

Research on Wood Defect Detection Algorithms Based on Improved Neural

Yuan Zhang^{1, a}, Xuewen Ding^{1, 2, *}, Xiaokai Jiang¹, Shaosai Wang¹

¹ School of Electronic Engineering, Tianjin University of Technology and Education, Tianjin 300222, China

² Tianjin Yunzhitong Technology Co., Ltd., Tianjin 300350, China

* Corresponding author: Xuewen Ding (Email: dingxw@tute.edu.cn), ^azy_514@163.com

Abstract: To address the challenges of complex texture interference and the difficulty of detecting small-scale defects on wood surfaces, this study proposes an improved neural network-based detection algorithm, YOLOv11-RABRA. First, a Single Class Improve (SI) strategy is employed to selectively augment minority defect categories, thereby alleviating class imbalance and enhancing the model's learning capability for rare defects. On this basis, a C3K2_RepVGG module is designed by integrating RepVGG re-parameterization, which optimizes the feature extraction and processing pipeline and strengthens defect feature representation. Furthermore, an ADown adaptive lightweight down-sampling mechanism is introduced to effectively preserve fine-grained features such as cracks, achieving a balance between accuracy and computational efficiency. Finally, a C2BRA dual-route attention module is applied to enhance defect-focused feature aggregation while suppressing interference from complex wood textures. Experimental results demonstrate that, compared with the baseline YOLOv11 model, the proposed algorithm achieves improvements of 13.7 percentage points in mAP50 and 12.8 percentage points in precision, significantly enhancing the accuracy and robustness of wood surface defect detection. In summary, the proposed method achieves advances in both detection performance and lightweight design, providing a feasible solution for high-precision detection in intelligent manufacturing and industrial quality inspection.

Keywords: Wood Surface Defect Detection; Neural Network Improvement; SI; C3K2_RepVGG; Adaptive Downsampling (ADown); C2BRA.

1. Introduction

Wood is an important industrial material, plays a crucial role in various industries in China, including construction, decoration, furniture, and office applications. Due to significant differences in the requirements for wood's appearance, deformation, surface roughness, and resistance to external forces across different industries, researching wood defect detection technologies is of great practical significance. Currently, methods for detecting wood defects mainly include manual inspection, mechanical detection[1], modern technological detection, and machine vision detection[2]. Among these, traditional manual inspection methods are inefficient and highly subjective, making it difficult to meet the demands of modern industrial production. Mechanical detection often involves contact operations that require frequent adjustments of the object's position, resulting in relatively low detection efficiency. In contrast, modern technological detection methods such as laser, near-infrared spectroscopy[3], ultrasound, microwave, and X-ray[4] can significantly enhance production efficiency and automation levels; however, their widespread application in large-scale industrial settings is limited by high costs.

In the field of machine vision detection, traditional methods typically preprocess images of wood defects and select candidate regions using sliding window techniques. This technique involves sliding windows of various sizes and aspect ratios across the image with a fixed step size to extract features such as color, geometry, and texture; these features are then input into a classifier to determine the type of defect. Li et al.[5] proposed a detection method based on Linear Discriminant Analysis (LDA [6]) and compressed sensing images for rapid and accurate identification of wood defects.

Yuce et al. [7] utilized Principal Component Analysis to reduce the number of key variables and employed Artificial Neural Networks (ANN [8]) for real-time identification of veneer defects in wood manufacturing companies' quality control processes. However, the complex and variable texture backgrounds of wood surfaces along with diverse defect types pose numerous challenges for traditional machine vision detection methods when dealing with complicated scenes; they struggle to effectively address defects of varying sizes and shapes. With advancements in computational power and algorithm development, deep learning has been widely applied in the field of machine vision. Currently, deep learning object detection algorithms are primarily categorized into two-stage detection algorithms[9] and single-stage detection algorithms [10]. The detection process of two-stage algorithms includes region proposal generation, feature extraction, classification, and bounding box regression to achieve precise object detection. Xu Shanshan et al. [11] extracted wood defect features based on an incremental convolutional neural network structure and employed an incremental parameter training strategy to enhance the robustness of the algorithm. Xia et al.[12] combined ResNet50[13] with shape-variable convolutional networks and Feature Pyramid Networks (FPN[14]) to improve multi-scale defect feature representation capability. Li et al.[15], addressing irregular wood defects, utilized Generative Adversarial Networks to augment datasets and construct a Mask R-CNN [16] detection model. Single-stage detection algorithms perform object classification and bounding box regression directly on images without generating region proposals; this results in faster detection speeds and lower computational costs, making them particularly suitable for real-time detection tasks. Hu Zhongkang [17] designed a solid

wood board defect detection algorithm based on SSD[18], enhancing performance through the introduction of Feature Pyramid Networks (FPN) and ResNet. Jia Haonan et al. [19] achieved rapid recognition of multiple types of wood defects using an improved YOLOv5[20] model; Han et al. [21] developed an STC-YOLOv5 model specifically for small object detection issues; Zhang Zilei[22] improved the YOLOX model by incorporating Adaptive Spatial Feature Fusion (ASSF) [23] and Focal Loss [24], thereby enhancing detection performance.

Despite significant advancements in deep learning technologies for wood defect detection, existing algorithms still face limitations due to factors such as complex texture backgrounds, diversity in defect types, and some defects being small in size. Therefore, this study addresses key issues in wood defect detection by proposing a more targeted deep learning object detection algorithm aimed at improving adaptability and reliability in complex scenarios while providing more effective technical support for quality control in the wood industry. Additionally, this research is innovative in several ways: it enhances multi-scale defect feature representation capabilities through improved network structures; combines data augmentation with strategies for optimizing imbalanced samples to improve the detection performance for small sample defects; while ensuring accuracy it also considers speed to achieve an efficient yet robust method for detecting wood surface defects.

2. Improvements to the YOLOv11 Network

2.1. Overview of the YOLOv11 Network

Compared to previous YOLO series models, YOLOv11 has

achieved significant improvements in both detection accuracy and inference speed. This improvement is primarily attributed to systematic optimizations in its network architecture. The overall architecture still consists of a Backbone, Neck, and Head, but innovative designs have been introduced at key stages such as feature extraction, information fusion, and task separation, balancing detection efficiency and robustness.

In the Backbone, YOLOv11 replaces traditional bottleneck layers with C3k2 modules to effectively enhance feature extraction capabilities. Additionally, it introduces a Cross Stage Partial Squeeze-and-Attention (C2PSA) module that focuses on key information and suppresses redundant noise through pyramid decomposition and weighting mechanisms to improve detection accuracy and robustness. Combined with Spatial Pyramid Pooling Fast (SPPF), this further strengthens the model’s ability to express multi-scale features. The Neck section continues the C3k2 design while optimizing the feature fusion strategy for efficient integration of cross-layer features, thereby enhancing adaptability to targets of different sizes. The Head section employs a decoupled structure that separates classification and regression tasks: the regression head uses conventional convolution for bounding box prediction, while the classification head introduces Depthwise Separable Convolution (DWConv)[25], significantly reducing parameter size and computational complexity while maintaining recognition accuracy and enhancing overall inference efficiency.

In summary, YOLOv11 implements targeted improvements at multiple critical stages of its network structure, achieving a coordinated optimization of detection accuracy, speed, and robustness. Its approach not only possesses application potential in general object detection tasks but also provides new ideas and technical support for real-time detection research in complex scenarios.

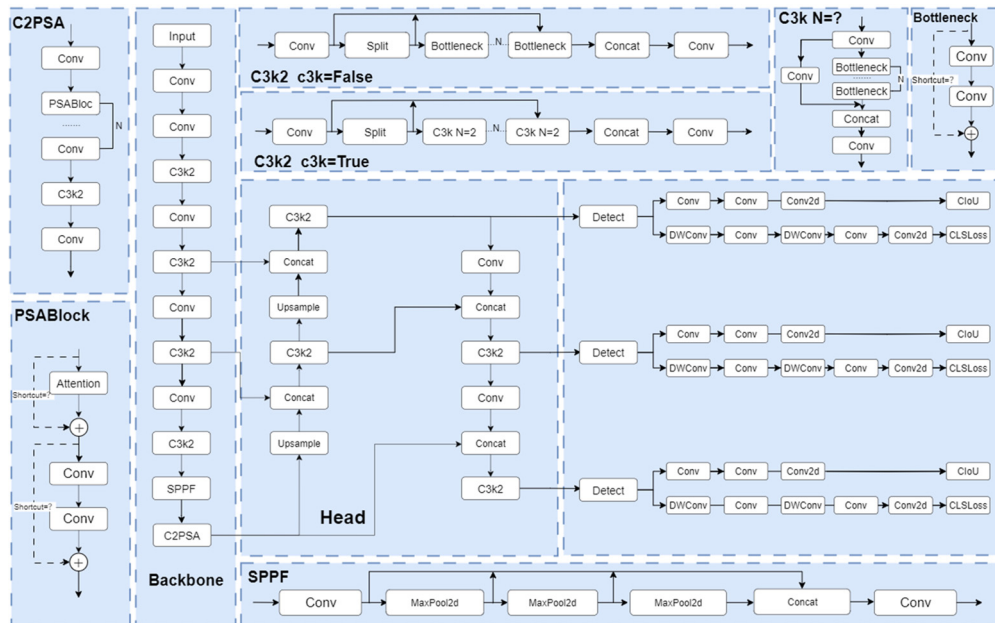


Figure 1. YOLOv11 network structure

2.2. Yolov11-RABRA Algorithm Network

In wood surface defect detection, although YOLOv11 achieves a good balance between detection accuracy and inference speed, it still has certain limitations under complex texture backgrounds, particularly in recognizing millimeter-

scale cracks, small knots, and multi-scale defects, leading to potential loss of detail features and false detections.

To address these issues, this paper proposes an improved algorithm called YOLOv11-RABRA. In the Backbone section, the original C3k2 is replaced with a C3k2_RepVGG module that integrates RepVGG structures to enhance feature

extraction capabilities through a multi-branch structure during training while balancing speed and accuracy through structural reparameterization during inference; an ADown method is introduced to prevent loss of fine crack features caused by traditional convolution compression; simultaneously, the C2BRA module replaces C2PSA to optimize the attention coordination mechanism that highlights defect areas while suppressing background interference effectively reducing missed detections and false positives. In the Head section, the feature fusion layer is changed from C3k2 to C3k2_RepVGG to maintain consistent feature processing logic while efficiently aggregating small,

medium, and large-scale features during multi-scale fusion for accurate identification of defects such as cracks, knots, and extensive decay. The overall structure of YOLOv11-RABRA is illustrated in Figure 2.

Through these improvements, YOLOv11-RABRA achieves simultaneous enhancements in detection accuracy and inference efficiency in wood defect detection, particularly demonstrating greater adaptability and robustness in recognizing fine cracks and multi-scale defects. This model embodies both theoretical innovation and engineering application value, providing a feasible solution for efficient intelligent detection in the wood industry.

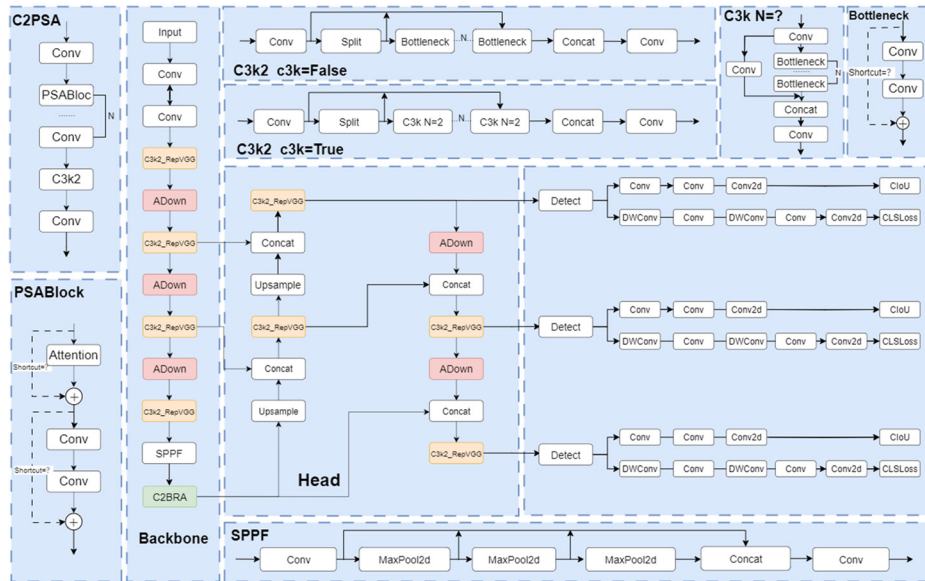


Figure 2. YOLOv11-RABRA Network Structure

3. Algorithm Improvements

3.1. Feature Fusion Module Based on Structural Reparameterization and Dual Feature Fusion Custom Convolution (C3K2_RepVGG)

The feature extraction capability of the C3K2 module in YOLOv11 is limited, and the efficiency of multi-scale feature fusion is low, resulting in uneven detection accuracy for defects of varying sizes and shapes, making it difficult to balance complex texture feature learning with inference speed. To address this, this paper combines the reparameterization advantages of RepVGG with feature fusion optimization concepts to design the C3K2_RepVGG module for performance breakthroughs.

As a lightweight neural network, RepVGG achieves a synergistic optimization of efficiency and performance through the separation of training and inference. During the training phase, it borrows from the residual concepts of ResNet to construct a parallel structure consisting of a 3×3 dense convolution branch, a 1×1 local branch, and an identity mapping branch to enhance multi-level semantic representation, as shown in Figure 3. In the inference phase, structural reparameterization techniques are employed to enhance efficiency through two conversion steps: first, mathematically merging Conv2d and BN layers into a single 3×3 convolution layer while constructing an equivalent identity mapping convolution layer for pure BN branches to dynamically calibrate distributions; subsequently,

reparameterization aggregates multiple branches into a single-path 3×3 convolution layer that normalizes non-standard convolution kernels through padding and stride adjustments while introducing compensation mechanisms to ensure accuracy, as shown in Figure 4. The basic unit employs a 3×3 convolution + ReLU structure that adapts to GPU parallel computing characteristics while balancing speed and efficiency.

Building upon this foundation, this paper further proposes the C3K2_RepVGG module, enhancing detection performance through three optimizations: first, integrating the feature separation and recursive fusion mechanism of the C2f module to reduce feature flow redundancy; second, introducing the C3k module to support customizable kernel sizes, thereby enhancing adaptability for multi-scale defects; third, retaining the Bottleneck structure to improve feature hierarchy, complementing RepVGG’s global feature extraction. During training, rich feature representations are learned through multi-branch parallel learning; during inference, these are compressed into a single branch via reparameterization while maintaining feature representation capabilities while reducing computational complexity. In summary, the C3K2_RepVGG module embodies both theoretical innovation and engineering application value, providing a feasible solution for efficient intelligent detection of wood surface defects.

3.2. ADown (Adaptive Downsampling) Module

Traditional convolutional downsampling often loses small defect details, leading to high miss rates in detecting small

targets such as millimeter-scale cracks. To address this issue, this paper introduces the ADown (Adaptive Downsampling) module to optimize the Backbone network. This module serves as an adaptive lightweight downsampling mechanism that selectively retains detail features such as fine cracks

while avoiding information loss caused by forced compression from traditional convolutions, significantly enhancing detection capabilities for millimeter-scale defects on wood surfaces.

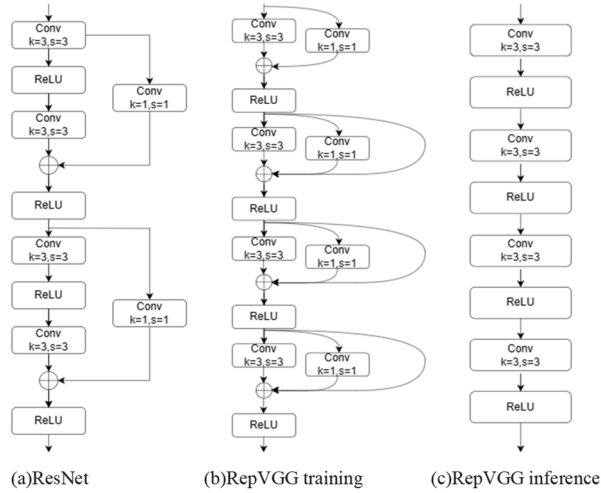


Figure 3. Schematic diagram of RepVGG Architecture

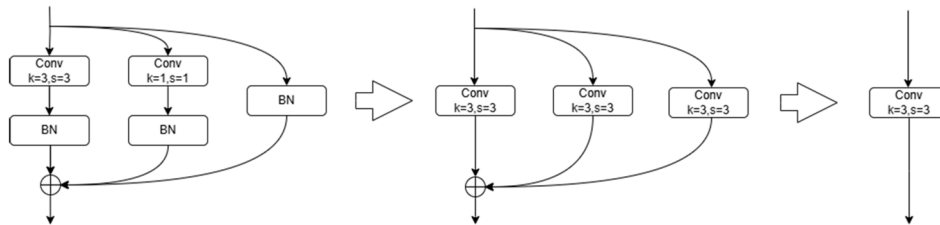


Figure 4. Structural Reparameterization process

The ADown module employs a multi-branch convolution structure that can dynamically adjust kernel sizes ($3 \times 3 / 5 \times 5$) and stride parameters ($1/2$) based on input features. The specific process is illustrated in Figure 5: after input feature map X (with c_1 channels) undergoes average pooling ($k=2, s=1$), it is split into sub-feature maps X_1 and X_2 (each with $c_1/2$ channels); both branches are processed through Conv1 (a 3×3 convolution with $s=2, p=1$) and “Maxpool (a 3×3 pool with $s=2, p=1$) + Conv2 (a 1×1 convolution)” respectively to achieve downsampling by half while avoiding loss of edge textures effectively reducing miss rates for wormholes and micro-cracks. To enhance feature expression, the module achieves multi-scale fusion through dual branches: the local branch uses a 1×1 convolution to extract fine features while the global branch employs a grouped convolution with a size of 3×3 to capture contextual information; subsequently, channel mixing breaks group isolation and embeds an ECA attention mechanism to strengthen defect area representation effectively integrating high-frequency edges with global semantics to improve defect recognition under complex texture backgrounds.

In terms of lightweight design, the ADown module combines grouped convolutions with depthwise separable convolutions and adopts a non-redundant connection structure that reduces parameter size and computational load by 40%–50% compared to traditional strided convolutions under similar downsampling conditions. Through dynamic adjustments during training and structural optimizations during inference working in synergy, this module significantly enhances computational efficiency while

ensuring detection accuracy. In summary, the ADown module demonstrates promising application prospects in both theoretical innovation and engineering practice providing an efficient and feasible solution for wood defect detection.

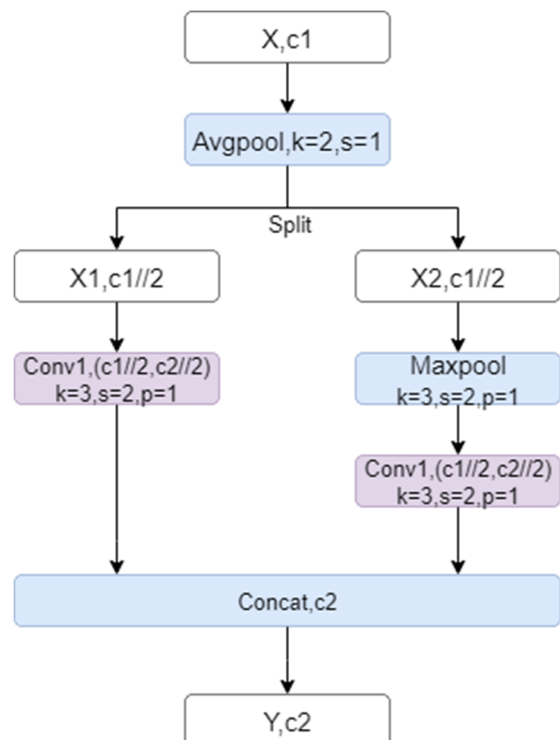


Figure 5. Structure of ADown Module

3.3. C2BRA Convolutional Enhanced Dual-Branch Bi-Level Routing Attention Module

To address issues where traditional algorithms struggle with distinguishing defects from background interference due to complex wood grain patterns (leading to high false positive rates by misclassifying normal textures as defects), this paper introduces the Bi-level Routing Attention (BRA) mechanism proposed by BiFormer and deeply integrates it with detection algorithms to design the C2BRA module suitable for wood defect detection. This mechanism utilizes a dual-layer dynamic sparse structure composed of “region-level routing” and “token-level attention” for optimization significantly enhancing detection performance.

The specific process is illustrated in Figure 6: first, for feature maps $X \in \mathbb{R}^{H \times W \times C}$ extracted from the backbone network, the BRA module performs region partitioning and input projection dividing these maps into $S \times S$ non-overlapping regions each containing HW/S^2 feature vectors forming region-level feature tensors $X^r \in \mathbb{R}^{S^2 \times HW/S^2 \times C}$; then linear projections generate query matrix Q , key matrix K , and value matrix V . Next comes region-to-region routing operations: by performing mean pooling on Q and K within each region we obtain region-level query matrix Q^r and key matrix K^r calculating region affinity matrix A^r as described in Equation (1); then using top-k operations we select the top k most relevant regions generating routing index matrix I^r which facilitates preliminary localization of defect-related areas effectively filtering out most irrelevant backgrounds. Finally, token-to-token attention calculations are performed aggregating key-value pairs from relevant regions via I^r yielding aggregated key matrix K^g and value matrix V^g calculated according to Equations (2) and (3); ultimately output features O are computed using Equation (4), completing weighted fusion where locally enhanced context $LCE(V)$, achieved via deep convolutions further strengthens local details such as defect edges significantly improving YOLOv11’s ability to recognize subtle defects.

This structure has three major advantages: first it overcomes limitations imposed by local receptive fields in convolutions effectively capturing long-range associations between defects within wood textures; second it implements a query-aware dynamic sparsity mechanism allowing different regions to select differentiated focus objects based on semantic needs; third it suppresses interference from complex textures through sparse computations significantly reducing false positive rates while avoiding high complexity and memory overhead associated with global attention thus achieving balance between accuracy and real-time performance. In conclusion, the C2BRA module embodies both theoretical innovation and engineering application value providing a feasible yet efficient solution for wood defect detection.

$$A^r = Q^r (K^r)^T \quad (1)$$

$$K^g = \text{gather}(K, I^r) \quad (2)$$

$$V^g = \text{gather}(V, I^r) \quad (3)$$

$$O = \text{Attention}(Q, K^g, V^g) + \text{LCE}(V) \quad (4)$$

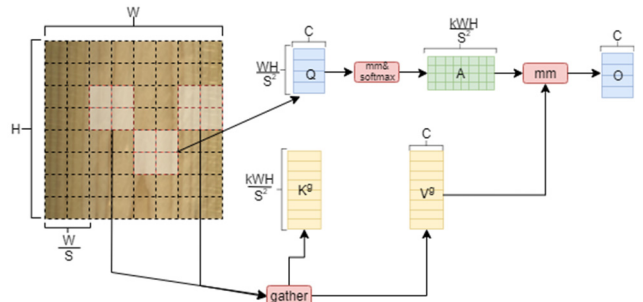


Figure 6. Schematic Diagram of BRA Attention Mechanism Process

4. Experimental Comparison and Analysis

4.1. Dataset Construction

This study utilizes a complete wood surface defect dataset provided by the Czech Technical University in Ostrava, consisting of a total of 20,275 images, of which 18,283 contain one or more defects and 1,992 are defect-free. The original annotations of the dataset cover a total of 12 defect types including live knots, dead knots, and others, with all images having a resolution of $2,800 \times 1,024$. Statistical analysis revealed two main issues with the dataset: first, some categories have significantly more samples than others leading to severe imbalance; second, certain defect categories exhibit highly similar visual features which can confuse the model during training. These issues result in insufficient learning for low-sample categories and demonstrate uneven detection accuracy under complex backgrounds.

To address these problems, this paper implemented systematic processing during data selection and augmentation stages. Initially, images with blank labels or those that were blurry or severely damaged were removed to ensure data quality. Subsequently, various data augmentation strategies were introduced including sharpening, brightness adjustment, random hue modification, and horizontal flipping to expand sample size and increase diversity. Particularly for defect categories with fewer samples, targeted augmentation strategies were employed effectively alleviating class imbalance issues and providing more sufficient and balanced sample support for subsequent training.

Following these preprocessing and augmentation steps, the label distribution of the dataset was significantly balanced with each defect category having over 1,000 samples. The final constructed dataset contains 5,007 high-quality images as shown in Table 1 covering all target defect categories and is divided into training set, validation set, and test set in an 8:1:1 ratio. This series of processing operations not only ensured the reliability and balance of the data but also enhanced the model’s ability to learn from different defect categories. In summary, the constructed dataset possesses favorable characteristics in terms of sample size, category distribution, and data quality providing a solid foundation for the training and evaluation of subsequent wood defect detection models.

Table 1. Comparison of Dataset Labels

Dataset Name	Quartzity	Live Knot	Marrow	resin	Dead_Knot	knot_with_crack	Knot_missing	Crack
Czech Dataset	503	1075	2276	11985	21224	1181	2169	3455
Final Dataset	1144	1068	2267	3560	6389	1175	2113	3434

4.2. Experimental Environment Setup and Training Parameter Configuration

The experiments were conducted on a high-performance computing platform equipped with an Intel Core™ i7-12700K processor (12th generation), NVIDIA GeForce RTX 3060 graphics card, and 32GB RAM; the operating system

was Windows 11 while the development environment included Python 3.8, PyTorch 1.8.0, CUDA 11.0, and cuDNN 8.1.5. The initial learning rate was set to 0.01 with a cosine decay function applied for learning rate adjustments at a period of 0.01. Input image size was set to 1280×1280 with a batch size of four; training was conducted for a total of 300 epochs; experimental environment details and related configurations are shown in Table 2.

Table 2. Experimental Environment

Name	Related configuration
Operating system	Windows 11 (64 bits)
Python/Pytorch	3.8/1.8.0
CUDA/CUDNN	11.0/8.1.5
CPU	12th Gen Intel® Core™ i7-12700K 3.61 GHz
GPU	NVIDIA GeForce RTX 3060
Memory	32GB

4.3. Evaluation Metrics

This study employs three key evaluation metrics to assess the effectiveness of the improved algorithm: precision, recall, and mean average precision (mAP). The mathematical definitions and calculation formulas for each metric are as follows:

(1) Precision: This refers to the proportion of correctly detected targets by the algorithm out of all detected targets. TP represents true positives (the number of targets correctly predicted as positive), while FP represents false positives (the number of targets incorrectly predicted as positive). Its mathematical formula is as follows:

$$\text{Precision} = \frac{TP}{TP+FP} \quad (5)$$

(2) Mean Average Precision: This refers to the average value of average precision calculated across multiple categories used to measure overall retrieval performance across all classes. Q_R is the total number of categories; q

represents each category; $AP(q)$ is the average precision for category q ; its mathematical expression is as follows:

$$mAP = \frac{1}{|Q_R|} \sum_{q \in Q_R} AP(q) \quad (6)$$

(3) FPS (Frames Per Second): This measures how many frames per second the model can process reflecting both real-time performance and inference efficiency of the algorithm.

4.4. Comparative Experiments

To further validate the superiority of the proposed YOLOv11-RABRA algorithm in object detection tasks this paper conducts comparative experiments against representative models such as RT-DETR, SSD, and various YOLO series models using the same wood surface defect dataset; experimental results are shown in Table 3. As seen from Table data there are significant differences among different models regarding detection accuracy parameter scale and inference speed.

Table 3. Comparative Experimental Results

Model	Precision /% P	Recall rate /% R	mAP50/%	Parameters /M	FPS / (frame · s ⁻¹)
RT-DETR	60.8	58.4	57.9	8.7	10.15
SSD	61.3	59.9	58.1	9.5	10.28
YOLOv3	63.8	60.5	59.5	9.8	12.05
YOLOv5	62.4	60.1	58.8	6.2	15.00
YOLOv6	61.7	59.8	59.2	4.2	15.58
YOLOv7-tiny	63.4	60.2	58.9	3.9	15.24
YOLOv8	62.9	60.4	60.7	3.1	12.84
YOLOv9-tiny	63.5	58.9	61.8	2.8	14.74
YOLOv10	63.2	59.7	62.4	2.9	19.87
Yolov11	64.9	61	64.3	2.5	36.16

In terms of accuracy performance YOLOv11 achieved a precision rate of 64.9%, recall rate of 61%, and mean average precision (mAP50) reaching up to 64.3%, significantly outperforming most models such as RT-DETR and SSD effectively enhancing recognition completeness and accuracy for wood surface defects while reducing occurrences of missed detections and false positives. Regarding model lightweighting YOLOv11 has only 2.5M parameters far less than RT-DETR's 8.7M demonstrating higher structural compactness along with better computational resource adaptability. Meanwhile its frames per second (FPS) reaches 36.16 frames · s⁻¹ showcasing notable advantages in inference speed meeting industrial requirements for online

detection and real-time alerts.

In conclusion, YOLOv11 achieves a good balance between detection accuracy, model lightweighting, and inference efficiency, possessing theoretical performance advantages while also meeting engineering deployment requirements. This model demonstrates outstanding comprehensive competitiveness in wood surface defect detection tasks providing reliable technical support for building intelligent detection systems.

4.5. Ablation Experiments

To systematically validate the effectiveness of the improved algorithm in wood surface defect detection tasks,

ablation experiments were designed to quantify the independent contributions and synergistic effects of each improvement module. The experiments focused on typical defect categories such as live knots, dead knots, and cracks, employing a stepwise introduction of improvement modules to construct multiple comparative models. Table 4 summarizes the experimental results under different combinations of improvement modules, highlighting differences in precision, recall, mean average precision (mAP50), parameter size, and inference speed among the models, clearly demonstrating the independent contributions and synergistic effects of each module.

The baseline model (Model 1), without any improvements, exhibited insufficient feature extraction integrity, weak detail retention capabilities, and limited differentiation between defects and wood grain backgrounds, resulting in low precision and recall levels that could not meet high-accuracy detection requirements, thereby validating the necessity for improvements. Building upon this foundation, Single Class Improve (SI) was first introduced to effectively alleviate class imbalance issues by augmenting data specifically for categories with fewer samples. Although SI increased parameter count and complexity, it significantly contributed to improvements in detection accuracy. Further integration of the C3K2_RepVGG module (Model 3) into the model enhanced feature expression capabilities through reparameterization structures and optimized feature fusion mechanisms, allowing continuous optimization based on data augmentation and validating the synergistic effects between SI and C3K2_RepVGG. The addition of the ADown module (Model 4) further improved detail retention through an adaptive downsampling mechanism, significantly enhancing performance especially for fine-grained defects like small cracks while avoiding information loss associated with traditional downsampling thus achieving a balance between

detection accuracy and efficiency. Finally, the introduction of the C2BRA module (Model 5) strengthened the focus on defect areas through a dual-layer routing attention mechanism significantly suppressing interference from complex wood grain backgrounds. This model achieved optimal overall performance: a precision rate of 77.7%, recall rate of 75.4%, mAP50 reaching 78%, with parameter size controlled at 2.15M and frame rate recovering to 17.44 FPS.

Based on quantitative result validation, further visual analysis was conducted to assess each module's contribution to detection performance. Figure 7 illustrates the detection performance of models before and after improvements on typical defects such as dead knots, cracks, resin pockets, and pith; this qualitatively verifies the feature enhancement and false positive suppression effects of the improvement modules. Results indicate that the improved model exhibits more stable and accurate performance on defects with fuzzy boundaries or smaller sizes particularly showing significant reductions in missed detections and false positives in scenarios involving dead knots accompanied by cracks or knots.

In summary: SI establishes data foundations; C3K2_RepVGG enhances feature representation; ADown achieves a balance between detail retention and lightweighting; C2BRA improves defect-background differentiation. The synergistic effects among various modules across data augmentation, feature extraction, downsampling optimization, and attention mechanisms ultimately led to comprehensive enhancements in wood surface defect detection performance. In conclusion, the improved algorithm achieves comprehensive optimization in terms of accuracy efficiency and lightweighting for wood defect detection providing feasible solutions for high-precision detection in intelligent manufacturing and industrial quality inspection.

Table 4. Ablation Experimental Results

Model	SI	C3K2_RepVGG	ADown	C2BRA	Precision /%	Recall rate /%	mAP50 /%	Parameters /M	FPS / (frame · s ⁻¹)
1	-	-	-	-	64.9	61	64.3	2.58	36.16
2	√	-	-	-	73.9	70.8	74.5	2.82	20.17
3	√	√	-	-	74	73.7	75.7	3.32	19.85
4	√	√	√	-	74.6	74.4	77.7	2.13	17.16
5	√	√	√	√	77.7	75.4	78	2.15	17.44

5. Conclusion

To address issues such as low recognition rates for small defects and strong background interference in wood defect detection, this paper proposes the improved YOLOv11-RABRA algorithm. In terms of method design, the first step is to alleviate class imbalance through the SI module, enhancing the learning ability for low-sample defect categories. Next, the C3K2_RepVGG module is introduced, utilizing structural reparameterization to strengthen feature extraction and fusion capabilities. Simultaneously, the ADown adaptive downsampling mechanism is incorporated to effectively retain fine-grained features such as cracks, achieving a balance between detection accuracy and

lightweighting. Finally, the C2BRA attention mechanism is employed to enhance the focus on defect areas significantly suppressing interference from complex wood grain backgrounds, thereby comprehensively improving the model's robustness and detection performance. Experimental results indicate that compared to the original YOLOv11, YOLOv11-RABRA achieves improvements of 13.7%, 12.8%, and 14.4% in mAP50, precision, and recall respectively while maintaining real-time detection capabilities, validating the independent contributions and synergistic effects of each improvement module. In summary, YOLOv11-RABRA achieves a good balance between detection accuracy, efficiency, and lightweighting providing an efficient and feasible technical pathway for wood defect detection while also laying a methodological foundation for high-precision

detection applications in intelligent manufacturing and industrial quality inspection.

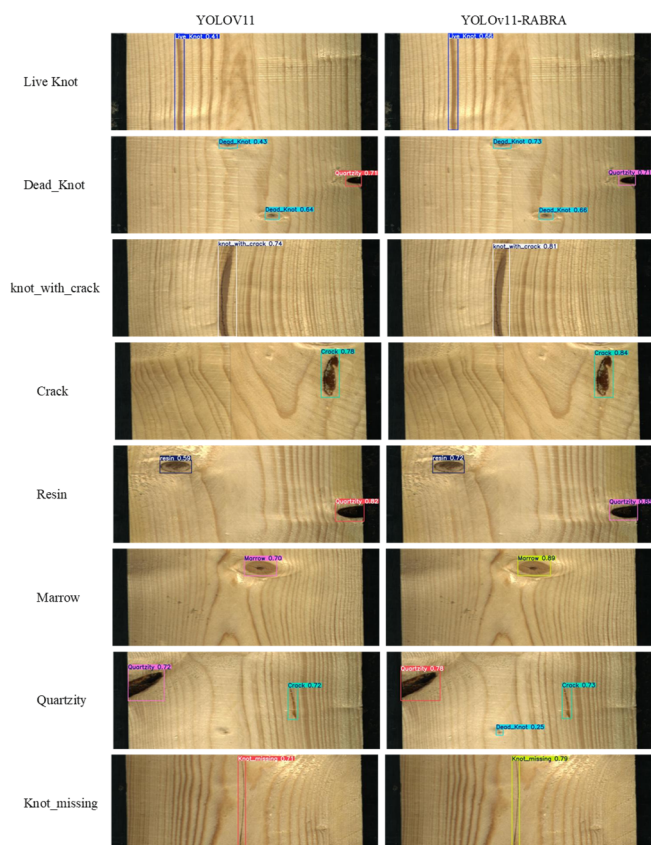


Figure 7. Visualization Results

References

- [1] WANG Jun-dong, ZHAO Qian, LIU Ying, et al. Design of Laser Scanning and Detection Devices for Wood Surface Defects [J]. *Forestry Machinery & Woodworking Equipment*, 2016, 44(11): 24-27.
- [2] SHEN Yin-xi, LIU Ying, YANG Yu-tu. Surface defect detection of solid wood board based on improved YOLOv5 algorithm [J]. *Forestry Machinery & Woodworking Equipment*, 2024, 52(3): 24-29.
- [3] ZHOU Zhu, YIN Jianxin, ZHOU Suyin, ZHOU Houkui. Knot detection on coniferous wood surfaces based on near infrared spectroscopy [J]. *Journal of Zhejiang A&F University*, 2017, 34(3): 520-527.
- [4] Qi D W, Zhang P, Zhang X F, et al. Edge detection of wood defects in X-ray wood image using neural network and mathematical morphology [C]// Technical Committee on Control Theory, Chinese Association of Automation. Proceedings of the 29th Chinese Control Conference. Beijing: China Machine Press, 2010: 2450-2455.
- [5] Li C, Zhang Y, Tu W, et al. Soft measurement of wood defects based on LDA feature fusion and compressed sensor images[J]. *Journal of Forestry Research*, 2017, 28(6): 1285-1292.
- [6] XIE Le; HENG Xi-dan;LIU Yang;JIANG Qi-long;LIU Dong. Transformer fault diagnosis based on linear discriminant analysis and step-by-step machine learning [J]. *Journal of Zhejiang University: Engineering Science*, 2020, 54(11): 2266-2272.
- [7] Yuce B, Mastrocinque E, Packianather M S, et al. Neural network design and feature selection using principal component analysis and Taguchi method for identifying wood veneer defects[J]. *Production & Manufacturing Research*, 2014, 2(1): 291-308.
- [8] Luo Jiancheng, Zhou Chenghu, Yang Yan. ANN Remote Sensing Classification Model and Its Integration Approach with Geo-knowledge [J]. *NATIONAL REMOTE SENSING BULLETIN*, 2001, 5(2): 122-129.
- [9] ZHAO Shuxuan;ZHANG Jie;WANG Junliang;XU Chuqiao. Fabric Defect Detection Algorithm Based on Two-stage Deep Transfer Learning [J]. *Journal of Mechanical Engineering*, 2021, 57(17): 86-97.
- [10] WANG Ning;ZHI Min. Review of One-Stage Universal Object Detection Algorithms in Deep Learning [J]. *Journal of Frontiers of Computer Science and Technology*, 2025, 19(5): 1115-1140.
- [11] Shanshan Xu, Yingan Liu, Sheng Xu. Wood defects recognition based on the convolutional neural network [J]. *Journal of Shandong University (Engineering Science)*, 2013, 43(2): 23-28.
- [12] Xia B, Luo H, Shi S. Improved Faster R-CNN based surface defect detection algorithm for plates[J]. *Computational Intelligence and Neuroscience*, 2022: 1-11.
- [13] LI Haiying; SUN Yue; ZHANG Xiao;SONG Jiancheng. Fault Diagnosis of Vacuum Contactor Based on Modal Time-frequency Diagram and ResNet50 [J]. *High Voltage Engineering*, 2023, 49(5): 1831-1840. DOI:10.13336/j.1003-6520.hve.20221305.
- [14] ZHAO Fei;ZHANG Wenkai;YAN Zhiyuan;YU Hongfeng; DIAO Wenhui. Multi-feature Map Pyramid Fusion Deep Network for Semantic Segmentation on Remote Sensing Data [J]. *Journal of Electronics & Information Technology*, 2019, 41(10): 2525-2531.
- [15] Li D, Xie W, Wang B, et al. Data augmentation and layered deformable Mask R-CNN-based detection of wood defects[J]. *IEEE Access*, 2021, 9: 108162-108174.
- [16] Yu Jiayong, Li Feng, Xue Xiankai, et al. Intelligent Identification of Bridge Structural Cracks Based on Unmanned Aerial Vehicle and Mask R-CNN [J]. *China Journal of Highway and Transport*, 2021, 34(12): 80-90. DOI:10.19721/j.cnki.1001-7372.2021.12.007.
- [17] Hu Zhongkang. Research on Optimization System of Solid Wood Based on Machine Vision [D]. Nanjing: Nanjing Forestry University, 2019.
- [18] CHEN Dehai; SUN Shiru;WANG Yuzhao;LEI Zhijun. Research on small target detection based on improved SSD algorithm [J]. *Transducer and Microsystem Technologies*, 2023, 42(3): 65-68+72.
- [19] Jia Haonan, Xu Huadong, Wang Lihai, et al. Quantitative identification of surface defects in wood paneling based on improved YOLOv5 [J]. *Journal of Beijing Forestry University*, 2023, 45(4): 147-155.
- [20] QIU Tianheng;WANG Ling;WANG Peng;BAI Yan'e. Research on Object Detection Algorithm Based on Improved YOLOv5 [J]. *Computer Engineering and Applications*, 2022, 58(13): 63-73.
- [21] Han S, Jiang X, Wu Z. An improved YOLOv5 algorithm for wood defect detection based on attention[J]. *IEEE Access*, 2023, 11: 71800-71810.
- [22] Zhang Zilei. Research on Defect Detection Algorithm of Sawn Timber Based on Deep Learning [D]. Harbin: Harbin University of Science and Technology, 2023.
- [23] GENG Yanli;LIN Yanbo;FU Yanfang;YANG Shucui. Study on pigs target detection algorithm based on Transformer and adaptively spatial feature fusion [J]. *Journal of Northeast Agricultural University*, 2023, 54(1): 88-96.
- [24] Yang Feifan, Xu Weicheng, Chen Shengde, et al. Rice disease image classification based on 1 fusion of Focal Loss and typical convolutional neural network structure [J]. *Jiangsu Agricultural Sciences*, 2023, 51(14): 198-204. DOI: 10.15889/j.issn.1002-1302.2023.14.027.
- [25] Cheng R J, Yang Y, Li L W, et al. Lightweight residual network based on depthwise separable convolution for hyperspectral image classification [J]. *Acta Optica Sinica*, 2023, 43(12): 311-320.