DWBA-YOLO: A Dual-Layer Weighted Background-aware Network for Multi-Scale Particleboard Surface Defect Detection

Anning Ding, Chao He, Yongkang Kang, and Huaqiong Duo *

Automated defect detection is crucial for particleboard manufacturing, enabling precise quality control and improved production efficiency. However, existing approaches face three key challenges: small-scale defects, low visual contrast between defects and surrounding regions, and severe texture interference from complex backgrounds, which collectively undermine feature extraction and multi-scale representation. To address these issues, this study developed DWBA-YOLO, a multi-scale surface defect detection network tailored for complex texture scenarios. First, an Adaptive Dual-layer Weighting Mechanism (ADWM) was introduced, where Intra-Feature Weighting suppresses texture-dominated channel responses while Cross-Feature Weighting adaptively calibrates contributions from different pyramid levels. Second, an Adaptive Spatial Feature Fusion head was designed to learn spatially varying fusion weights and to mitigate cross-scale inconsistencies while maintaining lightweight overhead. Third, Normalized Wasserstein Distance was incorporated to enhance small-scale defect localization. Extensive experiments demonstrated the effectiveness of the method. On a proprietary particleboard defect dataset, DWBA-YOLO improved recall by 4.7%, precision by 4.2%, mAP@50 by 3%, and mAP@50:95 by 2.5% compared with YOLOv8n, while reducing computational complexity by 43%. These results indicate that DWBA-YOLO is effective and practical for real-time particleboard defect detection.

DOI: 10.15376/biores.21.2.3856-3871

Keywords: Particleboard surface defect detection; Deep learning; YOLO; Multi-scale feature fusion; Adaptive Dual Weighting Module (ADWM); Object detection

Contact information: College of Materials Science and Art Design, Inner Mongolia Agricultural University, Hohhot 010018, Inner Mongolia Autonomous Region, China;
** Corresponding author: duohuaqiong@imau.edu.cn*

INTRODUCTION

Particleboard is an important engineered wood product widely used in furniture manufacturing and architectural decoration. During high-throughput production, surface defects may develop that degrade mechanical performance and service safety, making reliable defect detection critical. However, particleboard surface inspection remains highly challenging. As illustrated in Fig. 1, defects present three key difficulties: (1) small scales (defined as <32×32 pixels or <0.92 cm^2 per MS COCO metrics), primarily including Oil Stain, Char Spot, and Oversized Chip in the present dataset (Fig. 1a); (2) low visual contrast that makes defects barely distinguishable from surrounding regions (Fig. 1b); and (3) strong texture interference from wood-grain backgrounds (Fig. 1c). These factors collectively complicate reliable detection. Undetected defects can shorten product service

life and pose safety risks. For instance, surface looseness facilitates moisture penetration, leading to hygroscopic swelling, adhesive deterioration, and increased risk of structural failure.

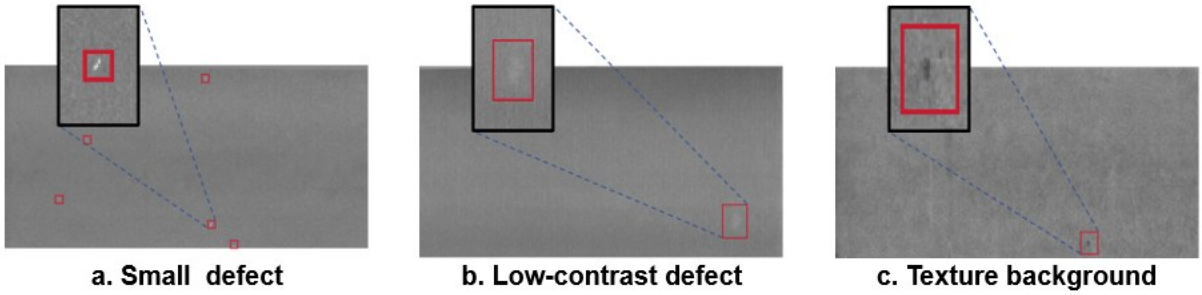

Fig. 1. Representative examples of three key challenges in particleboard surface defect detection

Accordingly, automated visual inspection systems have become increasingly important for in-line quality control. Nevertheless, achieving reliable detection of small-scale defects while maintaining real-time efficiency under complex textures remains challenging. Recent advances in industrial defect inspection have explored segmentation-based approaches (Liu *et al.* 2024), robust optimization methods (AlBahar *et al.* 2022), and multi-modal signal processing (Liu *et al.* 2023). However, particleboard surface inspection poses unique challenges, requiring lightweight real-time models that remain robust to both texture interference and minute defects.

To address the need for real-time defect detection, recent studies have adapted YOLO-based detectors (Redmon *et al.* 2016; Jocher *et al.* 2023) to handle the texture-heavy backgrounds of wood panels. Zhao *et al.* (2021) pioneered this approach by integrating Ghost Bottlenecks and SELayer into YOLOv5, achieving 92.1% mAP@0.5 with 62.5% model compression. However, their attention mechanism applies uniform channel weighting, failing to adaptively suppress texture-dominated channels in complex backgrounds. Zhao *et al.* (2022) further advanced these efforts by coupling semantic segmentation with detection using YOLOv5-Seg-Lab and introducing channel attention (FCAM), reporting 93.20% mAP at 56.02 FPS. Yet, their fusion strategy relies on direct concatenation without spatial adaptation, leading to cross-scale semantic conflicts where the same spatial location is treated as foreground at one scale but background at another.

More recently, He *et al.* (2025) proposed LE-YOLO, which introduces the Normalized Wasserstein Distance (NWD) to enhance small-scale defect localization, achieving a compact size of 2.10 M parameters and 5.5 GFLOPs. However, its evaluation was limited to three defect categories, and a relatively low mAP@50:95 of 0.55 indicates insufficient localization precision under strict IoU thresholds. Yan *et al.* (2025) achieved a high mAP@0.5 of 0.944 by combining traditional pre-screening with an enhanced YOLOv10 detector. Despite this accuracy gain, the reliance on pre-screening introduces computational overhead and complicates end-to-end deployment. Other lightweight approaches, such as GBCD-YOLO (Zheng *et al.* 2024) and loss-refined models (Wang 2025), reported improved efficiency. However, they still employ fixed feature fusion weights, limiting their ability to dynamically prioritize informative scales amidst severe texture noise.

Beyond wood-based materials, methods for other industrial surfaces—such as steel (Hoang *et al.* 2025; Ma *et al.* 2025) and copper (Liang *et al.* 2025; Wan *et al.* 2025)—have

emphasized adaptive feature fusion strategies (Liu *et al.* 2019). Notably, Ma *et al.* (2025) proposed ELA-YOLO with linear attention, achieving 81.7% mAP at 101 FPS on NEU-DET, while Wan *et al.* (2025) introduced the dual-branch DFSDNet, reporting 88.53% mAP on KUST-DET. Although these methods demonstrate cross-scenario robustness, they inherit a critical limitation when applied to wood: their reliance on global fusion weights prevents adaptive responses to local texture complexity, which is particularly severe in particleboard with chaotic surface patterns and scale-varying defects.

Despite achieving high mAP@0.5 scores often exceeding 90%, existing methods face three persistent limitations regarding the challenges shown in Fig. 1: (1) uniform channel weighting fails to suppress texture noise, reducing sensitivity to smaller defects; (2) naïve pyramid fusion causes cross-scale semantic conflicts; (3) IoU-based losses degrade with minimal target overlap, destabilizing localization supervision.

To address these issues, the authors propose DWBA-YOLO, a Dual-layer Weighted Background-Aware network built on the YOLOv8 baseline (Jocher *et al.* 2023) for real-time particleboard inspection. The contributions are threefold:

1. Adaptive Dual-layer Weighting Mechanism (ADWM): A Dual-branch module where Intra-Feature Weighting (IFW) suppresses texture noise, and Cross-Feature Weighting (CFW) adaptively balances pyramid levels for varying defect scales.
2. Adaptive Spatial Feature Fusion (ASFF) head: A head that learns location-specific weights to mitigate cross-scale semantic conflicts, improving robustness against small-scale defects.
3. Normalized Wasserstein Distance (NWD) loss: An auxiliary localization metric providing stable supervision for smaller targets when IoU-based metrics fail.

Extensive experiments demonstrate that DWBA-YOLO achieves state-of-the-art accuracy with reduced computational costs, making it highly suited for texture-heavy industrial inspection.

EXPERIMENTAL

Particleboard Surface Defects Dataset

The particleboard surface defect dataset comprises eight categories (Fig. 2): Char Spot, Chip Out, Dent, Glue Spot, Oil Stain, Oversized Chip, Sanding Mark, and Soft Spot.

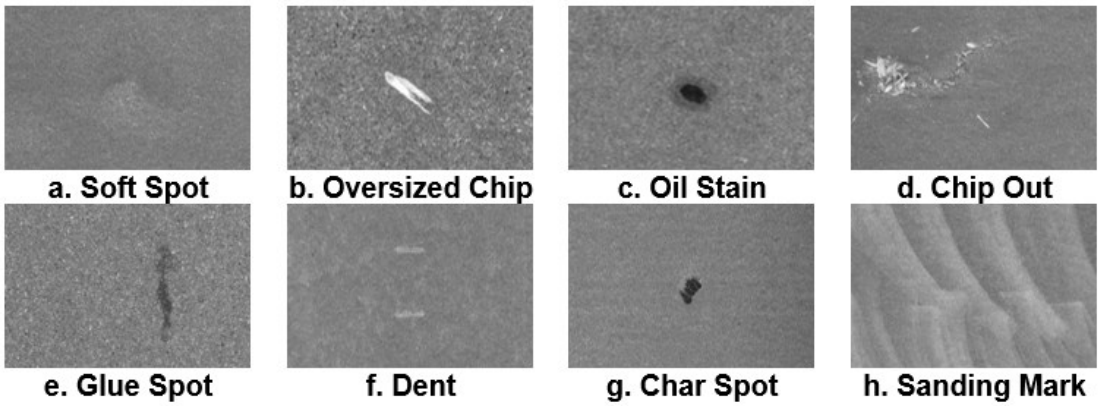

Fig. 2. Representative examples of the eight particleboard surface defect categories

Images of standard 2440×1220mm boards were captured using a Hikrobot MV-CL043-A1GC industrial camera, yielding raw 8133×4096 pixel images with a spatial resolution of approximately 0.30 mm/pixel. To prevent feature degradation from global down-sampling, images were not resized; instead, they were cropped to retain only the regions of interest (ROIs), resulting in 640×640 pixel patches that preserve the original physical scale. This produced 4,262 patches with 6,542 bounding-box annotations, partitioned into training, validation, and testing sets (8:1:1 ratio) *via* stratified sampling. The dataset exhibits instance-level class imbalance, with Oversized Chip accounting for 45.3% of annotations and Oil Stain only 5.6% (Table 1). To mitigate this during training, minority classes (<600 annotations) were oversampled 4 to 8 times to approximate the dominant class size. Validation and test sets retained their original distributions.

Table 1. Distribution of Defect Categories in the Particleboard Surface Defects Dataset

Defect Category	Annotations	Images	Defect Category	Annotations	Images
Char Spot	533 (8.1%)	494	Oil Stain	367 (5.6%)	347
Chip Out	452 (6.9%)	405	Oversized Chip	2961 (45.3%)	1685
Dent	759 (11.6%)	102	Sanding Mark	389 (5.9%)	359
Glue Spot	560 (8.6%)	453	Soft Spot	520 (8.0%)	426

DWBA-YOLO

YOLOv8 is a widely adopted one-stage object detection framework offering a strong balance between accuracy and inference efficiency, making it a practical baseline for industrial visual inspection. However, particleboard surface defect detection remains challenging because defects are often tiny, low-contrast, and irregularly shaped, amidst strong interference from complex textures and illumination variations. To address these issues, the present approach employed an accuracy-efficiency balanced detection network on top of YOLOv8. Specifically, an Adaptive Dual-layer Weighting Mechanism (ADWM) was introduced into the feature fusion pathway to enhance cross-level information interaction. Furthermore, an Adaptive Spatial Feature Fusion (ASFF) detection head was designed to suppress cross-scale inconsistency under complex textured backgrounds.

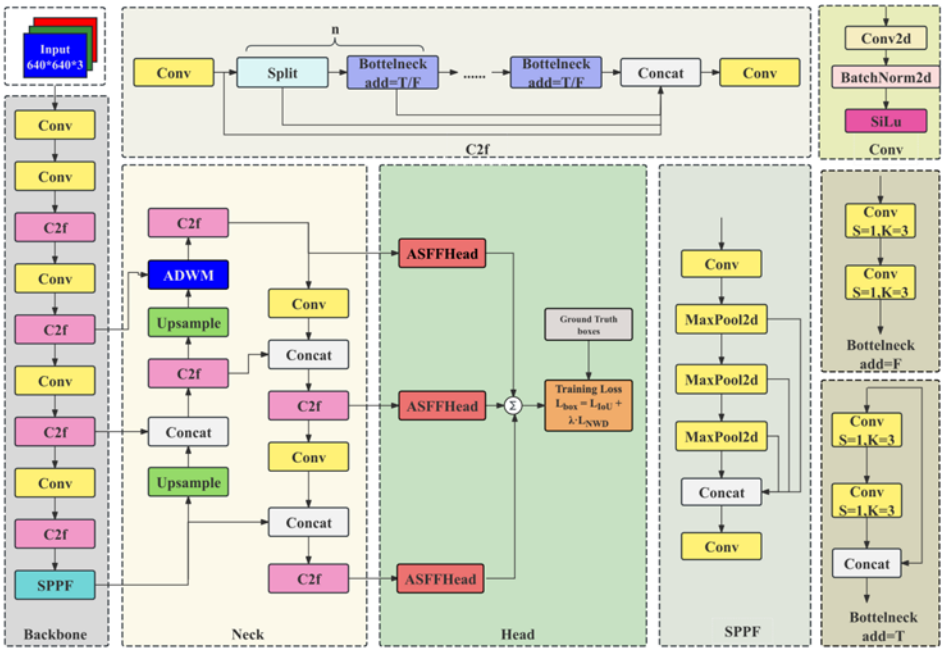

Fig. 3. Overall architecture of DWBA-YOLO. The network extends YOLOv8 with ADWM, an ASFF head, and NWDLoss

In addition, a Normalized Wasserstein Distance (NWD) term was incorporated into the localization loss to strengthen small-scale and low-contrast defect recognition. The overall architecture of the proposed method is illustrated in Fig. 3. This section details the core components of the approach, including ADWM, the ASFF head, and the integration of NWDLoss into the training objective.

Adaptive dual-layer weighting mechanism (*ADWM*)

Feature fusion in one-stage detectors typically relies on concatenation or summation, implicitly assuming (i) uniform channel importance and (ii) equal reliability across pyramid levels after spatial alignment. However, these assumptions often fail for texture-dominant industrial surfaces with small-scale defects. Texture-driven channels can suppress low-contrast defect cues, and optimal pyramid levels vary by defect scale; shallow features favor smaller targets, while deeper features may introduce background dominance. To address these issues, the current version used the Adaptive Dual-layer Weighting Mechanism (ADWM) (Huang *et al.* 2025), which decouples fusion into intra-feature channel reweighting and cross-feature level reweighting, yielding a more selective and instance-aware representation. Figure 4 illustrates the ADWM workflow (Fig. 4a) and the channel-weight generator (CACW) structure (Fig. 4b).

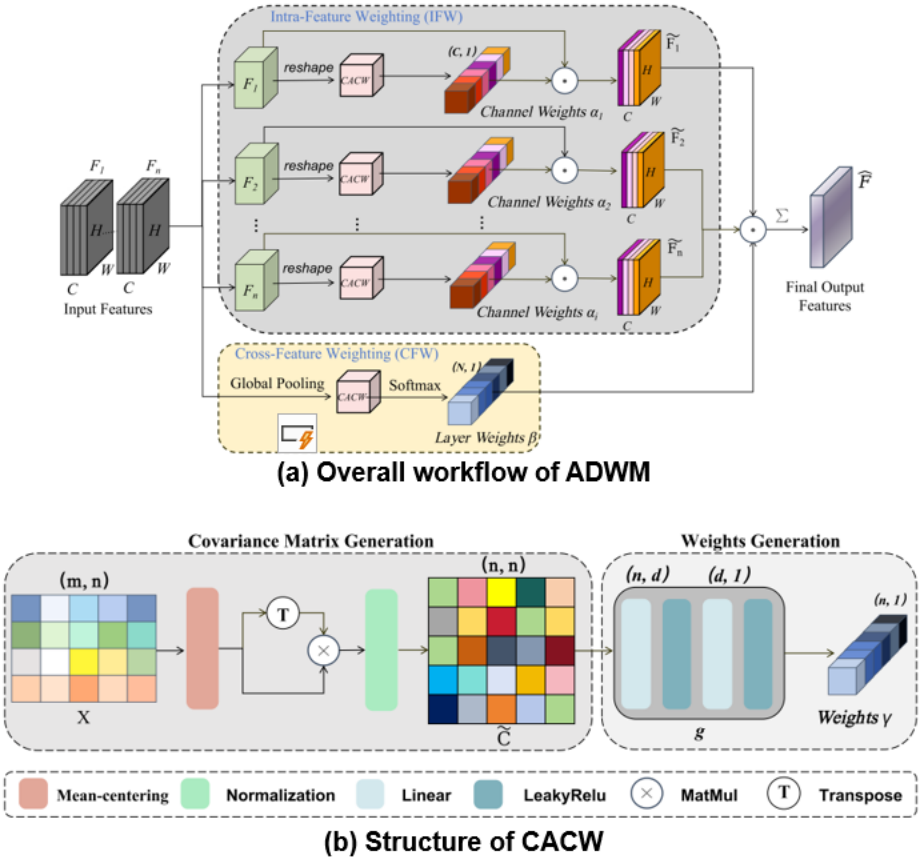

Fig. 4. Overall architecture of the proposed Adaptive Dual-layer Weighting Mechanism (ADWM)

(1) CACW: Cross-channel Adaptive Collaborative Weighting

To explicitly capture inter-channel dependencies and suppress texture noise, CACW computes a covariance-style representation. As shown in Fig. 4(b), given a reshaped 2D input $X \in R^{m \times n}$ (where m and n denote the flattened spatial resolution and channel dimensions, respectively), the input X is first mean-centered. The covariance-style matrix is then calculated *via* matrix multiplication between the mean-centered X and its transpose, followed by a normalization operation to yield the dependency matrix $\hat{C} \in R^{n \times n}$. Finally, a weight generation block g, comprising linear transformations and LeakyReLU activations, projects $\hat{C}$ into the final adaptive weights $\gamma \in R^{n \times 1}$.

(2) IFW: Intra-Feature Weighting

The IFW branch is designed to dynamically highlight defect-relevant channels while dampening background interference. For the i-th input feature map $F_i \in R^{C \times H \times W}$ from the feature pyramid, it is reshaped and fed into CACW to generate the channel-specific weight vector $\alpha_i \in R^{C \times 1}$. The refined feature $\widehat{F}_i$ is obtained via element-wise multiplication (denoted by $\odot$) along the channel dimension:

$$\widehat{F}_i = F_i \odot \alpha_i \quad (1)$$

Operationally, this acts as a numerical filter, suppressing texture noise by multiplying background-dominated channels by near-zero values.

(3) CFW: Cross-Feature Weighting

To adaptively calibrate the contributions of different pyramid levels, the CFW branch generates level-specific layer weights β. Crucially, to maintain the integrity of the multi-scale context, these layer weights are derived directly from the original input features $F_1, \ldots, F_N$ rather than the IFW-filtered outputs. Each original feature F_i is spatially pooled into a global descriptor z_i, which are then stacked to form the aggregate representation Z. The CACW module processes Z, and a Softmax function is applied to predict the normalized level-wise weights $\beta \in R^{N\times 1}$:

$$\beta = \text{Softmax}(\text{CACW}(Z)) \quad (2)$$

This step automatically calculates the exact proportion of micro-details versus broader wood context needed for the current local region.

(4) Fusion

To balance accuracy and efficiency, two adjacent pyramid levels (N=2) are selected for fusion. The final fused feature is computed as:

$$F_{out} = \sum_{i=1}^{N} \beta_i \hat{F}_i \quad (3)$$

Unlike conventional single-stage attention mechanisms, ADWM explicitly decouples intra-feature channel selection (IFW) and cross-feature level calibration (CFW). This dual-layer formulation avoids the premature loss of cross-scale semantics and achieves a more stable, instance-aware representation under complex texture interference.

Adaptive spatial feature fusion (*ASFF*)

In pyramid-based one-stage detectors, feature maps at different pyramid levels are supervised by scale-specific targets. Consequently, a single spatial location may exhibit inconsistent semantics, acting as foreground at one level and background at another. Naïve fusion strategies (*e.g.* summation or concatenation) mix these cross-scale conflicts, thereby degrading feature representations. This issue is particularly pronounced in texture-heavy industrial surface inspection, where strong background patterns and subtle defect regions coexist.

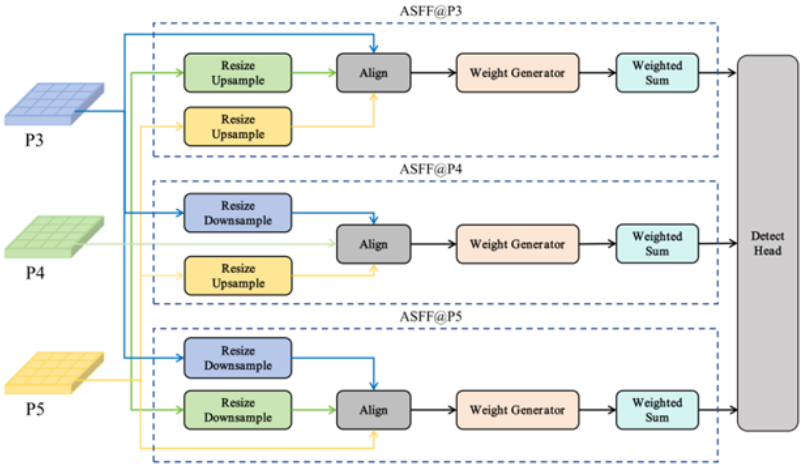

Fig. 5. Illustration of the Adaptive Spatial Feature Fusion (ASFF) module

To address this issue, the present version adopted the Adaptive Spatial Feature Fusion (ASFF) head as a lightweight module inserted immediately before the detection

layers, enabling spatially adaptive fusion across pyramid levels. As shown in Fig. 5, rather than enforcing uniform fusion, ASFF explicitly models the relative importance of different feature levels at each spatial location.

Specifically, given multi-level features from the neck (P3, P4, and P5), ASFF performs a sequence of distinct operations. First (Spatial Alignment), it aligns features from different pyramid levels to a target resolution *via* simple upsampling or downsampling, applying lightweight 1×1 channel alignment when required. Second (Weight Generation): ASFF predicts spatially varying fusion weights across pyramid levels, which are normalized at each spatial location. Third (Weighted Fusion): these weights combine the aligned features pixel-by-pixel. This sequence allows the detector to locally emphasize the most informative scale while suppressing misleading or contradictory responses from other levels.

During implementation, three ASFF units execute parallel fusion processes. Rather than generating a single mixed output, this yields three separate fused feature maps (denoted as P3', P4', and P5') at their respective resolutions. These distinct features are subsequently fed to the detection head in the original pyramid order, keeping the original classification and regression branches unchanged. By enabling per-location scale selection with minimal computational overhead, ASFF mathematically prevents micro-defects from being blurred by the surrounding wood background. This improves cross-scale semantic consistency and robustness to texture interference, making it highly suited for detecting small-scale and low-contrast defects in real-time industrial scenarios.

Normalized Wasserstein Distance (NWD)

In particleboard inspection, small-scale defects occupy a few pixels and exhibit high intra-class variations. Amidst texture-dominant surfaces, illumination fluctuations, and sensor noise, the overlap between predicted and ground-truth boxes becomes unstable, yielding unreliable regression cues. Consequently, standard IoU-family losses degrade; for tiny targets, minor localization noise disproportionately skews dimensional penalties, over-punishing near-accurate predictions.

To improve robustness, the present work incorporated the Normalized Wasserstein Distance (NWD) as an auxiliary regression signal. Instead of relying solely on overlap, NWD measures the discrepancy between boxes, providing smoother supervision when overlap is small or unstable. Following prior work, CIoU and NWD were combined using a balancing factor μ to form the final box regression loss.

$$L_{box} = (1-\mu)L_{CIoU} + \mu L_{NWD} \quad (4)$$

The NWD loss is defined as,

$$L_{NWD} = 1 - \exp(-\widetilde{W}_2) \quad (5)$$

where W_2 denotes the normalized Wasserstein-2 distance between the predicted box $b_p = (x_p, y_p, w_p, h_p)$ and the ground-truth box $b_g = (x_g, y_g, w_g, h_g)$:

$$W_2^2(b_p, b_g) = (x_p - x_g)^2 + (y_p - y_g)^2 + \left(\frac{w_p - w_g}{2}\right)^2 + \left(\frac{h_p - h_g}{2}\right)^2 \quad (6)$$

$$\widetilde{W}_2 = \frac{\sqrt{W_2^2}}{C} \quad (7)$$

where C is a normalization constant related to the image size (or diagonal length) to reduce scale sensitivity, and μ is fixed to μ=0.4 unless otherwise specified.

Experimental Details

All experiments were conducted on a Windows 10 workstation with an Intel Core i5-13600KF CPU, 32 GB RAM, and a single NVIDIA GeForce RTX 4060 Ti GPU (8 GB VRAM). Models were implemented in Python 3.10.19 using PyTorch 2.5.1 with CUDA 12.1. All compared models were initialized with official COCO-pretrained weights and fine-tuned on the self-collected defect dataset under identical training protocols to ensure fair comparison. COCO pretraining provided generic visual representations that stabilize convergence and is widely adopted in industrial defect detection. For newly introduced modules (*e.g.*, ADWM, ASFF) that were not present in the pretrained checkpoint, parameters were randomly initialized; all compatible layers inherited pretrained weights.

The input image size was set to 640 × 640. Training was performed for 200 epochs with a batch size of 8, using SGD with an initial learning rate of 0.01, momentum of 0.937, and weight decay of 5×10^{-4}. To stabilize the early stage of training, a warmup strategy was employed for the first 3 epochs, during which the warmup momentum was initialized at 0.8 and the bias learning rate at 0.0. Automatic mixed precision (AMP) was enabled. To ensure fair comparisons and reproducibility across variants, the random seed was set to 0 for all experiments, ensuring deterministic behavior in weight initialization, data shuffling, and data augmentation operations.

Online augmentations were applied during training to improve robustness, including HSV perturbation, random translation and scaling, horizontal flipping, Mosaic, RandAugment, and random erasing. Notably, to allow the model to better adapt to the true data distribution, Mosaic was explicitly disabled during the final 10 epochs. Bounding boxes were updated accordingly for geometric transforms. No augmentation was applied to validation or test sets. During the evaluation phase, Non-Maximum Suppression (NMS) was utilized with a strict IoU threshold of 0.7 to compute the precision metrics accurately.

Evaluation metrics

To evaluate detection effectiveness and computational efficiency, Precision (P), Recall (R), F1-score, and mean Average Precision (mAP) were determined. Efficiency-related indicators, including model size, number of parameters, and GFLOPs, were reported to reflect deployment cost and computational complexity. Precision and Recall are defined as follows,

$$\text{Precision (P)} = \frac{\text{TP}}{\text{TP+FP}} \tag{8}$$

$$\text{Recall (R)} = \frac{\text{TP}}{\text{TP+FN}} \tag{9}$$

where TP, FP, and FN denote true positives, false positives, and false negatives, respectively. The F1-score is the harmonic mean of Precision and Recall.

$$\text{F1} = \frac{2 \cdot \text{P} \cdot \text{R}}{\text{P+R}} \tag{10}$$

Average Precision (AP) and mean Average Precision (mAP) were used for overall detection performance. AP for each class is computed as the area under its precision-recall (PR) curve, and mAP is obtained by averaging AP across all C classes.

$$\text{mAP} = \frac{1}{\text{C}} \sum_{\text{i}=1}^{\text{C}} \text{AP}_i \tag{11}$$

where C is the total number of defect categories, and AP_i denotes the Average Precision for the i-th category. Following common practice in object detection, mAP@0.5 (IoU =

0.5) and mAP@0.5:0.95 (averaged over IoU thresholds from 0.5 to 0.95 in steps of 0.05) were reported to assess performance under both lenient and strict localization criteria.

Efficiency metrics (model size, parameters, and GFLOPs) are reported to characterize inference complexity and deployment overhead. GFLOPs for convolutional layers are computed as,

$$\text{GFLOPs} = \frac{H \times W \times C_{in} \times C_{out} \times k \times k}{10^9} \quad (12)$$

where H and W denote the spatial dimensions of the feature map, C_{in} and C_{out} represent the number of input and output channels, and k indicates the convolution kernel size. These metrics provide a comprehensive assessment of both accuracy and computational cost.

RESULTS AND DISCUSSION

Baseline Selection and Comparison

To establish a solid and reproducible baseline for particleboard surface defect detection, this project benchmarked from three commonly adopted paradigms: one-stage CNN-based detectors, two-stage detectors, and transformer-based detectors, under an identical experimental protocol. Specifically, an evaluation was made of a series of real-time YOLO-family one-stage models (YOLOv5n, YOLOv8n, YOLOv9n, YOLOv10n, YOLOv11n, YOLOv12n, and YOLOv13n) to cover recent architectural evolutions, alongside Faster R-CNN (two-stage) and RT-DETR (transformer-based) as complementary baselines.

For fairness, all models were trained and evaluated using the same data split, input resolution, optimizer configuration, training schedule, and evaluation protocol. Performance was assessed in terms of detection accuracy (Precision, Recall, F1-score, mAP@0.5, and mAP@0.5:0.95) and computational efficiency (model size, parameter count, and GFLOPs; FPS is reported when supported by the original implementation). Quantitative results are summarized in Table 2.

Table 2. Performance Comparison of Representative Object Detection Models on the Particleboard Surface Defect Dataset

Model	P (%)	R (%)	F1 (%)	mAP@50 (%)	mAP@50:95 (%)	GFLOPs (G)	Parameters (M)
YOLOV5n	87.3	76.4	81.4	87.2	51.8	7.2	2.5
YOLOV8n	89.2	81.0	84.9	87.9	56.8	8.2	3.0
YOLOV9n	86.9	81.2	83.9	87.1	55.0	7.9	2.0
YOLOV10n	79.7	76.3	78.0	81.2	50.7	8.4	2.7
YOLOV11n	89.4	78.9	83.8	86.2	52.7	6.5	2.6
YOLOV12n	83.8	80.3	82.0	85.0	52.8	6.5	2.6
YOLOV13n	88.7	76.2	82.0	84.6	50.8	6.3	2.4
Faster R-CNN	84.2	78.3	81.1	84.3	50.7	28.5	19.7
RT-DETR	89.1	79.6	84.0	85.7	55.3	73.64	38.72

Table 2 summarizes the performance of representative object detectors evaluated under an identical experimental protocol. Overall, one-stage YOLO-style detectors demonstrated a more favorable balance between detection accuracy and computational efficiency than the two-stage (Faster R-CNN) and transformer-based (RT-DETR)

baselines. For example, YOLOv8n achieved 56.8% mAP@0.5:0.95 with 8.2 GFLOPs, whereas Faster R-CNN attained a lower mAP@0.5:0.95 of 50.7% at a substantially higher cost of 28.5 GFLOPs, and RT-DETR required 73.64 GFLOPs to reach a comparable mAP@0.5:0.95 of 55.3%. These results highlight the efficiency advantage of one-stage detectors for real-time industrial inspection.

Within the YOLO family, different variants exhibited clear trade-offs between accuracy and efficiency. Models with lower computational costs tended to reduce GFLOPs but often suffered from degraded localization accuracy under stricter IoU criteria. For instance, YOLOv11n and YOLOv13n reduced GFLOPs to 6.5G and 6.3G, respectively, but their mAP@0.5:0.95 drops to 52.7% and 50.8%, compared with the 56.8% achieved by YOLOv8n.

Among all evaluated candidates, YOLOv8n provided the most balanced performance. It delivered consistently strong Precision, Recall, and F1-score, alongside the highest mAP@0.5:0.95 among lightweight YOLO variants, while maintaining moderate model complexity. This balance is particularly important for particleboard surface defect detection, where small-scale defects and texture-dominant backgrounds require both robust feature representation and efficient inference.

In contrast, Faster R-CNN and RT-DETR incurred substantially higher computational overhead without exhibiting a consistent accuracy advantage under the evaluated settings. Based on these observations, YOLOv8n as the baseline for subsequent experiments, ensuring that later performance gains can be attributed to the proposed modules rather than increased backbone capacity.

Module Comparative Experiments

To evaluate various feature extraction, attention, and fusion designs, Table 3 compares single-module integrations on the YOLOv8n baseline. For fairness, all variants were trained under an identical experimental protocol (*e.g.,* data split, resolution, training schedule) while other network structures and hyperparameters remain unchanged. This controlled setup isolates each module's impact.

Table 3. Component-wise Comparison of Feature Extraction, Attention, and Fusion Modules Integrated into YOLOv8n Baseline

Model	P (%)	R (%)	F1 (%)	mAP@50 (%)	mAP@50:95 (%)	GFLOPs (G)	Parameters (M)
Baseline	89.2	81.0	84.9	87.9	56.8	8.2	3.0
MEUM	90.4	82.9	86.5	88.5	57.4	8.2	3.0
PATConv	90.6	80.9	85.4	89.3	58.1	4.1	3.1
RCSSC	91.1	83.6	87.2	89.3	57.2	6.3	6.6
DySnakeConv	89.3	84.4	86.8	89.1	58.0	4.1	3.2
LKLGL	90.6	84.4	87.3	89.1	58.1	5.1	5.2
RHDWT	88.5	84.6	86.5	89.1	58.5	5.5	5.2
LRSA	90.5	78.1	84.0	86.3	53.0	4.2	3.2
ECA	90.5	82.0	86.0	88.4	57.2	4.0	3.0
PSConv	91.2	81.5	86.1	89.7	58.3	4.2	3.2
HLKConv	87.4	81.0	84.1	87.8	53.0	6.1	2.9
CBAM	90.3	83.2	86.6	89.6	58.1	4.1	3.1

ADWM	92.1	83.0	87.7	90.8	59.1	4.0	3.0
ASFFhead	92.6	84.9	88.5	90.2	58.6	5.1	4.3
BiFPN	91.4	83.5	87.2	89.4	58.7	4.0	2.7
RFAHead	91.3	82.3	86.6	89.1	57.0	5.6	4.6

Table 3 compares representative feature extraction, attention, and fusion modules when integrated into the YOLOv8n baseline under identical training and evaluation settings.

Among attention mechanisms, ECA and CBAM delivered modest improvements over the baseline, with mAP@0.5:0.95 reaching 57.2% and 58.1% respectively, while LRSA showed reduced performance under strict localization criteria at 53.0% mAP@0.5:0.95, suggesting that uniform channel reweighting may amplify texture-dominated responses rather than suppressing them. Convolutional variants present varying accuracy–efficiency trade-offs: PSConv improves localization with moderate overhead, whereas RCSSC, LKLGL, and RHDWT achieve higher accuracy at the cost of increased parameters.

The proposed ADWM achieved the highest detection performance (90.8% mAP@0.5 and 59.1% mAP@0.5:0.95) while maintaining competitive efficiency (3.0M parameters, 4.0 GFLOPs). This demonstrates that explicitly decoupling intra-feature channel selection and cross-feature level calibration can provide more effective suppression of texture interference than single-stage attention or convolution-based alternatives.

Among head-level fusion modules, ASFFHead attained the highest Recall (84.9%) and F1-score (88.5%), but it introduced substantial overhead (4.3M parameters), while BiFPN offered comparable accuracy with lower complexity. Overall, ADWM provided the most favorable accuracy-efficiency balance, making it well-suited for resource-constrained industrial deployment.

Ablation Studies

To quantify the contribution of each proposed component, ablation studies were conducted on the YOLOv8n baseline. As summarized in Table 4, ADWM, ASFF, and NWD were introduced individually or in combination under identical training and evaluation settings. This controlled setup isolated individual performance gains and facilitates a systematic assessment of their complementary effects on accuracy and computational efficiency.

Table 4. Ablation Study of ADWM, ASFF, and NWD Integrated into the YOLOv8n Baseline

ADWM	ASFF	NWD	P (%)	R (%)	F1 (%)	mAP@50 (%)	mAP@50:95 (%)	GFLOPs (G)	Params (M)
×	×	×	89.2	81.0	84.9	87.9	56.8	8.2	3.0
√	×	×	92.1	83.0	87.7	90.8	59.1	4.0	3.0
×	√	×	92.6	84.9	88.5	90.2	58.6	5.1	4.3
×	×	√	91.7	84.5	87.9	89.9	58.6	4.0	3.0
√	√	×	93.2	82.1	87.3	89.3	58.1	3.6	4.9

√	×	√	93.1	84.2	88.4	90.1	58.7	3.2	4.3
×	√	√	92.4	84.6	88.3	90.4	59.1	4.2	3.6
√	√	√	93.4	85.7	89.3	90.9	59.3	4.7	3.5

Table 4 reports an ablation study of ADWM, ASFF, and NWD on the YOLOv8n baseline under identical training and evaluation settings, so that the observed performance differences primarily reflect the contribution of each component.

Enabling ADWM alone yielded the largest single-module gain, increasing mAP@0.5 from 87.9% to 90.8% and mAP@0.5:0.95 from 56.8% to 59.1%. By replacing the standard concatenation-based feature fusion with an adaptive dual-weighting mechanism, ADWM maintained 3.0M parameters while reducing computational cost to 4.0 GFLOPs, compared to 8.2 GFLOPs for the baseline. By comparison, enabling ASFF alone mainly improved sensitivity, with Recall rising from 81.0% to 84.9% and the F1-score from 84.9% to 88.5%, indicating fewer missed detections due to enhanced cross-scale interaction. Incorporating NWD alone provided consistent localization benefits, with Recall reaching 84.5% and mAP@0.5:0.95 improving to 58.6%, supporting its role as a more stable regression signal for small-scale defects.

When components were combined, their synergistic effects become evident across the feature representation, fusion, and supervision stages. Specifically, ADWM first filters out chaotic wood-grain texture to provide clean, instance-aware features. ASFF then takes these refined features and resolves cross-scale spatial conflicts without blending away micro-defects. Crucially, the full potential of these spatially aligned, texture-free features is only unlocked when NWD is integrated. Without NWD, standard IoU-based losses struggle to optimize the fine-grained small-target features generated by ADWM and ASFF; NWD provides the stable localization gradients necessary to complete this pipeline. This three-way synergy culminated in the best overall performance being achieved when all three components were jointly enabled, where mAP@0.5 and mAP@0.5:0.95 reached 90.9% and 59.3%, respectively, together with the highest Recall of 85.7% and F1-score of 89.3%, at a moderate cost of 4.7 GFLOPs and 3.5M parameters. Overall, these results demonstrate that ADWM-enhanced feature representation, ASFF-based multi-scale interaction, and NWD-based localization supervision are strictly mutually reinforcing for robust particleboard surface defect detection.

Visual Analytics

To complement the quantitative evaluation, Fig. 6 visualizes response heatmaps for representative defect categories, with particular emphasis on small-scale defects, low-contrast defects, and easily confused cases. Compared with the baseline, the proposed method exhibited stronger activation on defect regions of interest and reduced spurious responses on texture-dominant backgrounds, indicating improved localization focus under challenging conditions. These qualitative observations are in line with the quantitative improvements reported in Tables1-3, especially under stricter localization criteria.

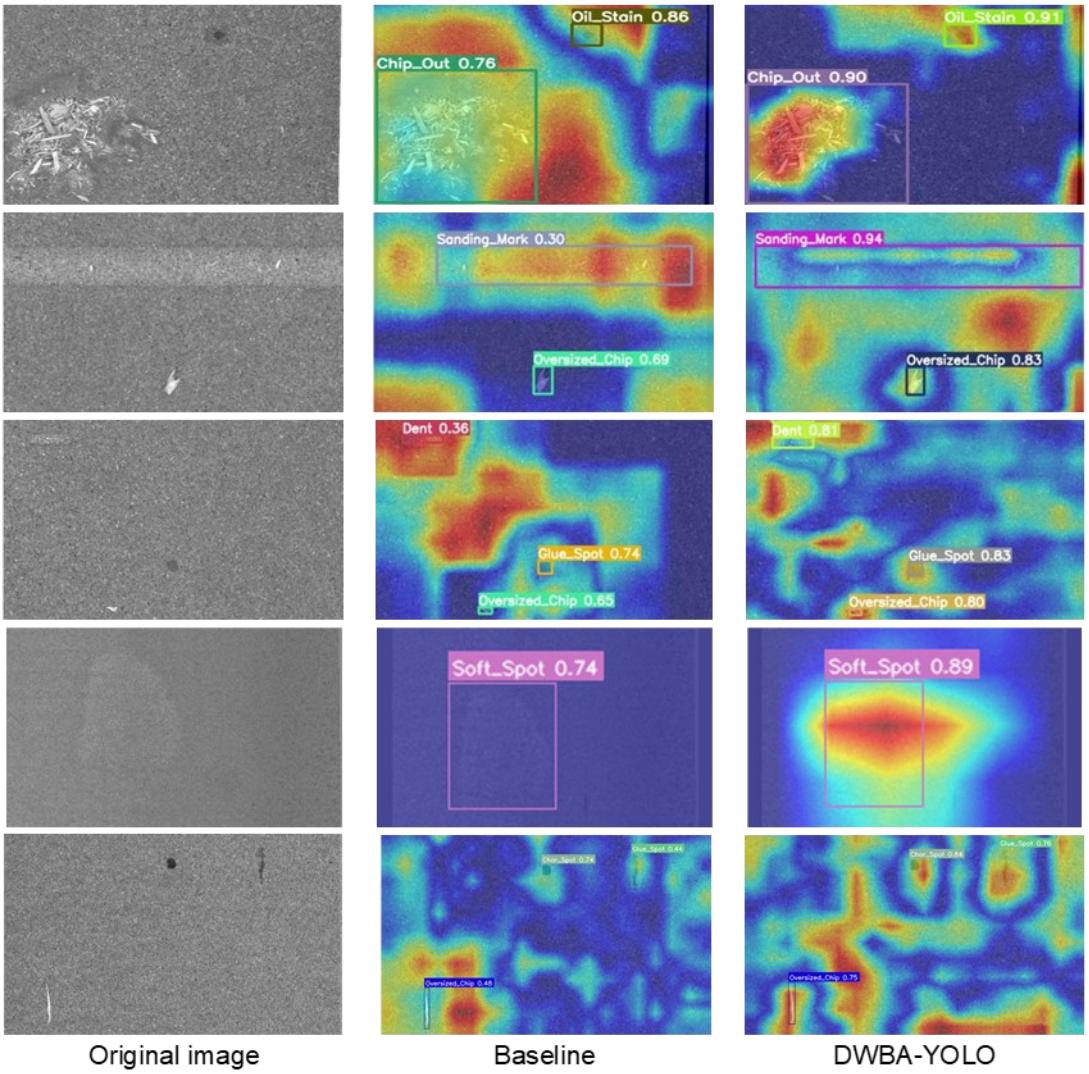

Fig. 6. Response heatmap visualization of representative surface defect samples

The heatmaps visualize the spatial distribution of activation intensity, indicating the regions emphasized during defect detection. Compared with the baseline, the proposed method produces more concentrated responses around small-scale defects and low-contrast defect regions while reducing spurious activations on texture-dominant backgrounds.

CONCLUSIONS

1. This paper presents a lightweight and effective defect detection framework for particleboard surfaces addressing small-scale defects, low-contrast defects, and texture-dominated backgrounds. Built on a YOLOv8n baseline, three complementary designs were integrated to improve accuracy without substantially increasing deployment cost.
2. An Adaptive Dual-layer Weighting Mechanism (ADWM) was used to enhance feature discrimination by suppressing texture-dominated responses and adaptively balancing

information across pyramid levels, while ASFF performs spatially adaptive multi-scale fusion to alleviate cross-scale semantic conflicts and reduce missed detections.

3. The Normalized Wasserstein Distance (NWD) was introduced as an auxiliary localization supervision signal to provide smoother and more reliable regression for small-scale defects when overlap-based losses become unstable.
4. Extensive experiments, including module-wise comparisons and ablation studies, verified that each component contributed consistent gains, and their joint integration achieved the best overall performance with a favorable accuracy-efficiency trade-off. These results indicate the proposed method is well suited for real-time industrial surface inspection and can be extended to other textured materials.

Despite these promising results, several limitations remain. First, because the model was trained under controlled industrial lighting, its robustness to natural illumination and generalization to other wood products (*e.g.*, MDF) require validation. Second, while designed to be lightweight, latency benchmarks on edge devices (*e.g.*, Jetson Nano) are pending. Finally, extreme minority classes (*e.g.*, Oil Stain) still exhibit lower accuracy. To address these issues, future work will focus on domain adaptation for cross-material generalization, model compression for sub-2 GFLOPs edge deployment, advanced re-balancing algorithms for minority classes, and multi-modal thermal fusion for moisture-related defects.

ACKNOWLEDGMENTS

The authors are grateful for the financial support from the Inner Mongolia Natural Science Foundation for the project "Construction of an Intelligent Inspection Model for Surface Defects of Particleboard Made from Psammophytic Shrubs" (Project No. 2024LHMS03063).

REFERENCES CITED

AlBahar, A., Kim, I., and Yue, X. (2022). "A robust asymmetric kernel function for Bayesian optimization, with application to image defect detection in manufacturing systems," *IEEE Transactions on Automation Science and Engineering* 19(4), 3222-3233. https://doi.org/10.1109/TASE.2021.3114157

He, C., Kang, Y., Ding, A., Jia, W., and Duo, H. (2025). "LE-YOLO: A lightweight and enhanced algorithm for detecting surface defects on particleboard," *BioResources* 20(3), 7179-7193. https://doi.org/10.15376/biores.20.3.7179-7193

Hoang, V. T., Le, V. T., Dinh, N., Tran-Trung, K., Van, B. N., Hong, H. D. T., and Huong, T. H. (2025). "Deep learning-based faulty wood detection with area attention," *Computers, Materials & Continua* 85(1), 1495-1514. https://doi.org/10.32604/cmc.2025.066506

Huang, J., Chen, H., Ren, J., Peng, S., and Deng, L. (2025). "A general adaptive dual-level weighting mechanism for remote sensing pansharpening," in: *Proceedings of the Computer Vision and Pattern Recognition Conference*, article ID 00698. https://doi.org/10.1109/CVPR52734.2025.00698

Jocher, G., Chaurasia, A., and Qiu, J. (2023). Ultralytics YOLOv8, GitHub repository, (https://github.com/ultralytics/ultralytics), Accessed January 27, 2026.

Liang, X., Li, Y., Wang, X., Liu, P., Shen, Y., and Guo, J. (2025). "Adaptive shape imitation and selective semantic guidance for industrial surface defect detection," *Expert Systems with Applications* 279, article ID 127334. https://doi.org/10.1016/j.eswa.2025.127334

Liu, J., Shen, X., Wang, J., Jiang, L., and Zhang, H. (2023). "An intelligent defect detection approach based on cascade attention network under complex magnetic flux leakage signals," *IEEE Transactions on Industrial Electronics* 70(7), 7417-7427. https://doi.org/10.1109/TIE.2022.3201320

Liu, S., Huang, D., and Wang, Y. (2019). "Learning spatial fusion for single-shot object detection," arXiv preprint, article ID 1911.09516. https://doi.org/10.48550/arXiv.1911.09516

Liu, T., He, Z., Lin, Z., Cao, G.-Z., Su, W., and Xie, S. (2024). "An adaptive image segmentation network for surface defect detection," *IEEE Transactions on Neural Networks and Learning Systems* 35(6), 8510-8523. https://doi.org/10.1109/TNNLS.2022.3230426

Ma, R., Chen, J., Feng, Y., Zhou, Z., and Xie, J. (2025). "ELA-YOLO: An efficient method with linear attention for steel surface defect detection during manufacturing," *Advanced Engineering Informatics* 65, article ID 103377. https://doi.org/10.1016/j.aei.2025.103377

Redmon, J., Divvala, S., Girshick, R., and Farhadi, A. (2016). "You only look once: Unified, real-time object detection," in: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, Las Vegas, NV, USA, article ID 10.1109. https://doi.org/10.1109/CVPR.2016.91

Wan, F., Zhang, G., and Li, Z. (2025). "Dfsdnet: A dual-branch multi-scale feature fusion network for surface defect detection of copper strips and plates," *Computers in Industry* 167, article ID 104265. https://doi.org/10.1016/j.compind.2025.104265

Wang, X. (2025). "Research on wood defects feature imbalance optimization and recognition," *IEEE Access* 13, 23841-23850. https://doi.org/10.1109/access.2025.3538285

Yan, J., Yang, C., Yan, Q., Zhang, T., and Qu, W. (2025). "Multi-data fusion approach for surface defect detection and quality grading of particleboards," *Expert Systems with Applications* 293, article ID 128706. https://doi.org/10.1016/j.eswa.2025.128706

Zhao, Z., Ge, Z., Jia, M., Yang, X., Ding, R., and Zhou, Y. (2022). "A particleboard surface defect detection method research based on the deep learning algorithm," *Sensors* 22(20), 7733. https://doi.org/10.3390/s22207733

Zhao, Z., Yang, X., Zhou, Y., Sun, Q., Ge, Z., and Liu, D. (2021). "Real-time detection of particleboard surface defects based on improved YOLOV5 target detection," *Scientific Reports* 11, article ID 21777. https://doi.org/10.1038/s41598-021-01084-x

Zheng, Y., Wang, M., Zhang, B., Shi, X., and Chang, Q. (2024). "GBCD-YOLO: A high-precision and real-time lightweight model for wood defect detection," *IEEE Access* 12, 12853-12868. https://doi.org/10.1109/ACCESS.2024.3356048

Article submitted: January 18, 2026; Peer review completed: February 7, 2026; Revised version received and accepted: February 26, 2026; Published: March 11, 2026.
DOI: 10.15376/biores.21.2.3856-3871