ISSN: 2472-3703 DOI: 10.6911/WSRJ.202510_11(10).0008

Research Progress of Deep Learning-Based Wood Surface Defect Detection

Yuan Zhang[1, a], Xuewen Ding[1, 2, b, *]

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

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

[a]zy_514@163.com, [b]dingxw@tute.edu.cn

* Corresponding Author

Abstract

The detection of surface defects in wood is crucial for enhancing product value and manufacturing efficiency. Traditional methods exhibit inherent limitations in both efficiency and robustness, whereas deep learning techniques offer innovative solutions to these challenges. This paper provides a comprehensive review of the research progress in this field, beginning with an overview of the evolution from convolutional neural networks (CNNs) to object detection and image segmentation models, as well as their adaptability in wood defect identification. It then focuses on the role of key optimization strategies, such as attention mechanisms and multi-scale feature fusion, in improving detection performance under complex textures. Finally, this paper highlights future directions, including the integration of multimodal information, the development of self-supervised learning, and the construction of end-to-end systems, all aimed at providing forward-looking technological references for the intelligent upgrading of the wood industry.

Keywords

Wood surface defects; deep learning; object detection; attention mechanism; multi-scale feature fusion.

1. INTRODUCTION

Wood, renowned for its renewability, biodegradability, and superior properties, is extensively utilized in the production of furniture, architectural decorations, bridges, and structural engineering applications. It serves as a vital green material in the pursuit of the "dual carbon" strategic objectives. Wood products offer an environmentally sustainable alternative to high-carbon materials such as steel and cement, positioning them as particularly promising within the construction sector and contributing to long-term carbon sequestration efforts. As global awareness of the role of wood in emission reduction expands, the wood industry—recognized as a low-carbon sector—has gained increasing recognition [1].However, during processing, transportation, and storage, wood frequently develops surface defects, including live knots, pith, and dead knots. These defects not only diminish the aesthetic and commercial value of the wood but also compromise its structural integrity, reduce its service life, and, in some cases, lead to its disposal. Consequently, the development of precise and efficient wood defect detection technologies is essential for advancing intelligent manufacturing and optimizing resource utilization.

ISSN: 2472-3703 DOI: 10.6911/WSRJ.202510_11(10).0008

Traditional detection methods primarily rely on manual visual inspection or algorithms based on basic image features. For instance, the HOG+SVM classification method identifies defects[2] by extracting edge and texture features, while threshold segmentation and grayscale analysis detect defect[3] areas through grayscale variations. However, these methods share several limitations: they are sensitive to lighting variations, struggle with complex textures, depend on the operator's expertise, and lack robustness when detecting small or multi-class defects. As industrial automation and intelligence demands increase, these traditional techniques are inadequate in meeting the modern production requirements for high precision, consistency, and real-time performance.

In recent years, deep learning techniques have substantially advanced wood defect detection. Convolutional neural networks(CNNs)[4] automatically extract multi-level features, facilitating end-to-end learning that enhances defect recognition accuracy and robustness. Faster R-CNN[5] improves defect localization accuracy through region proposal networks, while the YOLO series enables efficient real-time detection. Transformer-based models[6] enhance the modeling of global context information. Additionally, optimization strategies such as feature pyramids, multi-scale fusion, and attention mechanisms have further bolstered detection performance, particularly for small defects in complex textured backgrounds. Lightweight networks and transfer learning approaches have also improved the adaptability of models for deployment on edge devices and in scenarios with limited sample sizes. Table 1 summarizes the key distinctions between traditional and deep learning-based detection methods.

Table 1. Methods for Identifying Wood Defects in Images

Category	Method Name	Proposed Time	Advantages	Disadvantages
Traditional Image Recognition Methods	BP Neural Network	1986	Versatile, adaptable	Prone to overfitting, computationally intensive
	RBF Neural Network	1985	Effective for small datasets	Struggles with large datasets, kernel optimization challenging
	Support Vector Machine (SVM)	1968	Strong generalization, good for high-dimensional data	Needs careful kernel selection, expensive optimization
	Extreme Learning Machine (ELM)	2004	Fast and simple, suitable for large datasets	Requires large data, limited generalization
Convolutional Neural Network -based Recognition Methods	CNN	1980s-1990s	Superior feature extraction, automatic learning	High computational demand
	R-CNN	2014	Combines CNN and RPN for precise detection	Complex preprocessing, resource-intensive
	Faster R-CNN	2015	Efficient and accurate with RPN integration	Requires labeled data, may not generalize well
	SSD	2016	Balances accuracy and real-time speed	Struggles with generalization in complex datasets

ISSN: 2472-3703 DOI: 10.6911/WSRJ.202510_11(10).0008

Despite significant progress, wood defect detection still faces numerous challenges, including imbalanced defect sample distribution, complex natural textures, significant interference from lighting and background, and the stringent requirements of real-time performance and lightweight solutions in industrial applications. This paper provides a comprehensive review of deep learning-based methods for wood defect detection, analyzing them from three dimensions: methodological framework, optimization strategies, and practical applications. Furthermore, it identifies the research bottlenecks and development trends. The paper is organized as follows: Chapter 2 introduces deep learning-based surface defect detection methods and key optimization strategies; Chapter 3 discusses the current challenges and future research directions.

2. SURFACE DEFECT DETECTION METHODS BASED ON DEEP LEARNING

2.1. Traditional Methods and Limitations in Surface Defect Detection

Before the widespread application of deep learning technologies, wood surface defect detection primarily relied on two methods: manual visual inspection and traditional machine[7] vision algorithms. Manual visual inspection is typically performed by experienced inspectors under specific lighting conditions, who observe the wood surface to identify and assess defects. The effectiveness of this method is highly dependent on the inspector's expertise and physical condition. It is often characterized by inefficiency, high labor intensity, strong subjectivity, and susceptibility to fatigue, making it difficult to meet the modern production line's demands for efficiency and consistency.

Traditional machine vision methods[7] aim to reduce reliance on manual inspection through automation. The typical process includes image acquisition and preprocessing, feature engineering, and classifier design. In the feature engineering stage, researchers often rely on hand-crafted texture features (e.g., gray-level co-occurrence matrix, local binary pattern [LBP]), shape features (e.g., edges, contours, Hu moments), and statistical features (e.g., color moments, gradient histograms). These features are then input into classifiers such as support vector machines (SVM), decision trees, or k-nearest neighbors (K-NN) for defect detection.

Although these methods have shown some success in controlled experimental environments, their limitations remain significant. Hand-crafted features have limited expressiveness and struggle to effectively capture the complex and variable natural textures of wood. They are sensitive to lighting, noise, and viewing angle, with poor robustness. Furthermore, their ability to detect small defects (e.g., insect holes or microcracks) is limited, resulting in a high false-negative rate. Additionally, these methods are dependent on prior knowledge and manual design, leading to long development cycles and poor portability, making them unsuitable for different wood species or novel defects.

2.2. Fundamental Theories and Core Models of Deep Learning

2.2.1 Basic Principles of Convolutional Neural Networks (CNNs)

Convolutional Neural Networks (CNNs)[4] are the core model of deep learning in the field of image processing. They effectively handle high-dimensional image data by employing mechanisms such as local connections, weight sharing, and spatial downsampling, while also providing translation invariance. A typical CNN architecture consists of an input layer, convolutional layers, activation functions, pooling layers, and fully connected layers, as shown in Figure1. Convolutional layers extract local features by sliding a set of learnable kernels over the input image. Shallow networks capture basic features such as edges, corners, and textures, while deeper networks can capture more abstract and semantic features. Activation functions introduce non-linear transformations, enabling the network to model complex mappings. Among them, ReLU and its variants are widely used due to their ability to mitigate the vanishing

ISSN: 2472-3703 DOI: 10.6911/WSRJ.202510_11(10).0008

gradient problem. Pooling layers reduce the spatial resolution of feature maps, expand the receptive field, and enhance the robustness of features to positional variations. Fully connected layers map high-dimensional features to target classes, performing the final classification or regression tasks. Batch Normalization accelerates model training and provides a regularization effect, while Dropout mitigates overfitting by randomly deactivating neurons. CNNs can autonomously learn the intrinsic features of wood defects, overcoming the reliance on manual feature engineering typical of traditional methods.

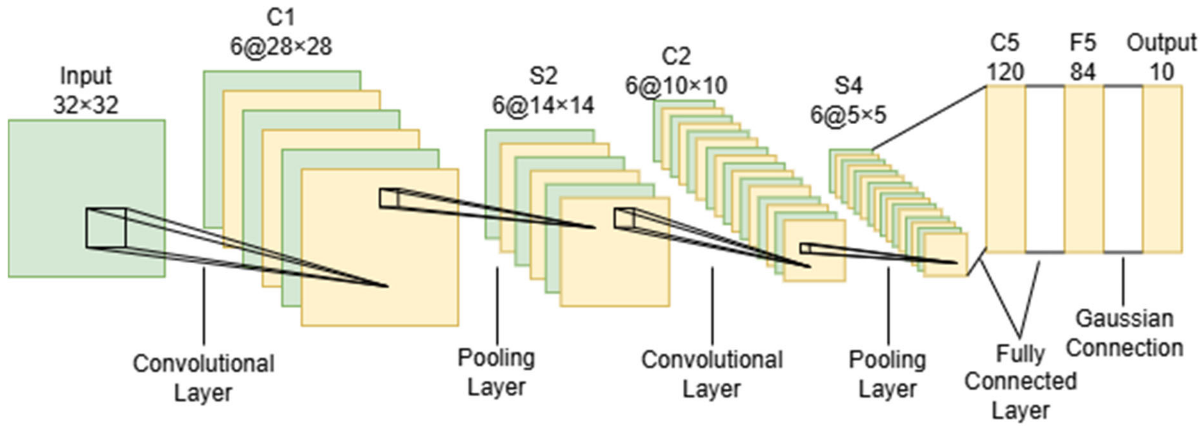


Figure 1. Schematic Representation of the CNN Architecture

2.2.2 Evolution of Mainstream Deep Learning Architectures

With the continuous advancement of deep learning theories, Convolutional Neural Network (CNN) architectures have evolved from shallow networks to deep networks, transitioning from general-purpose models to specialized architectures tailored for specific tasks. This evolution has significantly propelled the development of computer vision. In wood defect detection, these networks typically serve as the backbone for feature extraction, providing high-quality deep features that support subsequent classification or detection tasks.

VGGNet[8], by stacking multiple small convolution kernels (3×3) instead of using a single large kernel (e.g., 5×5 or 7×7), effectively controls the number of parameters while increasing the depth of the network. Its modular structure not only facilitates network design, transfer, and fine-tuning but also enables shallow convolutions to capture low-level features such as edges and textures, while deeper convolutions progressively extract higher-level semantic information. This structure enhances the model's ability to represent complex texture defects.

ResNet[9], by introducing residual modules, significantly alleviates the vanishing gradient problem in deep networks, making it possible to train networks with tens or even hundreds of layers, as illustrated in Figure 2. This powerful deep representation capability is crucial for learning and distinguishing defects with features similar to the natural wood texture (e.g., light cracks and wood grains).

When considering real-time industrial detection or deployment on mobile or embedded devices, lightweight networks (e.g., MobileNet[10] and ShuffleNet[11]) employ mechanisms like depthwise separable convolutions and channel shuffling to significantly reduce computational complexity while maintaining accuracy. These networks provide feasible solutions for deployment on embedded devices. Moreover, they not only reduce memory usage and latency but also integrate with model optimization techniques such as quantization and pruning, further accelerating computation. This makes them viable for large-scale wood defect detection and online quality control. Additionally, their multi-layer feature representation capabilities enable the simultaneous recognition of small defects and complex texture areas,

ISSN: 2472-3703 DOI: 10.6911/WSRJ.202510_11(10).0008

offering stable and reliable feature input for subsequent object detection and segmentation tasks.

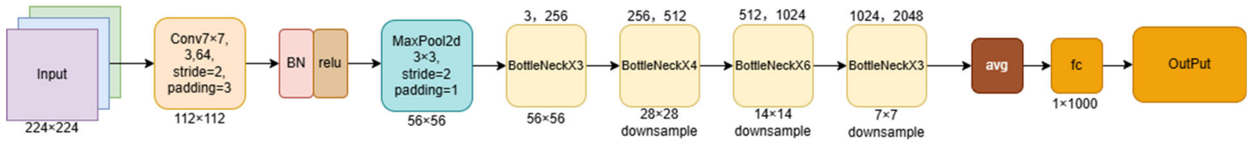


Figure 2. Schematic of the ResNet Network Architecture

2.2.3 From Classification to Detection: Object Detection Algorithm Framework

The task of object detection not only involves identifying the type of defects within an image but also requires precise localization of these defects in the spatial domain. In the field of deep learning, object detection algorithms are typically divided into two main categories: two-stage detectors and one-stage detectors.

Two-stage detectors, such as Faster R-CNN[5], are characterized by a core process that includes candidate region generation and region classification regression. Specifically, the Region Proposal Network (RPN) first generates numerous candidate boxes (anchors) on the input feature map, which are likely to contain the target. These redundant, highly overlapping candidate boxes are then filtered through non-maximum suppression (NMS). Subsequently, each candidate region is cropped and passed through the subsequent network for category prediction and bounding box regression optimization. In the context of wood defect detection, due to the complexity of defect shapes and significant size variations, researchers typically fine-tune the anchor box scales, ratios, and densities in the RPN to improve coverage of small defects (such as micro-cracks, insect holes, and localized decay). Additionally, the integration of a Feature Pyramid Network (FPN) allows the fusion of deep semantic information with shallow details, enabling the model to effectively respond to both large and small defects, thereby enhancing recall rates and localization accuracy across multiple scales.

In contrast, one-stage detectors, such as the YOLO series and SSD, simplify the object detection task by combining category prediction and bounding box regression into a unified model, eliminating the need for a candidate region generation step. This significantly boosts detection speed. In wood industry applications, YOLOv5[12] leverages adaptive anchor box computation, multi-scale training, random flipping, color perturbation, and an adaptive IoU loss function, which ensures high generalization ability in complex textured backgrounds and improves the accuracy of small-object detection. YOLOv8[13] incorporates an anchor-free design that directly predicts the center and dimensions of the target, thereby reducing dependence on prior anchor boxes. Combined with modules such as CSP and BiFPN, this design facilitates cross-scale feature fusion, preserving localization accuracy while improving inference speed. As a result, YOLOv8 is particularly effective for high-speed image processing in production lines and real-time wood defect detection, as demonstrated in Figure 3 Basic Structure of YOLOv8. This architecture enhances both localization and speed, making it highly suitable for industrial applications.

ISSN: 2472-3703 DOI: 10.6911/WSRJ.202510_11(10).0008

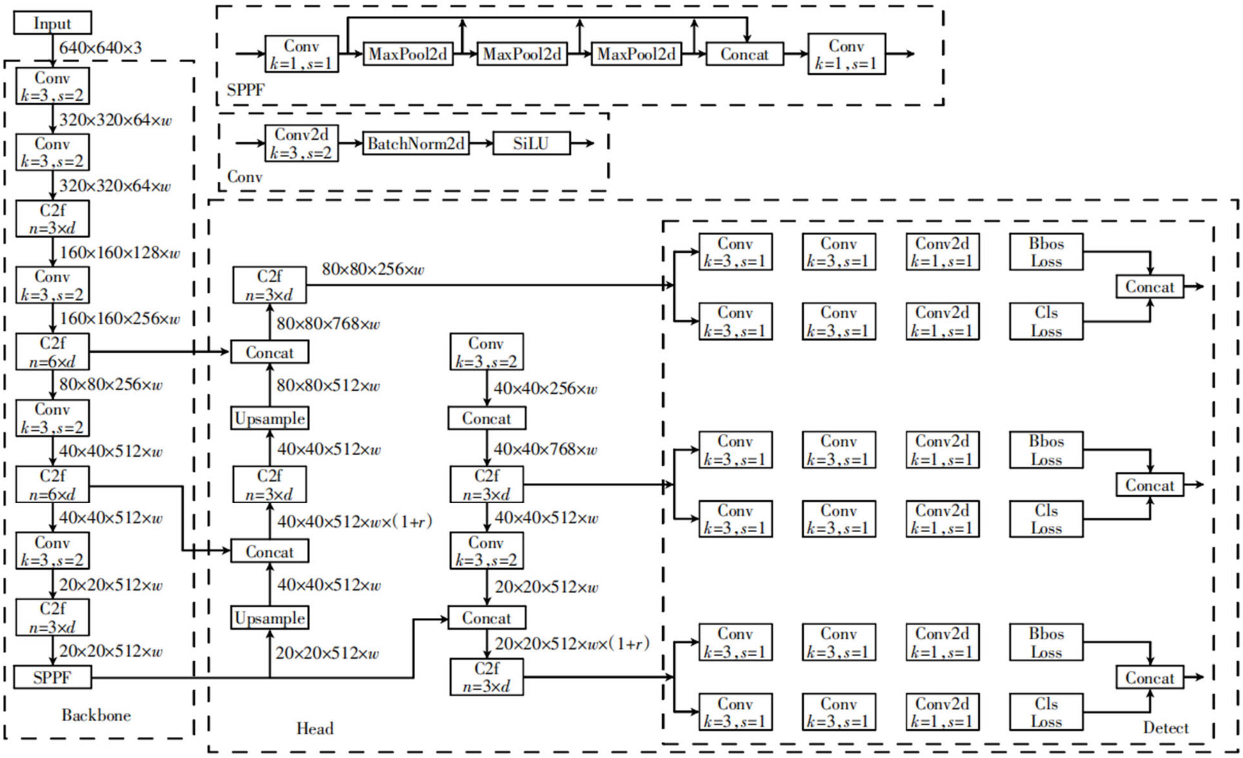


Figure 3. Basic Structure of YOLOv8

2.2.4 From Coarse to Fine: Semantic Segmentation and Instance Segmentation

In wood surface defect detection, many defects exhibit irregular shapes, complex boundaries, small sizes, and dense distributions. Traditional bounding box detection methods, which rely on approximate rectangular frames, struggle to capture the fine morphological details of defects accurately. This is particularly problematic when detecting small targets such as cracks, decay, discoloration, or insect holes, often resulting in missed or false detections. As a result, semantic segmentation [14] and instance segmentation [15] have become key techniques for improving detection accuracy, with a comparison of the two provided in Table 2.

Semantic segmentation aims to predict the category of each pixel in an image, labeling all pixels of the same class, but it cannot distinguish boundaries between objects of the same class. In wood defect detection, semantic segmentation enables pixel-level localization of defect regions by capturing defect contours and texture details, thereby enhancing the model's ability to perceive complex defect shapes. A typical network, such as U-Net, uses a symmetric encoder-decoder architecture. The encoder progressively extracts multi-scale features and compresses spatial information through multiple layers of convolution and pooling, while retaining high-level semantic representations. The decoder restores spatial resolution through upsampling and combines low-level details from the encoder using skip connections, achieving high-precision pixel-level predictions.

Instance segmentation builds upon semantic segmentation by further distinguishing individual objects within the same class, allowing for the independent segmentation of dense or adjacent defects. Mask R-CNN, a representative method for instance segmentation, extends the Faster R-CNN[5] framework by adding a mask prediction branch. It generates candidate bounding boxes through a region proposal network, then classifies, regresses boundaries, and predicts pixel-level masks for each candidate box to achieve fine-grained object-level segmentation. By combining multi-scale feature maps and RoIAlign operations, Mask R-CNN improves the model's response to defects of varying sizes, ensuring spatial accuracy for small and dense targets. This significantly alleviates the limitations of traditional bounding box

ISSN: 2472-3703 DOI: 10.6911/WSRJ.202510_11(10).0008

methods in segmentation accuracy and small target recognition, providing effective technical support for high-precision wood defect detection.

Table 2. Comparison of Common Segmentation Techniques for Wood Defect Detection

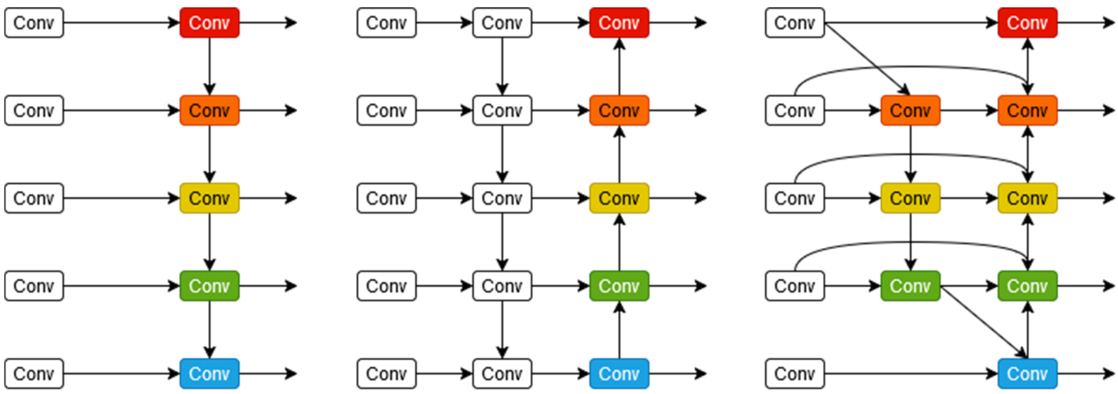
Segmentation Task Type	Related Models	Core Methodology	Advantages	Disadvantages
Semantic Segmentation	U-Net	The encoder extracts features and the decoder restores the resolution by upsampling, with high-level feature fusion achieved through skip connections.	High precision in localization, simple structure, effective for small samples.	Struggles to distinguish boundaries between similar objects and is sensitive to edge distortions.
Instance Segmentation	Mask R-CNN	Based on Faster R-CNN, Mask R-CNN adds a mask prediction branch and uses Region Proposal Networks (RPN) to generate candidate bounding boxes, followed by classification, bounding box regression, and pixel-level mask prediction.	Can segment instances of the same class, accurate for dense targets and small defects.	High computational load, requires large-scale and high-quality data, and is less effective on small target segmentation.

2.3. Key Optimization Strategies for Enhancing Detection Performance

Wood surface defect detection faces several challenges, including significant variations in target size, complex textures, strong background interference, and the high real-time performance demands of industrial environments. Defects typically occupy a small proportion of the image and are highly similar to the background texture, making it difficult for traditional convolutional networks to effectively extract features, which often leads to reduced accuracy in identifying small or low-contrast defects. To improve detection performance, researchers have proposed several strategies, including attention mechanisms, multi-scale feature fusion, model lightweighting, and data augmentation and transfer learning under small sample conditions.

The attention mechanism mimics human visual selective focus, enabling the network to assign higher weights to key regions and important features, thereby enhancing feature representation. Channel attention mechanisms, such as SENet, highlight high-information channels through global information aggregation and weight learning, while suppressing low-information channels. Spatial attention mechanisms generate spatial weight matrices to emphasize defect locations. Hybrid attention mechanisms, such as CBAM, combine channel and spatial attention to optimize feature representation across both dimensions. Coordinate attention further encodes horizontal and vertical position information, capturing long-range dependencies while preserving precise location information, making it particularly effective for detecting small, sparse defects (e.g., knots, insect holes, and cracks), and can be integrated with lightweight or efficient detection models.

Wood surface defects exhibit a large size variation, making it difficult for single-scale feature maps to effectively capture both small targets and large defects. Feature Pyramid Networks (FPN), through top-down paths and lateral connections, fuse deep high-semantic features with shallow high-resolution features, preserving both semantic information and spatial details. Shallow features aid in capturing the textures of small defects, while deep features provide semantic understanding of larger defects. Bidirectional Feature Pyramid Networks (BiFPN) build on this by introducing cross-scale weighted fusion, enabling bidirectional information flow and feature aggregation through learnable weights, which enhances the discernibility of small targets while suppressing shallow noise interference. The network structure of the feature fusion module is illustrated in Figure 4.



(a) FPN (b) PANet (c) BiFPN

Figure 4. Network Structure of Feature Fusion Modules

In industrial applications,deploymen is often constrained by computational resources. As a result, model compactness and inference speed become critical factors. Lightweight networks reduce computational complexity effectively through strategies such as replacing the backbone network, employing depthwise separable convolutions, and using channel shuffling. Network pruning removes redundant connections, quantization lowers the precision of weights and activations, and knowledge distillation trains the student model with guidance from a teacher model. These techniques enable lightweight models to maintain compact size while approximating the original performance, facilitating efficient real-time deployment.

Moreover, deep learning methods heavily rely on large-scale data, while the number of wood defect samples is limited. Solutions include traditional data augmentation techniques (e.g., rotation, flipping, scaling, cropping, color jittering, noise addition), generating defect images using Generative Adversarial Networks (GANs), and employing transfer learning. Transfer learning transfers features learned from pre-trained models on large datasets (e.g., ImageNet) to the wood defect detection task, followed by fine-tuning to achieve high-performance training under small sample conditions. The integration of these approaches significantly enhances the accuracy, robustness, and industrial applicability of wood surface defect detection.

2.4. Summary

This chapter provides a comprehensive review of the theoretical framework for surface defect detection based on deep learning. First, it discusses the limitations of traditional methods and emphasizes the necessity of incorporating deep learning. Next, it presents the foundational principles of convolutional neural networks (CNNs) and their applications in wood defect detection, including object detection, semantic segmentation, and instance segmentation models. Finally, key techniques for enhancing model performance are summarized, such as attention mechanisms, multi-scale feature fusion, model lightweighting, and optimization

ISSN: 2472-3703 DOI: 10.6911/WSRJ.202510_11(10).0008

strategies for scenarios with limited data. This theoretical framework provides a solid foundation for subsequent research, which aims to address practical challenges such as complex texture backgrounds, small target defects, and real-time detection.

3. CHALLENGES AND PROSPECTS

Despite the significant progress made by deep learning in the detection of wood surface defects, several challenges remain in achieving industrial-scale application. This chapter provides a systematic analysis of the key issues in current research and explores future development trends.

3.1. Major Challenges in Current Wood Defect Detection

Although deep learning has achieved notable progress in the detection of wood surface defects, it still faces multiple challenges in industrial applications, hindering its widespread implementation in intelligent manufacturing and industrial production. First, the lack of high-quality data remains a core bottleneck limiting model performance. The uneven distribution of wood defects—common defects such as knots account for more than 60% of defects, while critical defects like micro-cracks and insect holes comprise less than 5%—results in a significant reduction in the model's ability to recognize minority class defects. Additionally, the annotation process heavily relies on manual expertise, which is costly and lacks standardization, further limiting the construction of large-scale datasets and model training. Current mitigation strategies include traditional data augmentation, the use of Generative Adversarial Networks (GANs) to generate synthetic defect images, and transfer learning to leverage feature knowledge from other industrial domains. However, these methods still struggle to fully address the class imbalance problem, with detection performance for minority classes remaining limited. Future efforts should focus on building large-scale, multi-species, standardized datasets, promoting cross-domain data sharing, and exploring innovative sample generation methods based on diffusion models to meet the diverse defect detection needs of industrial environments.

Illumination variations, surface reflections, and complex background interference significantly affect the model’s generalization ability in industrial settings. Research shows that model performance can degrade by more than 25% in different scenarios, primarily due to over-reliance on domain-sensitive features. Existing methods attempt to improve robustness through domain adaptation and attention mechanisms, but they still face difficulties in feature decoupling in complex textured backgrounds. Potential solutions include developing domain-invariant feature learning architectures based on vision transformers, creating data augmentation strategies that incorporate physical priors, and establishing standardized testing benchmarks that cover multiple devices and environments to enhance model adaptability and reliability in heterogeneous industrial environments.

Detecting minor defects also presents a critical challenge. Studies show that when defect areas occupy less than 0.1% of the image, the model’s recall rate significantly declines. While multi-scale feature fusion techniques can improve the response to small targets, they incur high computational costs, making it difficult to balance accuracy with real-time performance. To address this, researchers have attempted to incorporate super-resolution reconstruction techniques to enhance detail perception, use hierarchical attention mechanisms to suppress background interference, and build hybrid architectures based on CNNs and transformers for collaborative modeling of local and global features. Additionally, improvements such as adaptive receptive field mechanisms and lightweight feature pyramid networks can enhance the detection of small defects while maintaining real-time performance, thus meeting the high precision requirements in complex textured environments.

ISSN: 2472-3703 DOI: 10.6911/WSRJ.202510_11(10).0008

Industrial applications also have stringent demands for real-time processing and model efficiency. Typically, the processing time for a single image needs to be under 100 ms, yet existing lightweight networks exhibit a significant decline in accuracy when compression rates exceed 40%. Optimization approaches include structured pruning, quantization, knowledge distillation, and neural architecture search. Combining edge computing with cloud collaboration, along with algorithm and hardware co-optimization, will contribute to efficient inference and meet the comprehensive requirements for speed, stability, and resource consumption in high-output production lines.

3.2. Future Development Trends and Technological Directions

The future progress of wood surface defect detection will exhibit a trend of multidimensional integration and system optimization. At the data level, multimodal information fusion will be a key approach to enhance the model's perceptual ability. RGB images provide texture and color information, infrared imaging detects internal material temperature and potential defect signals, and depth images offer structural and geometric features. Multimodal collaborative modeling can significantly enhance defect detection robustness, especially in industrial environments with varying lighting conditions or strong surface reflections. Additionally, self-supervised learning and few-shot learning strategies can extract useful features under limited labeled data conditions, reducing labor costs and improving model adaptability to new tree species or environments.

At the algorithmic level, deep learning networks are evolving towards stronger feature modeling capabilities and global information perception. Vision Transformers and their derivatives enable long-range dependency feature modeling, effectively capturing the global correlation between complex wood surface textures and subtle defects. Hybrid architectures combining CNNs and Transformers provide collaborative representations of local details and global context, improving model stability in detecting small defects and complex backgrounds. Techniques such as multi-scale feature fusion, attention mechanisms, and adaptive receptive fields will continue to be optimized to address challenges such as large defect size variation, complex textures, and strong background interference, achieving high-precision recognition of various defect types, including micro-cracks, wormholes, knots, and decay. Algorithm lightweighting and model compression technologies will ensure real-time deployment on edge devices and industrial production lines, balancing accuracy and processing efficiency.

At the system level, end-to-end intelligent inspection pipelines will gradually integrate the complete process from image acquisition, preprocessing, and defect recognition to explainability analysis, enabling automation and transparency in industrial quality control. The integration of cloud and edge computing architectures will facilitate large-scale data processing and real-time detection, meeting the speed and stability requirements of high-throughput production lines. Furthermore, cross-domain knowledge transfer, multimodal information sharing, and algorithm-hardware co-optimization are expected to further enhance model generalization capabilities, providing reliable data support for industrial decision-making. With the continuous evolution of deep learning algorithms, the proliferation of computational resources, and the growing demand for industrial applications, wood surface defect detection is poised to transition from laboratory research to industrial application, providing solid technical support for smart manufacturing and the green building materials industry.

REFERENCES

[1] Future projections of global wood carbon flow and stock and climate change mitigation potential considering waste wood recycling[J]. Biomass Bioenergy, 2025.

ISSN: 2472-3703 DOI: 10.6911/WSRJ.202510_11(10).0008

[2] Song L, Liu Y T. Research on weighted HOG and feature fusion behavior recognition method[J]. Software Guide, 2021, 20(11): 53-56.

[3] Wang X F, Liu C, Zhao Y Q. Threshold automatic selection algorithm based on grey system theory[J]. Computer Applications, 2010, 46(19): 154-156+219.

[4] Xiao Y Q, Yang H M. Research on application of object detection algorithm in traffic scene[J]. Computer Engineering and Applications, 2021, 57(6): 30-41.

[5] Xiao J Y, Ma Y, Zhou Z, et al. Wood surface knot defect detection based on Faster R-CNN model[J]. Wood Industry, 2020, 34(02): 45-48.

[6] Zhu Y F. Research on wood surface defect detection based on Transformer[D]. Liaoning University of Science and Technology, 2024.

[7] Xia J J. Research on wood surface defect detection based on machine vision[D]. Qilu University of Technology, 2025.

[8] Ansari M M, Kumar S, Bin Heyat M D, et al. SVMVGGNet-16: A novel machine and deep learning-based approach for lung cancer detection using combined SVM and VGGNet-16[J]. Current Medical Imaging, 2025, 21: e15734056348824.

[9] Gao M Y. Research on wood knot defect detection based on ResNet convolutional neural network[D]. Northeast Forestry University, 2022.

[10] Dong H S, Xian X F, Sun X, et al. Tea disease recognition based on multi-stage feature fusion MobileNet v3-ULAM[J]. Jiangsu Agricultural Sciences, 2025, 53(15): 200-211.

[11] Zhao Q, Qu J W. An improved SAM-ShuffleNet Ming Qinghua dragon pattern chronology model[J]. Modern Electronic Technology, 2025, 48(09): 116-123.

[12] Rong Q, Tian Q C, Tan R. Real-time wood surface defect detection method based on improved Yolov5s[J]. Forest Products Industry, 2025, 62(01): 64-71.

[13] Ge Y, Ji H, Liu X. Wood surface defect detection based on improved YOLOv8[J]. Signal, Image and Video Processing, 2025, 19: 663.

[14] Zhang C, Zhang S. A lightweight vehicle information detection method for bridges based on semantic segmentation and two-branch neural networks of monocular vision[J]. Signal, Image and Video Processing, 2025, 19: 1032.

[15] Deng Z C, Yang Z C, Chen C, et al. Multimodal plane instance segmentation with the Segment Anything Model[J]. Automation in Construction, 2025, 180: 106541.