detection methods are mainly divided into two-stage and single-stage detection algorithms. Two-stage algorithms, such as Region-based Convolutional Neural Network (R-CNN) (Girshick *et al.* 2014), Faster R-CNN (Ren *et al.* 2017), and Mask R-CNN (He *et al.* 2017), perform excellently in detection accuracy and are widely used in wood surface defect detection tasks (Gao *et al.* 2021; Li *et al.* 2021; Chen *et al.* 2023b; Zou *et al.* 2024). Fan *et al.* (2019) were the first to apply Faster R-CNN to wood defect detection, constructing a real-time defect detection system for solid wood flooring that meets industrial production requirements. They validated the practicality of the multi-stage Faster R-CNN object detection algorithm under deep learning for solid wood board defect detection. Xia *et al.* (2022) noticed that the texture features of wood often accompany wood defects and can interfere with the final recognition results. They proposed a Faster R-CNN surface defect detection algorithm that improves image texture background using bilateral filtering, enhancing the network's ability to process multi-scale defect features and achieving outstanding performance in detecting small defects. Hu *et al.* (2020) combined Progressive Growing of Generative Adversarial Networks (PGGAN) with the Mask R-CNN model and introduced transfer learning to identify and classify defects in poplar veneer, compensating for the traditional sample augmentation methods that lack fine defect details, poor defect image diversity, and limited sample distribution. However, two-stage algorithms typically have high computational complexity, making it difficult to fully meet the real-time requirements of wood processing.

In contrast, single-stage detection algorithms, such as You Only Look Once (YOLO) (Redmon *et al.* 2016) and its series of versions, as well as Single Shot MultiBox Detector (SSD) (Liu *et al.* 2016), convert the object detection problem into an end-to-end prediction task. This eliminates the cumbersome process of generating candidate regions, offering advantages in real-time performance, simple architecture, and computational efficiency. Wang *et al.* (2023) introduced a wood surface defect detection method called Omni-Dynamic Convolution Coordinate Attention-based YOLO (ODCA-YOLO), which incorporates an Omni-dimensional dynamic convolution-based coordinate attention

(ODCA) mechanism. This method effectively enhances the detection capability for small defect targets and was experimentally validated using an optimized wood surface defect dataset, fulfilling the practical requirements for accurate wood surface defect detection. Meng and Yuan (2023) proposed a YOLOv5 model based on a Semi-Global Network (SGN) for wood defect detection. By integrating a lightweight SGN into the backbone network to model global context, the method improves detection accuracy while reducing model complexity. Effectiveness was validated on a public wood defect dataset, significantly enhancing detection performance across various types of defects. Ding *et al.* (2020) utilized machine vision and deep learning techniques to detect three types of wood surface defects: live knots, dead knots, and cracks. They applied transfer learning to the SSD object detection algorithm and improved it by incorporating a DenseNet network, addressing the issues of high labor costs and low efficiency in wood defect detection. Furthermore, YOLO-based algorithms have demonstrated considerable potential in various fields. For example, Karimi *et al.* (2024) developed an automated defect detection system for Portuguese cultural heritage buildings, specifically targeting tile defects using YOLO. Additionally, Mishra and Lourenço (2024) offered a comprehensive review of artificial intelligence-assisted visual inspection techniques, emphasizing their application in the monitoring and preservation of cultural heritage (CH) sites. These studies highlight the effectiveness of YOLO-based frameworks in detecting small-scale defects, thereby reinforcing the relevance of our approach in optimizing YOLO for wood defect detection tasks.

However, wood, being a natural material, exhibits highly diverse grain patterns and structures, resulting in a complex and variable background that often leads to confusion between defect regions (especially small defects) and the inherent grain patterns of the wood. The variability in lighting conditions and the irregularities of the wood surface (such as knots, cracks, *etc.*) further complicate the background, making it challenging to distinguish defect regions from the surrounding wood. Moreover, wood surface defects are typically small in size, with many defects (*e.g.*, fine cracks, stains, knots) measuring just a few millimeters or even smaller. These defects often occupy only a small portion of the image and fall within the category of small-object detection. Small defects on wood surfaces tend to have low contrast, particularly in regions with dense natural grain patterns, where they may lack distinct edges or shapes. This makes the task of localization and classification more difficult, leading to a higher likelihood of false negatives (missed detections) and false positives (incorrect detections). These factors collectively render the detection of wood surface defects exceptionally challenging. To address the aforementioned issues, this study proposes an improved YOLOv8-based method for wood surface defect detection. This approach provides an accurate and efficient solution for the application of wood surface defect detection technology in the manufacturing industry. The main contributions of this paper are as follows:

1. A novel lightweight Multi-head Mixed Self-Attention (MMSA) module is designed and seamlessly integrated into the C2f module, resulting in the C2f-MMSA module. This integration significantly enhances the model's capacity to capture contextual and background information for small targets, effectively overcoming the limitations of the original C2f module in detecting small defects within complex backgrounds.
2. A learnable dynamic upsampling module is introduced to replace the upsampling module based on nearest-neighbor interpolation, alleviating the issue of feature information loss for small-scale wood surface defects during upsampling, and improving the model's feature representation capability and precision.

3. A Re-parameterizable Block is designed to accurately capture small-scale and multi-scale features in complex backgrounds, further exploring the fine-grained and multi-scale characteristics of wood surface defects.

## EXPERIMENTAL

### Wood Surface Defects Dataset

The dataset used in this study consists of wood surface defect images collected from real-world industrial environments. A custom-designed imaging system was employed for data acquisition (the field of view is 15 cm × 500 cm), capturing 10 categories of wood surface defects that comprehensively cover common defect types encountered in industrial settings (Kodytek *et al.* 2021). To mitigate potential biases arising from class imbalance and limited sample sizes, the dataset was refined to include seven predominant defect categories: Live\_Knot, Dead\_Knot, Knot\_with\_crack, Knot\_missing, Crack, Marrow, and Resin, comprising a total of 4,500 annotated images with a resolution of 2800×1024. The dataset was split into training and test sets in a 9:1 ratio, with 10% of the training set used for validation. The distribution of each defect type, including number of images, occurrence frequencies, and proportional representation within the dataset, is systematically summarized in Table 1.

**Table 1.** Distribution of Each Defect Type in Dataset

Defect Type	Number of Images with Defect	Number of Defect Occurrences	Images in Dataset (%)
Live_Knot	2956	5112	65.69
Dead_Knot	2075	3574	46.11
Knot_with_crack	327	574	7.27
Knot_missing	152	208	3.38
Crack	479	671	10.64
Marrow	281	316	6.24
Resin	623	719	13.84

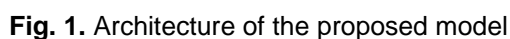
Note: The Images in dataset (%) means the proportion of images with such defects in the dataset.

### YOLOv8 Improvement

To address the aforementioned challenges, this study adopted YOLOv8 as the baseline model. A novel version of the model was developed with a particular emphasis on enhancing small object detection performance. This involved tuning the network's multiscale feature representation and adjusting detection heads to better capture fine-grained details, ensuring that small defects receive sufficient attention during inference. Details are as follows:

In the Backbone, a novel Multi-head Mixed Self-Attention (MMSA) mechanism was designed to effectively integrate channel attention and spatial information, thereby enhancing the representation of wood surface defect features under complex backgrounds. The MMSA module was incorporated into the C2f module to improve the model's ability to capture contextual and background information for small targets, effectively addressing the limitations of the original C2f module in detecting small defects under complex backgrounds.

For the detection Head, an additional small-object detection head with a resolution of  $160 \times 160$  was introduced. This design ensures that fine-grained features of small defects propagate through the downsampling pathway to other feature maps at scales of  $20 \times 20$ ,  $40 \times 40$ , and  $80 \times 80$ , thereby improving detection performance for small defects under complex backgrounds and reducing false positives and missed detections. The architecture of the proposed model is illustrated in Fig. 1.





### C2f-MMSA

The C2f is a key feature capture and fusion module in YOLOv8, employing the Cross Stage Partial Networks (CSPNet) design to enhance feature propagation and fusion (Varghese and Sambath 2024). However, when applied to small object detection tasks, particularly in the context of wood surface defects with complex backgrounds, the C2f structure presents several challenges:

- **Insufficient Fusion of Local and Global Information:** The C2f structure merges features from different layers through cross-stage fusion, with a primary focus on the propagation of local features. However, it does not fully consider global context information, which is crucial for accurate localization of wood surface defects, often embedded within complex backgrounds. These defects typically require global information from surrounding areas for precise detection.
- **Inadequate Focus on Small Targets:** C2f is designed to process larger or more conventional objects through hierarchical feature fusion. However, for small targets, particularly those occupying only a few pixels, the C2f structure may fail to capture fine-grained features adequately. In particular, when feature map sizes are reduced or when information is sparsely distributed, small target information may be lost or incorrectly localized.
- **Separation of Channel and Spatial Information:** While C2f strengthens cross-layer feature fusion, it does not specifically address the relationship between channel and spatial information. In small object detection tasks, the interaction between channel information and spatial information is crucial. However, C2f lacks the mechanisms to dynamically capture these interactions effectively.

To address these issues, this paper introduces a lightweight MMSA (Su *et al.* 2025) mechanism, combining the concepts of multi-head attention and multi-scale attention, inspired by the Efficient Channel Attention (ECA) design pattern. This mechanism effectively integrates local and global information within the image, as well as channel and spatial information, capturing global contextual information across both spatial and channel dimensions. With this mechanism, the model is better equipped to understand the context and background of small targets in the image, thereby overcoming the challenges encountered by the C2f in wood surface defect detection in complex backgrounds. Additionally, the incorporation of the multi-head attention mechanism allows the model to dynamically adjust the attention allocated to different features, enabling adaptive attention distribution across different scales and regions. This enhances the model's ability to focus on small defect regions and facilitates feature extraction across multiple scales.

The principle of the MMSA is illustrated in Fig. 2. First, the input image feature vectors undergo Local Max Pooling (LMP) to extract local spatial information, which is then transformed into a  $1 \times C \times k_s \times k_s$  vector. The structure comprises two branches: one captures global information, while the other focuses on local spatial information. The information extracted by LMP is transformed into a  $1 \times C \times k_s \times k_s$  vector, which is then processed by a 1D convolution (Conv1d). Afterward, de-pooling is applied to recover the original resolution of both vectors, followed by the fusion of attention information from the two branches.

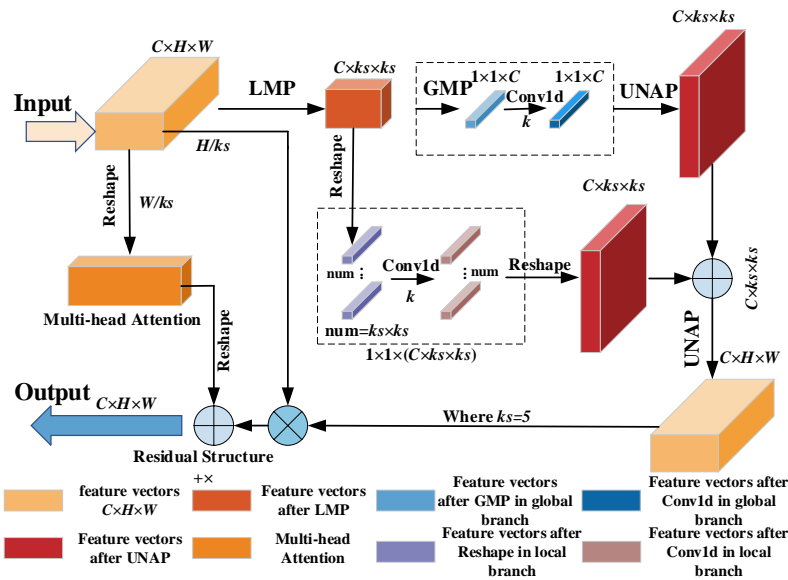


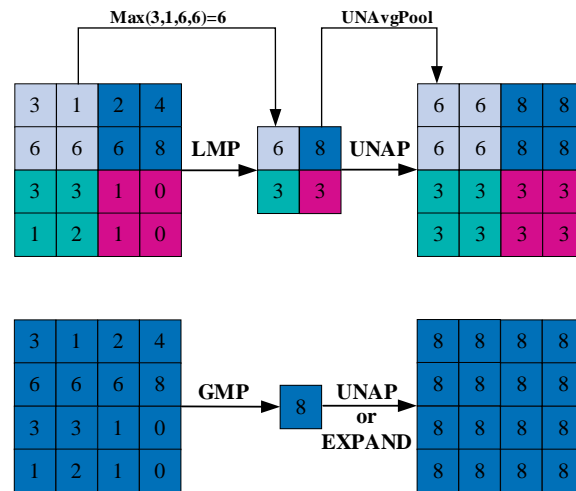
Fig. 2. Block diagram of MMSA

For Multi-head Attention, the input data is reshaped (feature vector transformation), and Multi-head Attention weights are computed to select feature vectors that meet the weight requirements. The feature vectors are then reshaped once again. Finally, the attention outputs from the three components are weighted and fused. The MMSA mechanism combines global channel attention, localized channel attention that refines spatial information, and multi-head attention results.

To achieve the fusion of spatial and channel attention, 1D convolution (Conv1d) as shown in Fig. 2 is employed. The size of the 1D convolution kernel  $K$  is proportional to the number of channels  $C$ . When capturing local cross-channel interactions, only the relationship between each channel and its  $K$  neighboring channels is considered. The selection of  $K$  follows the approach used in ECA (Wang *et al.* 2020).

Figure 3 illustrates the relationships between Global Max Pooling (GMP), LMP, and Unpooling Average Pooling (UNAP) within the MMSA structure. The GMP extracts the global maximum feature, producing a feature map of size  $1 \times 1$ . The LMP divides the entire feature map into  $k \times k$  small regions and performs  $k \times k$  max pooling within each region. The UNAP, also known as reverse pooling, focuses on preserving the attributes of the pooled features while expanding them to the desired size. The UNAP can be implemented using adaptive pooling to ensure that the output size matches the size of the original feature map.

When extending LMP, if the size of the pooled features is not  $1 \times 1$ , a direct expansion operation is not feasible. Instead, the UNAP process must be employed to restore the feature map to its original size. As shown in Fig. 3, UNAP restores the resolution of the original feature map by applying parameters derived from the unpooling operation, filling the corresponding positions with the pooled results.

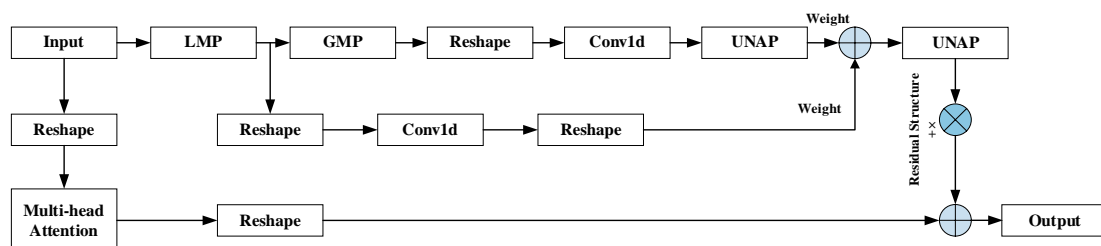


**Fig. 3.** Relationships between GMP, LMP, and UNAP within the MMSA

The structural relationships among LMP, GMP, and UNAP in MMSA are illustrated in Fig. 2 and Fig. 3 and can be summarized as follows:  $\text{LMP} \rightarrow (C, ks, ks) \rightarrow \text{GMP} \rightarrow (1, 1, C) \rightarrow \text{Conv1d} \rightarrow (1, 1, C) \rightarrow \text{UNAP}$ .

When expanding LMP, direct expansion is not feasible if the feature size is not  $1 \times 1$ . To address this limitation, the UNAP process is employed to restore the feature map to its original resolution. Specifically, during the  $\text{LMP} \rightarrow (C, ks, ks) \rightarrow \text{GMP} \rightarrow (1, 1, C) \rightarrow \text{Conv1d} \rightarrow (1, 1, C) \rightarrow \text{UNAP}$  sequence, the UNAP utilizes the parameters from the pooling operation to recover the resolution of the original feature map. Subsequently, the pooled results are placed at their corresponding locations. In this process, a Reshape operation is introduced within the  $\text{LMP} \rightarrow \text{GMP} \rightarrow \text{UNAP}$  pathway to facilitate proper feature alignment.

As depicted in Fig. 2, after extracting global attention and local attention within the MMSA module, the two attention mechanisms are adaptively fused using weighted summation. The resulting fused features are combined with the initial input via a residual connection, followed by integration with the multi-head attention output. This final step produces the output of the MMSA. Additionally, the MMSA adopts the Hard-sigmoid function as the normalization mechanism, effectively mitigating gradient vanishing issues during backpropagation.



**Fig. 4.** Processing flow of the MMSA

The detailed processing flow of the Multi-head Mixed Self-Attention (MMSA) mechanism is illustrated in Fig. 4. It involves the following steps:

First, the feature map undergoes both Local Max Pooling (LMP) and Global Max Pooling (GMP) to extract local and global contextual information. The pooled feature map



is then reshaped into a format suitable for Multi-Head Attention computation. Self-attention is computed based on this reshaped feature representation, capturing dependencies across different spatial regions. The self-attention results are subsequently transformed back to match the original feature map dimensions. Local and global attention weights are derived from the self-attention results, followed by the application of a Hard Sigmoid function to constrain these weights within the range of 0 to 1. To ensure scale consistency, the attention weights undergo adaptive average pooling. Next, the local and global attention weights are fused to construct a comprehensive attention map, which is then applied to the original feature map to emphasize critical features. Finally, the self-attention result is added back to the original feature map, producing the refined output feature map with enhanced feature representation.

The MMSA mechanism is integrated into the C2f, as shown in Fig. 5, and subsequently embedded into the Backbone of YOLOv8 to enhance the model's feature extraction and fusion capabilities.

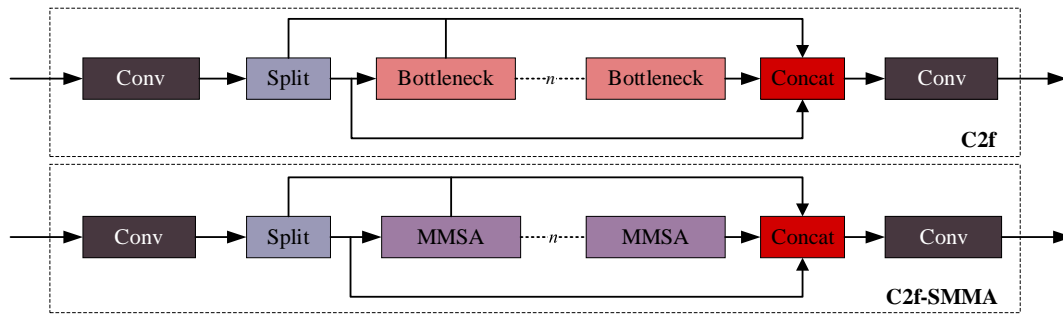


Fig. 5. Structure of C2f-SMMA

#### Dynamic upsampling

The YOLOv8 model employs nearest neighbor interpolation for feature map upsampling, facilitating feature fusion across different layers. While this method is computationally efficient and fast, the simplicity of nearest neighbor interpolation (Wang *et al.* 2019) leads to the loss of fine details, particularly affecting the features of small-scale spatial targets. This results in inadequate feature representation, which, as the network depth increases, severely impacts the detection accuracy of small targets. To address this issue, this paper introduces a super-lightweight, learnable dynamic upsampling (DySample) method (Liu *et al.* 2023), designed to mitigate feature loss and enhance both feature representation and detection accuracy. The specific structure is shown in Fig. 6.

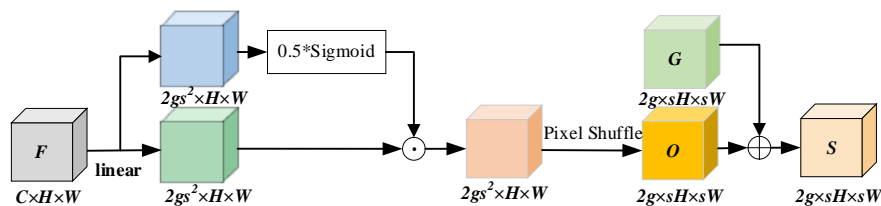


Fig. 6. Structure of DySample

Given the upsampling scale factor  $s$  and a feature map  $F$  of size  $C \times H \times W$ , the feature map is first divided along the channel dimension into  $g$  groups ( $g=4$ ), which helps further reduce computational complexity. A linear layer is then applied to generate offsets

of size  $2gs^2 \times H \times W$ . To increase the flexibility of the offsets, a Sigmoid function along with a static factor of 0.5 is used to produce a per-point “dynamic range factor,” with the dynamic range taking values within the range of  $[0, 0.5]$ . Finally, Pixel Shuffle ( $P_s$ ) is applied to reshape the offset  $\mathbf{O}$  into a size of  $2g \times sH \times sW$ . The mathematical expression for this process is as follows:

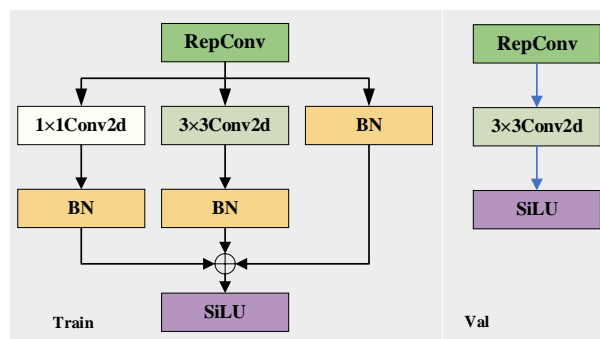
$$\mathbf{O} = P_s \left( 0.5 \times \text{Sigmoid}(\text{linear}(\mathbf{F}) \odot \text{linear}(\mathbf{F})) \right)$$

The sampling set  $\mathbf{S}$  is the sum of the offset  $\mathbf{O}$  and the original sampling grid  $\mathbf{G}$ , i.e.,  $\mathbf{S} = \mathbf{G} + \mathbf{O}$

### Re-parameterizable Block

In the Neck section, although the C2f partially facilitates feature extraction for wood surface defects, it struggles to accurately capture features of small-scale and multi-scale wood surface defects under complex backgrounds. This limitation hinders the effective exploration of fine-grained and multi-scale features. To address these challenges, the structural Re-parameterizable module (RepBlock) is introduced to alleviate these issues. RepBlock leverages a multi-branch structure during the training phase to enrich feature representation, capturing wood surface defect features from global to local scales and from small to large scales, thereby effectively improving detection accuracy.

During the inference phase, the reparameterization technique transforms the multi-branch structure into a more compact single-branch form, effectively accelerating model inference speed without compromising detection performance, thus meeting real-time requirements. The structure of Re-parameterizable Convolution (RepConv) during both training and inference phases is illustrated in Fig. 7. During training, the multi-branch structure consists of a  $3 \times 3$  convolution, a  $1 \times 1$  convolution, a residual structure, and batch normalization (BN) layers. After reparameterization, it is converted into a single-branch  $3 \times 3$  convolution for inference.



**Fig. 7.** Structure of RepConv during both training and inference phases

The structural reparameterization of RepConv is illustrated in Fig. 8. The specific steps are as follows:

(1) Convert the  $1 \times 1$  convolution and residual structure into a  $3 \times 3$  convolution. In the convolution conversion step, the original  $3 \times 3$  convolution remains unchanged, while the  $1 \times 1$  convolution is transformed into a  $3 \times 3$  convolution by padding zeros around it. The residual structure can be constructed with four convolution kernels, two of which have center values of 1, and the remaining two are set to 0. The output of the input feature matrix, processed through these four convolution kernels, is identical to the input.

(2) Fusion of the BN layer with the convolution layer: The structure is transformed

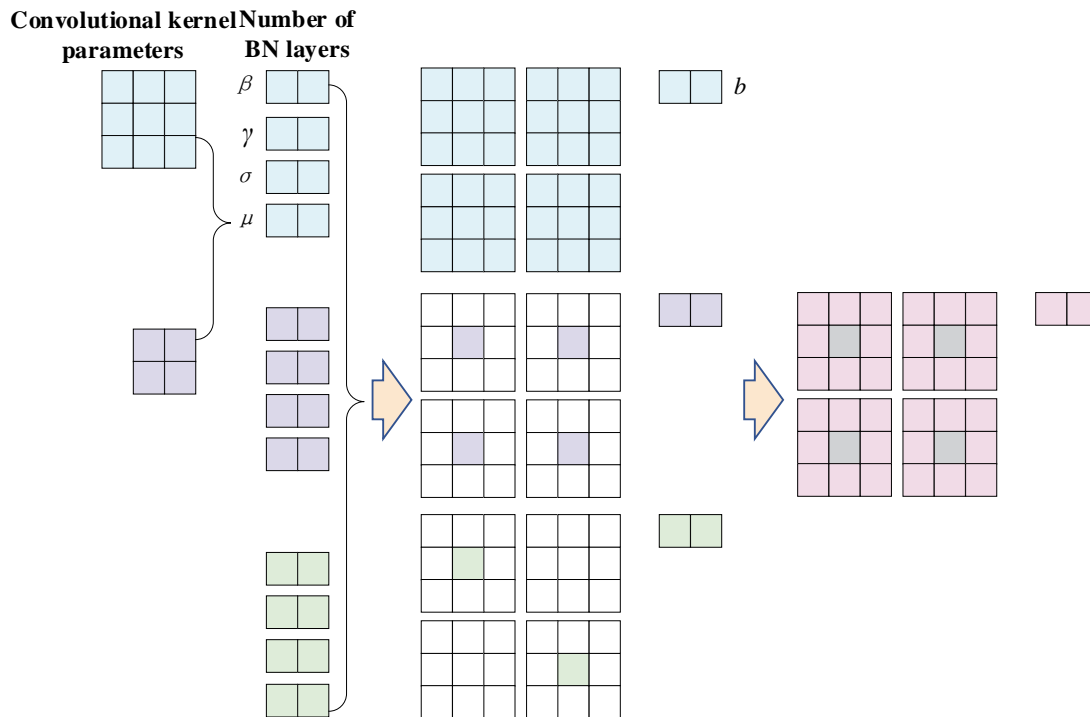
from the convolution plus BN layer into a convolution structure with bias. Let  $X \in \mathbb{R}^{H \times W \times C}$  be the input tensor. The computation for the BN layer is given by,

$$BN(X) = \gamma \frac{X - \mu}{\sigma} + \beta \quad (1)$$

where  $\mu$  represents the mean of the samples, and  $\sigma$  represents the variance of the samples.  $\gamma$  and  $\beta$  are learnable parameters, corresponding to the scaling and shifting factors, respectively. The computation for the convolution without bias is given by,

$$Conv(X) = X * W \quad (2)$$

where  $W$  is the weight matrix, which is used to perform a weighted sum on the input signals.



**Fig. 8.** Structural reparameterization of RepConv

The input tensor  $X$ , after passing through the convolutional layer and BN layer, can be expressed as:

$$BN(Conv(X)) = \gamma \frac{X * W - \mu}{\sigma} + \beta \quad (3)$$

That is,

$$BN(Conv(X)) = X * \left( \frac{\gamma}{\sigma} W \right) \gamma - \frac{\gamma \mu}{\sigma} + \beta \quad (4)$$

Let  $W_{fused} = \frac{\gamma}{\sigma} W$ ,  $b_{fused} = -\frac{\gamma \mu}{\sigma}$ ,  $W_{fused}$  and  $b_{fused}$  represent the fused convolution kernel weights and bias terms, respectively. The final result of the convolution and BN layer fusion is expressed as:

$$BN(Conv(X)) = X * W_{fused} + b_{fused} \quad (5)$$

By the fusion method described above, the  $3 \times 3$  convolutional layers and BN layers in Step (1) can be merged, reducing the number of parameters in the network.

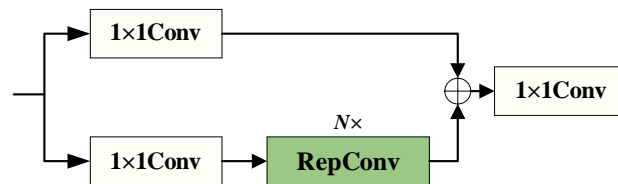
(3) Fusion of the convolution layers with their respective biases: the three sets of  $3 \times 3$  convolutional kernels and their corresponding biases are stacked together, resulting in a single  $3 \times 3$  convolutional kernel and bias. Let the convolution kernel parameters and biases of the three sets of  $3 \times 3$  convolutions be  $W_1, W_2, W_3$  and  $b_1, b_2, b_3$ , respectively. The output tensor  $Y \in \mathbf{R}^{H \times W \times C}$  after processing the input tensor  $X \in \mathbf{R}^{H \times W \times C}$  through the three sets of  $3 \times 3$  convolutions and their corresponding biases can be expressed as:

$$Y = (X * W_1 + b_1) + (X * W_2 + b_2) + (X * W_3 + b_3) \quad (6)$$

That is,

$$Y = X * (W_1 + W_2 + W_3) + (b_1 + b_2 + b_3) \quad (7)$$

Based on the above calculations, the RepConv multi-branch structure from the training phase is transformed into a single-branch convolutional structure through structural reparameterization. Subsequently, RepConv is incorporated into the RepBlock module, as shown in Fig. 9.



**Fig. 9.** Structure of RepBlock

The RepBlock module combines RepConv with a dual-path architecture. It consists of two parallel  $1 \times 1$  convolution layers: one directly passes the original information, while the other adjusts the channel size. This is followed by a module composed of multiple RepConv layers for in-depth feature extraction. Through the RepConv layers, RepBlock is capable of capturing complex features within the image with greater detail, which is crucial for improving detection accuracy. Additionally, the parallel residual connections help retain the original features, mitigating the vanishing gradient problem, and enhancing the stability and reliability of the model.

#### *Addition of small-object detection head*

In the dataset used in this study, small-object targets occupy a very small portion of the image. After setting the image size to  $640 \times 640$ , many targets are smaller than  $3 \times 3$  pixels. After multiple downsampling pooling operations, most of the features are lost, resulting in a high likelihood of false negatives. The detection heads in the baseline model have sizes of  $20 \times 20$ ,  $40 \times 40$ , and  $80 \times 80$ , and when using the smallest detection head ( $80 \times 80$ ) to detect each grid in the image, the receptive field is only  $8 \times 8$ . This limits the model's ability to recognize small targets.

To address this, a small-object detection head with a size of  $160 \times 160$  is added to the Head layer of the baseline model, improving the model's detection capability for small targets. The structure of the new detection head is shown in Fig. 2. First, the  $80 \times 80$  feature map from the second layer of the backbone network is stacked with the upsampled feature maps from the Neck layer. After passing through RepBlock and DySample processing, additional feature layers with small-object characteristics are obtained. These are then

concatenated with the  $160 \times 160$  feature map output from the second layer of the backbone network, enhancing the  $160 \times 160$  scale feature layer's ability to represent small targets related to wood surface defects. The added detection head allows small-object feature information to be propagated through the detection layers along the downsampling path to other feature layers at different scales. This enables small-object features to be extracted at deeper network layers, enhancing the detection of wood surface defects in complex backgrounds and effectively reducing both false positives and false negatives at different scales.

## RESULTS AND DISCUSSION

### Ablation Experiments

The software and hardware configuration used in the experiment is detailed in Table 2. The specific training hyperparameters are as follows: (1) Input image size: 640 pixels. (2) Number of iterations: 200. (3) Batch size: 8. (4) Initial learning rate: 0.01. (5) Weight decay coefficient: 0.0005. (6) Momentum: 0.937. To assess the accuracy and effectiveness of our method, two performance evaluation metrics were employed: Average Precision (AP) and mean Average Precision (mAP).

**Table 2.** Configuration of Software and Hardware Used in the Experiment

Device name	Parameter
GPU	NVIDIA GeForce RTX 4060 Laptop 8G
CPU	13th Gen Intel(R) Core(TM) i9-13900HX 2.2GHz
Computer operating system	Windows 11
Development environment software	PyCharm 2023.2.1
Programming language	Python 3.8
Deep learning framework	PyTorch 2.1.0
Computational acceleration	CUDA11.0

To evaluate the impact of each module in the improved model, an ablation study was conducted with YOLOv8 as the baseline model. This study aimed to validate the effectiveness of the proposed enhancements. The results of the ablation experiments are presented in Table 3. The table presents the results of different configurations, where each variant isolates the effect of an individual module to evaluate its contribution to overall performance. YOLOv8+C2f-MMSA indicates that the MMSA module is seamlessly integrated into the C2f module, replacing the C2f module in the backbone of YOLOv8. YOLOv8+Dysample indicates the incorporation of Dysample into the neck section for upsampling. YOLOv8+RepBlock denotes the integration of the RepBlock module into the neck section. YOLOv8+P2 denotes the incorporation of the small-object detection head module into the Head section. "This work" denotes the current proposed improved model.

**Table 3.** Results of Ablation Experiment

Methods	mAP	AP						
		Live_Knot	Marrow	Resin	Dead_Knot	Knot_with_Crack	Knot_Missing	Crack
YOLOv8	0.729	0.787	0.823	0.667	0.877	0.422	0.804	0.725
YOLOv8+C2f-MMSA	0.750	0.759	0.858	0.768	0.875	0.464	0.803	0.723
YOLOv8+Dysample	0.747	0.783	0.778	0.720	0.870	0.523	0.776	0.776
YOLOv8+RepBlock	0.754	0.784	0.829	0.762	0.862	0.492	0.809	0.739
YOLOv8+P2	0.735	0.786	0.842	0.700	0.868	0.507	0.706	0.733
This work	0.795	0.887	0.891	0.790	0.893	0.610	0.750	0.743

**Table 4.** Comparison of Various Detection Models

Methods	mAP	AP						
		Live_Knot	Marrow	Resin	Dead_Knot	Knot_with_Crack	Knot_Missing	Crack
YOLOv5	0.746	0.764	0.786	0.774	0.868	0.476	0.837	0.716
YOLOv7	0.734	0.767	0.814	0.724	0.858	0.529	0.782	0.665
YOLOv8	0.729	0.787	0.823	0.667	0.877	0.422	0.804	0.725
YOLOv9	0.757	0.789	0.805	0.786	0.875	0.515	0.759	<b>0.769</b>
YOLOv10	0.672	0.725	0.799	0.671	0.851	0.378	0.690	0.590
YOLO11	0.745	0.768	0.805	0.768	0.879	0.488	0.752	0.757
YOLOv12	0.746	0.773	0.835	0.760	0.882	0.493	0.814	0.666
Wang <i>et al.</i> (2024)	0.777	0.814	0.958	0.852	0.868	0.481	0.825	<u>0.642</u>
Xi <i>et al.</i> (2024)	0.784	0.854	0.871	0.779	0.831	0.595	<b>0.866</b>	<u>0.693</u>
This work	<b>0.795</b>	<b>0.887</b>	<b>0.891</b>	<b>0.790</b>	<b>0.893</b>	<b>0.610</b>	0.750	0.743



Integrating the MMSA module into the C2f structure (YOLOv8+C2f-MMSA) improved mAP by 2.1% compared to the baseline, demonstrating its effectiveness in enhancing feature extraction for small defects. Similarly, the introduction of Dysample (YOLOv8+Dysample) led to an 1.8% improvement, indicating that the learnable dynamic upsampling strategy mitigates information loss during feature scaling. The addition of RepBlock (YOLOv8+RepBlock) further enhanced performance by 2.5%, suggesting its ability to capture multi-scale defect features effectively. Finally, incorporating the small-object detection head (YOLOv8+P2) provided an additional boost of 0.6%, validating its role in refining small defect detection. By integrating all proposed enhancements, the final model (Ours) achieved the highest performance, surpassing the baseline by 6.6% in mAP. This demonstrates that the combined improvements contribute synergistically to the accuracy and robustness of wood surface defect detection.

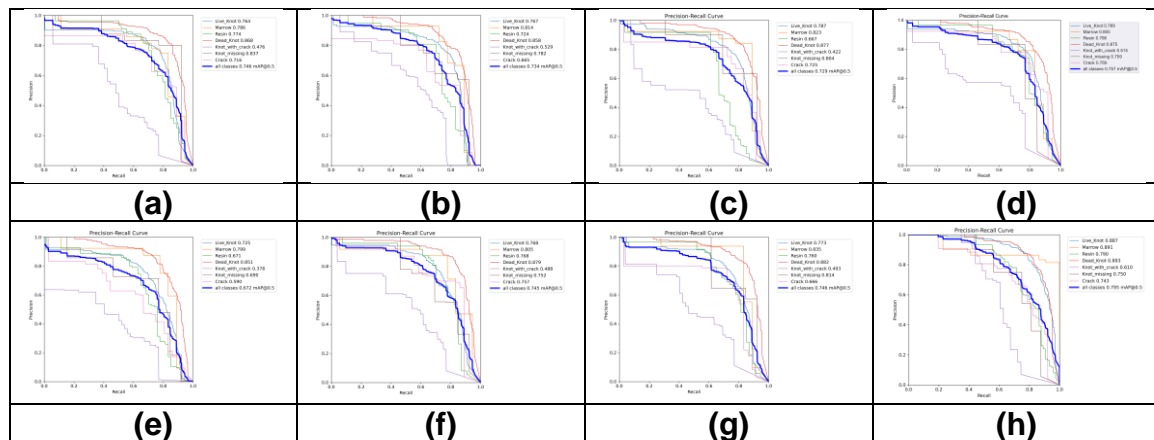
### Comparison Experiments with Benchmark Models

To further validate the effectiveness of the proposed model, a comparative experiment was conducted against several state-of-the-art wood surface defect detection models. The benchmark models selected for comparison include YOLOv5, YOLOv7, YOLOv9, YOLOv10, YOLO11, YOLOv12, and YOLOv8 (baseline). All models were trained and evaluated under identical conditions using the dataset mentioned in section 2.1 to ensure fairness.

The results, as shown in Table 4, indicate that the proposed model achieved the highest mAP, outperforming other models in detecting wood surface defects. Specifically, the present method improved the mAP by 6.6% over the YOLOv8 baseline, demonstrating the effectiveness of the introduced C2f-MMSA module, Dysample upsampling strategy, RepBlock, and small-object detection head in enhancing small defect recognition. In the AP results, the model of Xi *et al.* (2024) achieved the highest AP of 0.866 for the Knot\_Missing defect, while YOLOv9 attained the best AP of 0.769 for the Crack defect. For the remaining defect types, including Live\_Knot, Marrow, Resin, Dead\_Knot, and Knot\_with\_Crack, the proposed model consistently outperformed the other benchmark models, achieving the highest AP values across these categories. These results demonstrate the effectiveness of the proposed improvements in enhancing the detection performance for various wood surface defects.

The Precision-Recall (P-R) curve provides an intuitive visualization of the Average Precision (AP) values. It represents the trade-off between precision and recall, with a larger area under the curve indicating superior model performance. When the area reaches 1, it signifies that the model has perfectly detected all targets. Figure 10 illustrates the AP values for the seven types of wood defects evaluated in this study. Subfigures (a), (b), (c), (d), (e), (f), (g) and (h) correspond to the AP values obtained by YOLOv5, YOLOv7, YOLOv8, YOLOv9, YOLOv10, YOLO11, YOLOv12, and the proposed model, respectively.

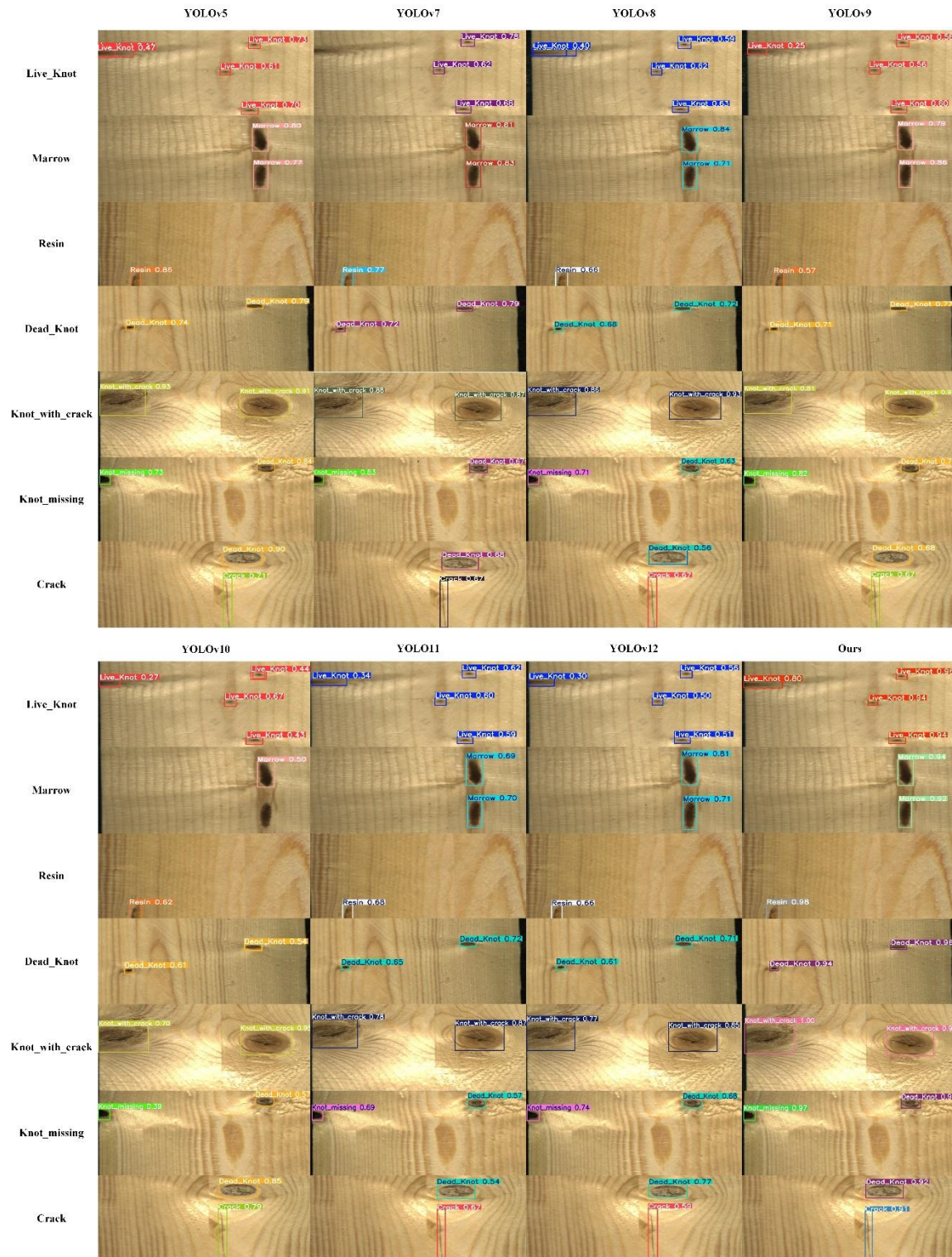
In addition, to validate the performance of the current model in complex backgrounds, the present results were compared with other YOLO-based studies addressing similarly challenging detection environments. For instance, tile defect detection in historical buildings Karimi *et al.* (2024) report overall accuracy of over 72%. For wood surface defect detection (Wang *et al.* 2024; Xi *et al.* 2024), researchers report mAP values of 77.7% and 78.4%, respectively. The present model achieved a comparable or higher detection accuracy (79.5% mAP), particularly under diverse wood textures and irregular defect patterns. These results highlight the robustness and generalization capacity of the present approach under real-world complexity.



**Fig. 10.** Precision–recall (P–R) curves

The proposed model achieved higher AP values than other benchmark models for almost all seven types of wood defects, except for Knot\_Missing and Crack. Despite the overall improvements achieved by the proposed model, the detection performance for Knot\_Missing and Crack defects did not surpass that of YOLOv5 and YOLOv9, respectively. This can be attributed to the following factors: (1) Knot\_Missing defects typically exhibit clear edges and relatively large missing regions, making them more distinguishable. YOLOv5, as a well-established model, may have been optimized for such easily identifiable defects, resulting in superior performance. (2) Crack defects, in contrast, are characterized by irregular, thin, and elongated structures, which can resemble natural wood grain patterns. The strong performance of YOLOv9 in this category suggests that its feature extraction and detection heads are more suited for capturing fine-grained and linear features. (3) The distribution of Knot\_Missing and Crack samples in the training dataset may impact the model's generalization capability. If these defect types are underrepresented or exhibit high variability, the model may struggle to learn a robust representation for them. (4) Knot\_Missing defects, being relatively large and distinct, may not benefit as significantly from the added feature extraction enhancements, as their characteristics are already well captured by standard detection modules.

The visual comparison results are illustrated in Fig. 11. Each detection box is associated with a confidence score, which quantifies the model's certainty regarding its detection outcome. This score ranges from 0 to 1, where higher values indicate greater confidence in the detection, whereas lower values suggest increased uncertainty in the model's predictions. The experimental results highlight the superior performance of the proposed model in wood defect detection. For instance, in detecting the Marrow defect, the confidence scores achieved by the proposed model were 0.94 and 0.92, compared to only 0.84 and 0.72 for the YOLOv8, reflecting improvements of 0.10 and 0.20, respectively. In detecting the Resin defect, the confidence score is the highest at 0.95. Furthermore, the proposed model exhibited no misclassifications or missed detections across all defect types, demonstrating its high reliability and accuracy in wood defect detection.

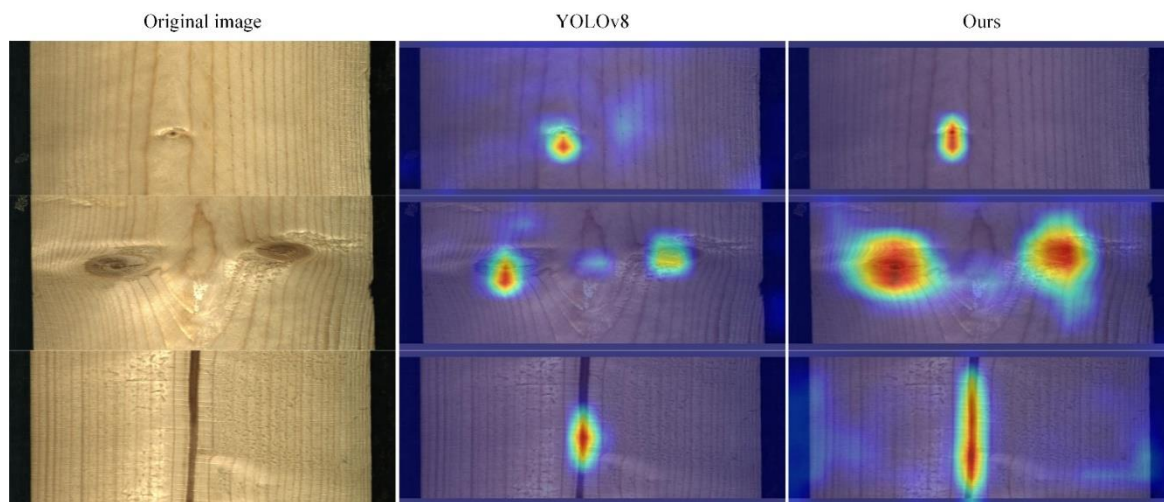


**Fig. 11.** Examples of visual comparison results

Figure 12 presents the visual results obtained using Grad-CAM for YOLOv8 and the proposed model in wood defect detection tasks. Grad-CAM is a visualization technique designed to enhance model interpretability by leveraging gradient information to generate heatmaps that highlight the most influential regions of the input image for a given



prediction. This approach provides an intuitive way to illustrate the model's attention distribution and explain its decision-making process. The proposed model exhibits a more precise focus on key defect areas during wood defect detection. The generated heatmaps display deeper color intensities, indicating a stronger response to the target regions, with attention concentrated on the critical defect features. In contrast, YOLOv8's heatmap presents a more dispersed color distribution, suggesting that its attention is spread across a broader area, potentially including irrelevant regions. This distinction highlights the superior feature extraction and defect localization capabilities of the proposed model, enabling more accurate identification of wood defects and ultimately enhancing detection accuracy and reliability.



**Fig. 12.** Grad-CAM comparison of wood defects

## CONCLUSIONS

The detection of wood surface defects is crucial for ensuring the quality and performance of wood products. To address the challenges posed by complex and variable backgrounds, as well as the small size of certain defects, this paper proposes an improved YOLOv8-based detection model that enhances accuracy while minimizing false positives and missed detections. Specifically, the MMSA module is integrated into the C2f structure in the Backbone to improve the model's ability to capture contextual and background information for small targets. In the Neck, the DySample module mitigates fine-grained detail loss during feature fusion, and the RepBlock module strengthens the extraction of multi-scale defect features. Additionally, a small-object detection head improves detection accuracy for small defects.

1. Ablation experiments demonstrate that the integration of different enhanced modules leads to statistically significant improvements in the detection accuracy of the baseline model, with varying degrees of enhancement observed across the modules (*e.g.*, +2.1% mAP with C2f-MMSA, +1.8% mAP with DySample, +2.5% mAP with RepBlock).
2. Comparison experiments further reveal that, compared to the baseline model, the proposed method achieves improved AP values across all defect categories except

Knot\_Missing. Similarly, when compared to YOLO variants (v5, v7, v9, v10, 11, and v12), the proposed model generally outperforms them in most defect categories. However, for Crack, YOLOv9 attains the highest AP value, while for Knot\_Missing, YOLOv5 performs best.

3. Visualization results further confirm that the proposed method effectively reduces both missed and incorrect detections, ensuring more accurate and reliable defect identification under complex wood surface conditions.

In summary, the proposed method effectively addresses the challenges of detecting small-target features in complex backgrounds for wood surface defect detection. It enhances detection accuracy and reliability, fulfilling the practical requirements of wood surface defect detection. Future work will explore adaptive training strategies and lightweight deployment frameworks to further improve performance in different scenarios of wood production.

## ACKNOWLEDGMENTS

This research was financially supported by the Science and Technology Planning Project of Guangxi Province, China (No. 2022AC21012) and the Industry-University-Research Innovation Fund Projects of China University in 2021 (No. 2021ITA10018).

## Author Contributions

W.D.: Writing—original draft, Investigation, Software, Methodology. J Y: Conceptualization, Writing—review and editing, Supervision, Data curation. All authors have read and agreed to the published version of the manuscript.

## Data Availability Statement

Data will be made available on request.

## Conflicts of Interest

The authors declare no conflicts of interest.

## REFERENCES CITED

- Chen, W., Liu, J., Fang, Y., and Zhao, J. (2023a). “Timber knot detector with low false-positive results by integrating an overlapping bounding box filter with faster R-CNN algorithm,” *BioResources* 18(3), 4964-4976. DOI: 10.15376/biores.18.3.4964-4976
- Chen, Y., Sun, C., Ren, Z., and Na, B. (2023b). “Review of the current state of application of wood defect recognition technology,” *BioResources* 18(1), 2288-2302. DOI: 10.15376/biores.18.1.2288-2302
- Connors, R. W., Mcmillin, C. W., Lin, K., and Ramon, E. (1983). “Identifying and locating surface defects in wood: Part of an automated lumber processing system,” *IEEE Transactions on Pattern Analysis and Machine Intelligence* 6, 573-583. DOI: 10.1109/tpami.1983.4767446
- Ding, F., Zhuang, Z., Liu, Y., Jiang, D., Yan, X., and Wang, Z. (2020). “Detecting defects on solid wood panels based on an improved SSD algorithm,” *Sensors* 20,

5315. DOI: 10.3390/s20185315
- Fan J., Liu, Y., Hu Z., H, Zhao, Qian., Shen, L., and Zhou, X. (2019). "Solid wood panel defect detection and recognition system based on Faster R-CNN," *Journal of Forestry Engineering* 4(03), 112-117. DOI: 10.13360/j.issn.2096-1359.2019.03.017
- Gao, M., Qi, D., Mu, H., and Chen, J. (2021). "A transfer residual neural network based on ResNet-34 for detection of wood knot defects," *Forests* 12(2), article 212. DOI: 10.3390/f12020212
- Girshick, R., Donahue, J., Darrell, T., and Malik, J. (2014). "Rich feature hierarchies for accurate object detection and semantic segmentation," in: *2014 IEEE Conference on Computer Vision and Pattern Recognition*, Columbus, OH, USA, pp. 580-587. DOI: 10.1109/CVPR.2014.81
- He, K., Gkioxari, G., Dollár, P., and Girshick, R. (2017). "Mask R-CNN," in: *2017 IEEE International Conference on Computer Vision (ICCV)*, Venice, Italy, pp. 2980-2988. DOI: 10.1109/ICCV.2017.322
- Hittawe, M. M., Sidibé, D., Beya, O., and Mériaudeau, F. (2017). "Machine vision for timber grading singularities detection and applications," *Journal of Electronic Imaging* 26(6), article 063015. DOI: 10.1117/1.JEI.26.6.063015
- Hu, K., Wang, B., Shen, Y., Guan, J., and Cai, Y. (2020). "Defect identification method for poplar veneer based on progressive growing generated adversarial network and MASK R-CNN model," *BioResources* 15(2), 3041-3052. DOI: 10.15376/biores.15.2.3041-3052
- Ji, M., Zhang, W., Han, J., Miao, H., Diao, X., and Wang, G. (2024a). "A deep learning-based algorithm for online detection of small target defects in large-size sawn timber," *Industrial Crops and Products* 222, article 119671. DOI: 10.1016/j.indcrop.2024.119671
- Ji, M., Zhang, W., Wang, G. F., Diao, X. L., Miao, H., and Gao, R. (2024b). "Machine vision for knot detection and location in Chinese fir lumber," *Forest Products Journal* 74(2), 185-202. DOI: 10.13073/FPJ-D-23-00050
- Jiang, X., Wang, J., Zhang, Y., and Jiang, S. (2024). "Defect detection in solid timber panels using air-coupled ultrasonic imaging techniques," *Appl. Sci.* 14, 434. DOI: 10.3390/app14010434
- Karimi, N., Mishra, M., and Lourenço, P. B. (2024). "Deep learning-based automated tile defect detection system for Portuguese cultural heritage buildings," *Journal of Cultural Heritage*, 68, 86-98. DOI: 10.1016/j.culher.2024.05.009
- Kodytek, P., Bodzas, A., and Bilik, P. (2021). "A large-scale image dataset of wood surface defects for automated vision-based quality control processes," *F1000Research* 10, article 581. DOI: 10.12688/f1000research.52903.2
- Li, D., Xie, W., Wang, B., Zhong, W., and Wang, H. (2021). "Data augmentation and layered deformable Mask R-CNN-based detection of wood defects," *IEEE Access* 9, 108162-108174. DOI: 10.1109/ACCESS.2021.3101247
- Li, H., Zhang, J., Xu, G., and Li, W. (2024). "Compression test and size effect study of defective wood," *Wood Material Science & Engineering* 1-9. DOI: 10.1080/17480272.2024.2383747
- Liu, W., Anguelov, D., Erhan, D., Szegedy, C., Reed, S., Fu, C. Y., and Berg, A. C. (2016). "SSD: Single shot multibox detector," in: *Computer Vision – ECCV 2016*, Amsterdam, pp. 21-37. DOI: 10.1007/978-3-319-46448-0\_2
- Liu, W., Lu, H., Fu, H., and Cao, Z. (2023). "Learning to upsample by learning to sample," In: *2023 IEEE/CVF International Conference on Computer Vision (ICCV)*,



- Paris, France, pp. 6004-6014. DOI: 10.1109/ICCV51070.2023.00554
- López, G., Basterra, L. A., Ramón-Cueto, G., and Diego, A. de. (2014). "Detection of singularities and subsurface defects in wood by infrared thermography," *International Journal of Architectural Heritage* 8(4), 517-536. DOI: 10.1080/15583058.2012.702369
- Meng, W., and Yuan, Y. (2023). "SGN-YOLO: Detecting wood defects with improved YOLOv5 based on semi-global network," *Sensors* 23, 8705. DOI: 10.3390/s23218705
- Mishra, M., and Lourenço, P. B. (2024). "Artificial intelligence-assisted visual inspection for cultural heritage: State-of-the-art review," *Journal of Cultural Heritage*, 66, 536-550. DOI: 10.1016/j.culher.2024.01.005
- Özcan, U., Kiliç, K., Kiliç, K., and Doğru, İ. A. (2024). "Using deep learning techniques for anomaly detection of wood surface," *Drvena Industrija* 75(3), 275-286. DOI: 10.5552/drvind.2024.0114
- Redmon, J., Divvala, S., Girshick, R., and Farhadi, A. (2016). "You Only Look Once: Unified, real-time object detection," in: *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Las Vegas, USA, pp. 779-788. DOI: 10.1109/CVPR.2016.91
- Ren, S., He, K., Girshick, R., and Sun, J. (2017). "Faster R-CNN: Towards real-time object detection with region proposal networks," *IEEE Transactions on Pattern Analysis and Machine Intelligence* 39(6), 1137-1149. DOI: 10.1109/TPAMI.2016.2577031
- Sandak, J., Sandak, A., Zitek, A., Hintestoisser, B., and Picchi, G. (2020). "Development of low-cost portable spectrometers for detection of wood defects," *Sensors* 20(2), article 545. DOI: 10.3390/s20020545
- Stängle, S.M., Brüchert, F., Heikkilä, A., Usenius, T., Usenius, A., and Sauter, U. H. (2015). "Potentially increased sawmill yield from hardwoods using X-ray computed tomography for knot detection," *Annals of Forest Science* 72, 57-65. DOI: 10.1007/s13595-014-0385-1
- Sun, P. A. (2022). "Wood quality defect detection based on deep learning and multicriteria framework," *Mathematical Problems in Engineering* 1-9. DOI: 10.1155/2022/4878090
- Su, Q., Mu, J., Liang, W., Wang, X., and Li, J. (2025). "Multi-head hybrid self-attention mechanism for object detection," *Laser & Optoelectronics Progress* 62(06), 0637006. DOI: 10.3788/LOP241509
- 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)*. 1-6. DOI: 10.1109/ADICS58448.2024.10533619
- Wang, J., Chen, K., Xu, R., Liu, Z., Loy, C. C., and Lin, D. (2019). "CARAFE: Content-aware reassembly of features," in: *2019 IEEE/CVF International Conference on Computer Vision (ICCV)*, Seoul, Korea, pp. 3007-3016. DOI: 10.1109/ICCV.2019.00310
- Wang, Q., Wu, B., Zhu, P., Li, P., Zuo, W., and Hu, Q. (2020). "ECA-Net: Efficient channel attention for deep convolutional neural networks," in: *2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, Seattle, WA, USA, pp. 11531-11539. DOI: 10.1109/CVPR42600.2020.01155
- Wang, R., Liang, F., Wang, B., and Mou, X. (2023). "ODCA-YOLO: An omni-dynamic

- convolution coordinate attention-based YOLO for wood defect detection,” *Forests* 14, article 1885. DOI: 10.3390/f14091885
- Wang, J., Zhang, C., Zhao, M., Zou, H., Qi, L., and Wang, Z. (2024). “A composite pulse excitation technique for air-coupled ultrasonic detection of defects in wood,” *Sensors* 24, article 7550. DOI: 10.3390/s24237550
- Wang, R., Liang, F., Wang, B., Zhang, G., Chen, Y., and Mou, X. (2024), “An efficient and accurate surface defect detection method for wood based on improved YOLOv8,” *Forests* 15(7), article 1176. DOI: 10.3390/f15071176
- Wyckhuysse, A., and Maldague, X. (2001). “A study of wood inspection by infrared thermography, part I: Wood pole inspection by infrared thermography,” *Research in Nondestructive Evaluation* 13, 1-12. DOI: 10.1007/s00164-001-0005-y
- Xia, B., Luo, H., and Shi, S. (2022). “Improved faster R-CNN based surface defect detection algorithm for plates,” *Computational Intelligence and Neuroscience* 1-11. DOI: 10.1155/2022/3248722
- Xi, H., Wang, R., Liang, F., Chen, Y., Zhang, G., and Wang, B. (2024). “SiM-YOLO: A wood surface defect detection method based on the improved YOLOv8,” *Coatings* 14(8), article 1001. DOI: 10.3390/coatings14081001
- Yi, L. P., Akbar, M. F., Wahab, M. N. A., Rosdi, B. A., Fauthan, M. A., and Shrifan, N. H. M. M. (2024). “The prospect of artificial intelligence-based wood surface inspection: A review,” *IEEE Access* 12, 84706-84725. DOI: 10.1109/ACCESS.2024.3412928
- Yu, H., Liang, Y., Liang, H., and Zhang, Y. (2019). “Recognition of wood surface defects with near infrared spectroscopy and machine vision,” *Journal of Forestry Research* 30, 2379-2386. DOI: 10.1007/s11676-018-00874-w
- Zhang, Y. G. (2017). *Restoration and Segmentation of Wood X-ray Images Based on Optimization Algorithm*, Master’s Thesis, Northeastern University, Shenyang, China.
- Zhu, Y., Xu, Z., Lin, Y., Chen, D. Ai, Z., and Zhang, H. (2024). “A multi-source data fusion network for wood surface broken defect segmentation,” *Sensors* 24, article 1635. DOI: 10.3390/s24051635
- Zou, X., Wu, C., Liu, H., Yu, Z., and Kuang, X. (2024). “An accurate object detection of wood defects using an improved Faster R-CNN model,” *Wood Material Science & Engineering* 20(2), 413-419. DOI: 10.1080/17480272.2024.2352605

Article submitted: April 10, 2025; Peer review completed: April 27, 2025; Revisions accepted: May 22, 2025; Published: May 23, 2025.

DOI: 10.15376/biores.20.3.5709-5730