

Wood Surface Defect Detection Based on Improved YOLOv8s

Shunyong Zhou, Hao Zhu, Xue Liu, Qin Hu, Huan Lu, Ziyang Peng

Abstract—Wood surface defect detection poses challenges due to the diverse range of defects, making accurate localization and identification difficult. In this study, we introduce an enhanced approach for detecting flaws on wood surfaces by leveraging an augmented version of the YOLOv8s algorithm. To improve the focus on problematic target qualities, we initially constructed a HAM (hybrid attention module) structure within the Backbone. This structure incorporates spatial and channel attention techniques, enhancing the ability to identify defects. Additionally, we enhance the feature fusion capabilities by augmenting the expansion convolution module, reducing information loss during the connection with the Neck network. This augmentation improves the target receptive field, ensuring critical information preservation for effective diagnosis of wood surface defects. Furthermore, we introduce ghost convolution to enhance feature expression while minimizing the number of parameters. This approach optimizes the model's overall performance. Through extensive testing, our proposed GH-YOLOv8s model demonstrates accurate detection of five distinct types of wood surface defects, including defect types such as Live Knot, Dead Knot, Resin, Knot with crack, and Crack, achieving a mean average precision (mAP) of 98.4%. This performance surpasses the original model by 2.0% and maintains a high FPS (Frames Per Second) rate of 163.9, this means it can achieve efficient object detection in real-time scenarios. Moreover, our approach outperforms commonly used target detection methods, establishing its superiority in wood surface defect detection.

Index Terms—Wood surface defects, HAM, Expansion convolution, Ghost convolution

Manuscript received August 11, 2023; revised December 28, 2023.

This work was supported in part by the Innovation Fund of Postgraduate, Sichuan University of Science & Engineering (Y2022129), the Sichuan Provincial Department of Science and Technology Projects (2020YFSY0027), the Innovation and Entrepreneurship Program for College Students (S202210622033) and The National Natural Science Foundation of China (61701330).

Shunyong Zhou is an associate professor at School of Automation and Information Engineering, Sichuan University of Science & Engineering, Yibin, China (corresponding author: tel: +086-15808229643; e-mail: zsy22219750517@163.com)

Hao Zhu is a postgraduate student at School of Automation and Information Engineering, Sichuan University of Science & Engineering, Yibin, China (e-mail: 1998hao1998@gmail.com).

Xue Liu is a postgraduate student at School of Automation and Information Engineering, Sichuan University of Science & Engineering, Yibin, China (e-mail: lx1352839149@163.com).

Qin Hu is a postgraduate student at School of Automation and Information Engineering, Sichuan University of Science & Engineering, Yibin, China (e-mail: 2094758200@qq.com).

Huan Lu is a postgraduate student at School of Automation and Information Engineering, Sichuan University of Science & Engineering, Yibin, China (e-mail: 15883522610@163.com).

Ziyang Peng is a postgraduate student at School of Automation and Information Engineering, Sichuan University of Science & Engineering, Yibin, China (e-mail: 19808146528m@sina.cn).

I. INTRODUCTION

THE "State of the World's Forests 2022" study emphasized the importance of forests and trees, highlighting the alarming destruction of over 420 million hectares of forests since the 1990s. To address this environmental crisis posed by biodiversity and harness sustainable economic benefits, it is crucial to implement effective measures for protecting forest resources [1]. A key aspect of forest preservation involves the prudent utilization of timber resources. Wood widely in various industries such as construction and furniture production, often exhibits multiple defects during its creation and processing, including Knot, Crack, and Resin. These defects might result in a decline in the strength of the wood's durability and even directly influence its service life as well as overall aesthetic appeal. Traditionally, detecting such defects on wood surfaces have relied on human visual inspection, which is limited in terms of accuracy and efficiency. However, with advancements in computer vision and image processing technologies, there is a growing interest in leveraging digital image processing technology to automate the detection of wood surface defects. This emerging field has gained significant attention as a research hotspot, offering promising avenues for improving the detection and assessment of wood imperfections.

For instance, Reference [2] suggested a technique based on feature vector texture feature classification and picture block percentile color histogram to identify knot defects. Reference [3] proposed a method utilizing Gabor filters and morphological processing for detecting linear objects. However, this method is limited to linear defects and struggles with detecting multi-category defects. Reference [4] introduced an automatic classification method for wood defects using a support vector machine (SVM), capable of classifying four defect types. Reference [5] presented an enhanced YOLOv3 [6] detection approach that offers more precise identification of defects. However, its performance is inferior to the RefineDet [7] model when using a picture input size of 320*320. Reference [8] employed the DenseNet [9] network in the single-stage SSD [10] network structure to fuse multi-layer feature maps, resulting in higher recall and accuracy rates but longer detection times. Reference [11] utilized the YOLOv5 [12] algorithm to detect wood defects such as Dark Knots, Decay Knots, Edge Knots, and Pin Knots. However, they solely applied the original algorithm and did not propose any corresponding improvements.

While various detection techniques for wood surface defects have been explored, issues related to detection efficiency persist due to variations in surface fault shapes. This paper aims to enhance the accuracy, efficiency, and

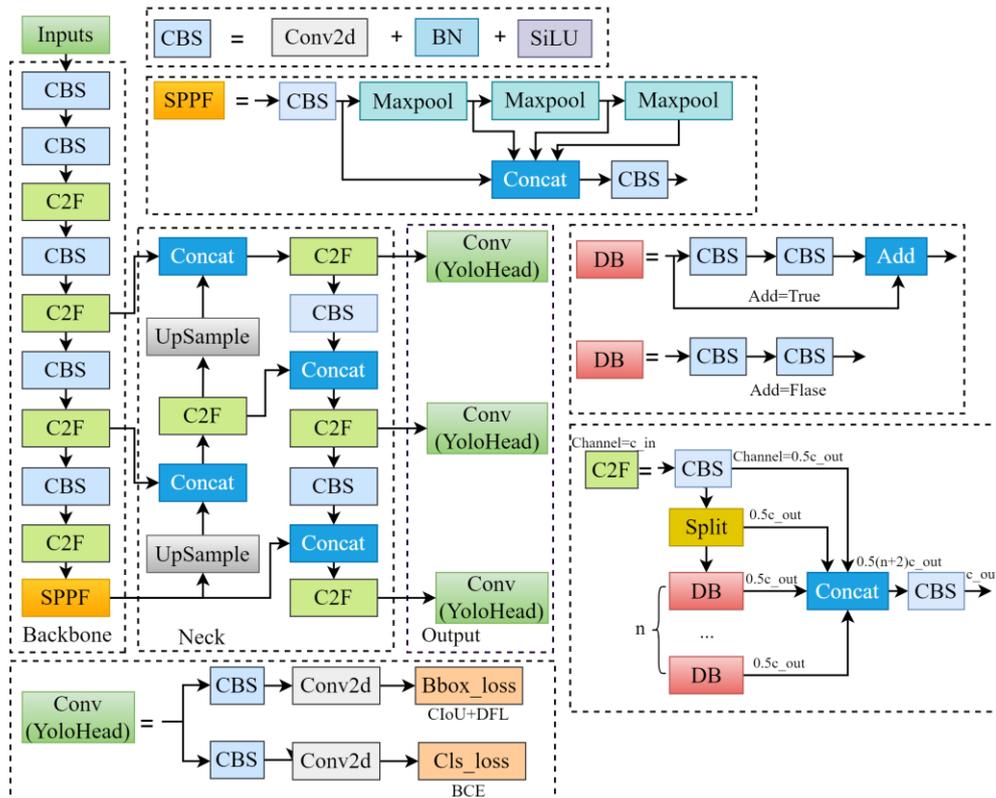


Fig. 1. The structure of the YOLOv8s model

automation of wood surface defect detection through a deep learning-based approach. Firstly, a ghost convolution [13] incorporating the Logish activation function [14] is employed to reduce parameters and calculations. Secondly, a hybrid attention mechanism (HAM) [15] module, combining spatial and channel attention, is used to retain critical spatial position information while focusing on essential channel features. Furthermore, a depthwise separable convolution block [16] with different expansion rates enhances the feature fusion capability of the Neck network. By implementing these optimization strategies, we can improve the YOLOv8s [17] detection performance on surface defects in the wood, which can fulfill positioning needs for wood defect identification and boost business automation in processing.

II. THE PRINCIPLE OF THE YOLOV8S MODEL

YOLOv8, developed by Ultralytics, is a novel algorithm in the YOLO series that incorporates new structural functionalities and optimization techniques from YOLOv5 and YOLOX [18]. These additions enhance the model's functionality, adaptability, and effectiveness. The YOLOv8 architecture consists of three main components: Backbone, Neck, and Output, as shown in Fig. 1.

In the initial stage, the input image undergoes feature extraction in the Backbone network. The extracted features are then passed to the Neck network, which further refines the details of the features. Finally, the Head network utilizes the refined features for target training and prediction.

The Backbone module consists of CBS (convolution layer, BN layer, SiLU activation function) and C2F structures. The primary part of CBS is to strengthen the semantic expression of features. Through a combination of the convolutional layer (CBS) and DB (a bottleneck block), the C2F module

performs a series of convolution operations on the input feature map to extract higher-level semantic features that capture the shape, texture, and other details of the detection target. YOLOv8s also incorporates the SPPF [19] module, which converts feature maps of varying sizes into fixed-size feature vectors. This is achieved by sequentially applying Maxpool with a 5*5 convolution kernel and cascading the results to expand the receptive field of feature map. The Neck network focuses on further extracting features obtained from the Backbone network and aims to combine feature information at different scales. YOLOv8 utilizes FPN [20] and PANet [21] structures to fuse features through upsampling and downsampling operations. This fusion process preserves the original features of the input image and increases the generalization capacity of the model. The YOLO Head serves as the classifier and regression component of YOLOv8, responsible for determining the presence of objects corresponding to the prior boxes on feature points and detecting and labeling them. YOLOv8s offers five models, namely YOLOv8n, YOLOv8s, YOLOv8m, YOLOv8l, and YOLOv8x. This work focuses on enhancing YOLOv8s as the core model, striking a balance between detection accuracy and speed.

III. IMPROVEMENTS TO THE YOLOV8S MODEL

The hybrid attention mechanism (HAM) module, the expansion convolution module (ECM), and ghost convolution are the key improvements made to the YOLOv8s model. These improvements have collectively resulted in the creation of GH-YOLOv8s, as illustrated in Fig. 2. In the following sections, we will delve into each component individually.

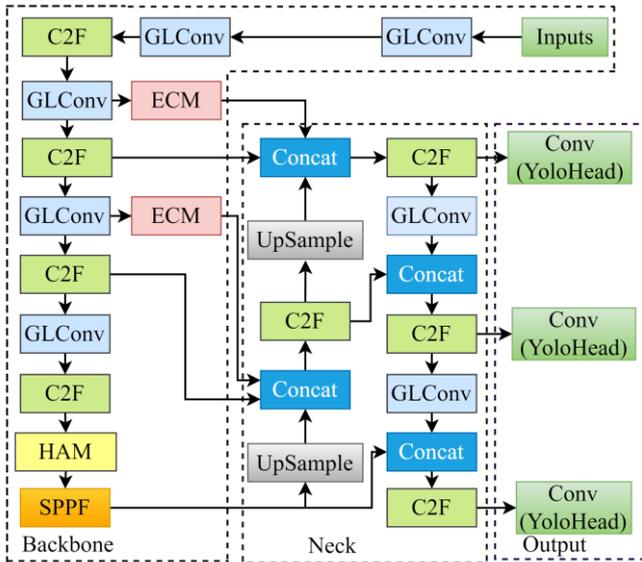


Fig. 2. The structure of the GH-YOLOv8s model

A. Logish Activation Function

The activation function plays a pivotal role within neural networks, enabling the introduction of nonlinear operations and breaking the limitations of simple linear transformations. This capability empowers the neural network to approximate complex functions effectively. Additionally, it addresses the challenges of disappearing or expanding gradients while enhancing the model's expressive capacity. In the YOLOv8s algorithm, the SiLU activation function is utilized. However, this function may encounter issues when the input values are extremely large or small, leading to outputs tending towards 0 or 1, potentially causing gradient vanishing or exploding challenges. To overcome these limitations, this article proposes the adoption of the Logish nonlinear activation function, which is particularly suitable for complex networks with higher learning rates. According to formula (1), the Logish function exhibits nonlinearity and nonmonotonicity. It begins by performing a logarithmic operation to narrow the numerical range of the sigmoid function. Subsequently, it applies significant regularization impact to the negative output using the variable "x".

$$f(x) = x \cdot \ln(1 + \text{Sigmoid}(x)) \quad (1)$$

Formula (2) is employed to illustrate the derivative of the Logish activation function, which plays a vital role in calculating gradients and updating network parameters during the backpropagation algorithm. This utilization guarantees a more stable training process for the neural network.

$$f'(x) = \ln\left(1 + \frac{1}{1+e^{-x}}\right) + \frac{x \cdot e^{-x}}{(1+e^{-x}) \cdot (2+e^{-x})} \quad (2)$$

The Logish activation function offers several advantages. Firstly, it ensures negative activation values and negative derivative values, which contribute to the robustness of the model during training. Additionally, it preserves partial sparsity, reducing parameter redundancy and enhancing the model's generalization capability. Importantly, switching to the Logish activation function does not introduce any increase in redundant calculations.

B. Lightweight Structure

Lightweight models offer efficient execution in resource-constrained environments and deliver faster inference speeds by minimizing model size and complexity. In this paper, the primary focus is on leveraging enhanced ghost convolution to achieve model lightweightness, as illustrated in Fig. 3. The ghost convolution process begins with a 1×1 convolution, utilizing the Logish activation function, to extract essential information from the input feature map. Subsequently, a depth-separable convolution, also employing the Logish activation function, is performed to generate a comparable feature map. Finally, we combine the 1×1 and depth separable convolution outputs.

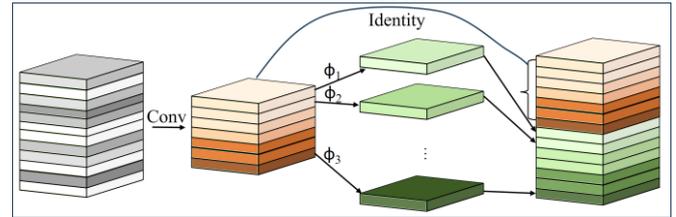


Fig. 3. The structure of ghost convolution

C. Attention Mechanism

The attention mechanism plays a crucial role in selecting important information from numerous features information while disregarding less relevant information. The typical attention mechanisms mainly include SE [22], ECA [23], and CBAM [24].

ECA enhances the SE attention mechanism by adaptively selecting the size of the one-dimensional convolution kernel. It facilitates cross-channel information interaction and avoids channel information loss during dimension reduction. Fig. 4 illustrates the schematic diagram of the ECA module.

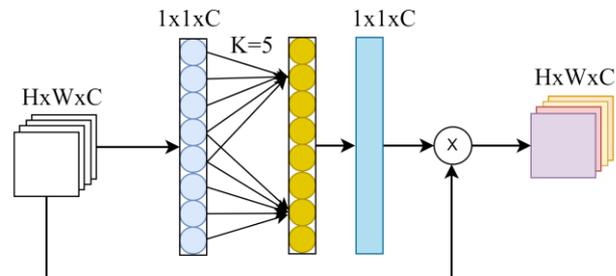


Fig. 4. The structure of the ECA attention mechanism

The ECA attention mechanism initially performs global average pooling (GAP) to obtain features. It then utilizes 1D convolution with a convolution kernel size "k" for adaptive feature selection. The weight "w" is obtained through the Sigmoid activation function, as shown in formula (2). Finally, the result is multiplied by the original feature map to get the output feature map. In the formula, σ represents the Sigmoid activation function, Conv1D denotes one-dimensional convolution, and k represents the size of the convolution kernel, which is proportional to the channel C, as shown in the formula (3-4). $|*|$ represents the odd number closest to *, and y and b represent the mapping slope and intercept, typically set as 2 and 1, respectively.

CBAM is a hybrid attention mechanism that integrates spatial and channel attention, as depicted in Fig. 5. The channel attention mechanism assigns different weights to

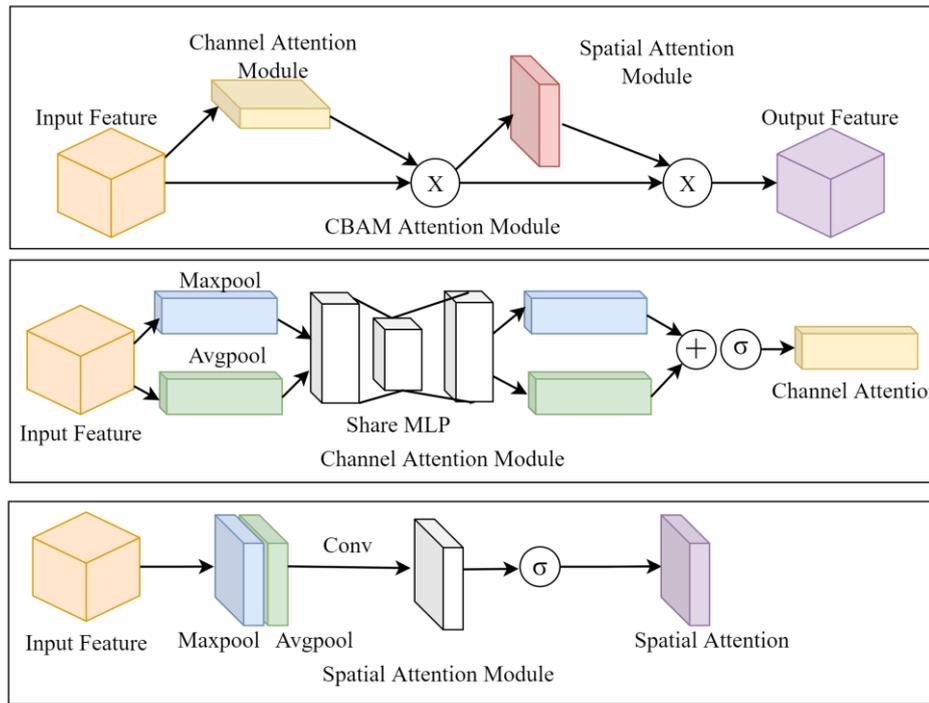


Fig. 5. The structure of the CBAM attention mechanism

various channels. It begins by applying Maxpool and Avgpool operations based on the width and height of the input feature map. The number of channels is adjusted through a neural network structure, and the output undergoes a summation operation. The resulting weight parameters are obtained through the Sigmoid activation function. Finally, the input feature map is multiplied by these weights to obtain the output feature map of the channel attention mechanism. The spatial attention mechanism assigns different spaces with different weights. It performs channel-based Maxpool and Avgpool operations on the feature map, resulting in two $H \times W \times 1$ feature maps. These maps are then concatenated along the channel direction. Subsequently, a 7×7 convolution reduces the dimensionality, followed by a Sigmoid function to generate weight parameters. These parameters are multiplied by the output feature map of the channel attention mechanism to obtain the output feature map of the CBAM attention mechanism.

This paper combines the ECA attention mechanism with the spatial attention mechanism within the CBAM attention mechanism, constructing a hybrid attention mechanism called HAM, as shown in Fig. 6. The spatial attention module utilizes the output of the ECA attention mechanism as input. By integrating multiple attention mechanisms, the hybrid attention mechanism enhances the model's performance from different aspects. This combination enables a more thorough and effective extraction of defect characteristics, thereby improving the model's generalization capacity.

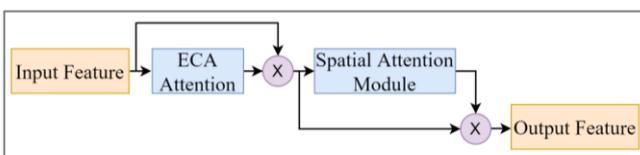


Fig. 6. The structure of the HAM attention mechanism

D. The Structure of ECM

The expansion convolution module, illustrated in Fig. 7, employs depthwise separable convolution to capture more precise feature information. It is seamlessly integrated with the Neck network to enhance feature fusion capabilities. The ECM structure uses a depthwise separable convolution, employing multiple expansion rates with a convolution kernel size of 3×3 . The feature maps obtained from different expansion rates are then connected using the Concat function. Subsequently, a depthwise separable convolution with 1×1 convolutions is utilized to actively restore the number of channels. This structural design aims to expand the model's receptive field and incorporate contextual information across different scales during feature extraction, thereby minimizing information loss. Simultaneously, it strengthens the connectivity between the Neck module and improves the pyramid feature fusion ability.

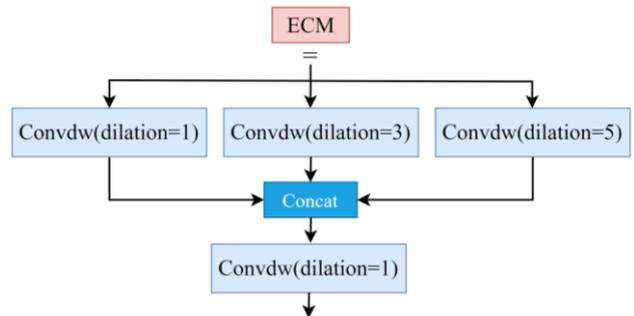


Fig. 7. The structure of ECM

IV. RESULT ANALYSIS

A. Lab Environment

The experimental environment actively utilizes a Windows 64-bit system, with the processor of Intel(R)

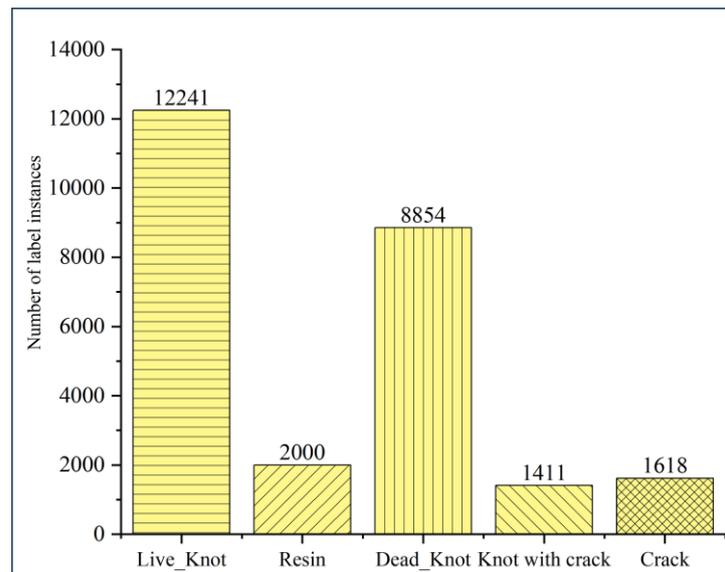


Fig. 8. Number of label instance

Xeon(R) Silver 4210R CPU @2.40GHz, 64GB of running memory. The used graphics card is NVIDIA RTX A5000, through PyTorch deep learning framework. GPU acceleration software was also used with CUDA11.6 and CUDNN8.4 versions.

B. Experimental Dataset

The experiment utilized a dataset sourced from a comprehensive collection of wood surface defects [25]. The dataset was captured using a three-line scanning camera, specifically the SW-4000TL-PMCL model, resulting in images with a resolution of 2800*1024. The dataset primarily consisted of defects in the form of Quartzity, Live Knot, Marrow, Resin, Dead Knot, Knot with crack, Knot missing, and Crack. However, the number of images containing more than three defect categories was quite small. Prior to the experiment, manual screening was performed on the dataset to remove any background noise or undesired features. Simultaneously, to tackle the issue of overfitting and augmenting the available data, data enhancement techniques such as cropping, translation, mirroring, and flipping were applied. These methods aimed to increase the dataset's quantity while enhancing the network's overall generalization capacity. As a result of this expansion process, the total number of datasets reached 14416, which were subsequently divided into an 80% training set and a separate verification set in a ratio of 8:2. In this experiment, we used these five sorts of labels as experimental data for training: Live Knot, Resin, Dead Knot, Knot with crack, and Crack. The number of labels is shown in Fig. 8.

C. Evaluation Index

To evaluate the performance of the model, the precision rate (Precision), the recall rate (Recall), and the mean Average Precision (mAP) are mainly used as the performance indicators of the algorithm, as expressed in formulas (3), (4) and (5).

$$Precision = \frac{TP}{TP + FP} \times 100\% \quad (3)$$

$$Recall = \frac{TP}{TP + FN} \times 100\% \quad (4)$$

$$mAP = \frac{\sum AP}{N(class)} = \frac{\sum \int_0^1 P(R) dR}{N(class)} \quad (5)$$

Among them, TP, FP, and FN respectively represent the correctly detected sample data, the wrongly detected sample data, and the incorrectly detected sample data. N denotes the number of types of target tags. AP corresponds to the area under the (Precision-recall) PR curve, where Recall is plotted on the horizontal axis and the Precision is plotted on the vertical axis. To enable performance comparison across all target classes, the evaluation metric employed for the model is the mean Average Precision (mAP) computed across all classes. Several hyperparameters are defined when conducting experiments. These include an initial learning rate of 0.01, an input image size of 640x640, a batch size of 16, 200 epochs, an Intersection over Union (IoU) threshold of 0.5, and the utilization of 4 workers.

D. Performance Analysis

This study employs advanced techniques, including ghost convolution with Logish activation functions, a hybrid attention mechanism (HAM), and an expansion convolution module (ECM), to enhance the YOLOv8 algorithm's precision and effectiveness in identifying wood surface defects. The GH-YOLOv8s model outperforms the original model in terms of assessment measures such as Precision, Recall, and mAP. Results presented in Table I demonstrate a noteworthy improvement, with Precision increasing by 2.7%, Recall by 3.6%, and mAP by 2.0%. These enhancements highlight the GH-YOLOv8s model's enhanced feature extraction capabilities compared to the original model.

TABLE I
MODEL PERFORMANCE COMPARISON

Model	Precision/%	Recall/%	mAP@0.5/%
YOLOv8s	96.0	93.3	96.4
GH-YOLOv8s	98.7	96.9	98.4

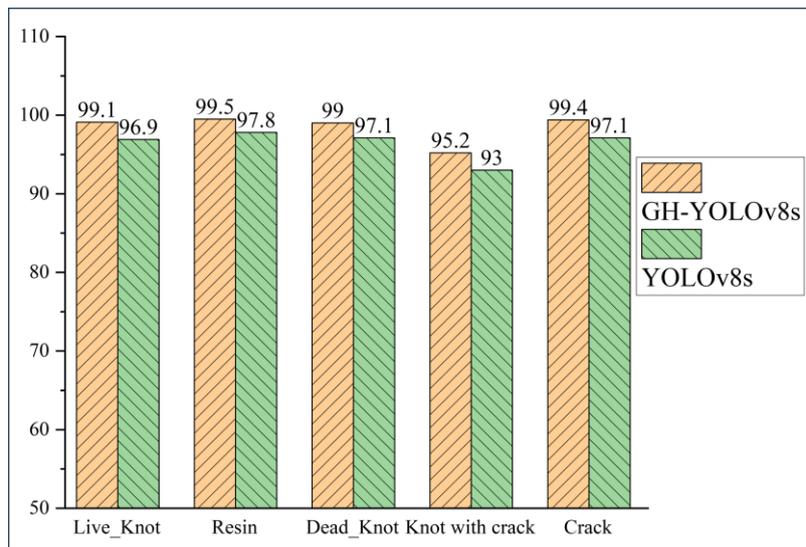


Fig. 9. Five types of label mAP comparison

TABLE II
ABLATION EXPERIMENT

GLConv	HAM	ECM	Precision/%	Recall/%	mAP@0.5/%
×	×	×	96.0	93.3	96.4
√	×	×	97.4	93.9	96.9
×	√	×	97.7	95.0	97.4
×	×	√	97.6	94.7	97.1
√	√	×	98.4	95.1	97.7
√	×	√	97.7	96.1	97.8
×	√	√	98.5	96.1	97.9
√	√	√	98.7	96.9	98.4

TABLE III
DETECTION RESULTS OF DIFFERENT ALGORITHMS

Model	Parameter/M	FPS	mAP@0.5/%	Weight/MB
GH-YOLOv8s	12.1	163.9	98.4	26.7
YOLOv7s	35.5	82.6	80.7	71.4
YOLOv6s	17.24	251.8	96.2	36.3
YOLOv5s	6.7	227.2	95.6	13.7
YOLOXs	8.94	88.9	77.9	68.5
YOLOv3s	61.5	30.5	82.7	235.1
SSD	24.1	50.9	82.6	100.1
Faster R-CNN	136.7	17.7	70.2	108.8

Moreover, the GH-YOLOv8s model excels in accurately and comprehensively capturing defect targets, leading to improved accuracy in wood surface defect detection. It efficiently performs precise defect detection tasks, as depicted in Fig.9, which showcases a comparison between the detection of each defect target using the GH-YOLOv8s model and the original model. The accuracy rate has significantly improved across the board.

E. Ablation experiment

To further validate the efficacy of the suggested approach, we conducted a set of ablation experiments, as illustrated in Table II. All improvement strategies employed in the experiments are based on the YOLOv8s model. In the table, the symbol "×" indicates the absence of a specific method, while the symbol "√" denotes its utilization. The results in the table clearly demonstrate the positive impact of incorporating the ghost convolution with Logish activation

functions. This addition has led to noticeable improvements in Precision, Recall, and mAP, allowing the model to effectively capture intricate relationships and nonlinear features within the feature maps. Furthermore, the inclusion of the HAM has resulted in a 1.0% increase in mAP, indicating that the model pays more focused attention to the defect targets. Moreover, the introduction of the ECM has enriched the contextual information and enhanced the feature representation. The structural integration of C2F and ECM within the Backbone, connecting it with the Neck network, facilitates the incorporation of additional semantic information into the feature pyramid layer. Alongside the other two modules, this integration effectively improves Recall and minimizes the loss of feature information. The collective contribution of these three modules significantly enhances the overall performance of the model and substantially augments its capability to identify wood surface defects.

F. Compare with Other Models

We compare the GH-YOLOv8s model with other popular target detection methods in the same configuration setting using a similar defect dataset to assess the efficacy of the model provided in this study, as summarized in Table III. In terms of mAP, the GH-YOLOv8s model exhibited remarkable improvement, surpassing YOLOv7s, YOLOv6s, YOLOv5s, YOLOXs, YOLOv3, SSD, and Faster R-CNN by 17.7%, 2.2%, 2.8%, 20.5%, 16.4%, 15.8%, and 17.1%, respectively.

Furthermore, the GH-YOLOv8s model offers significant advantages in terms of deployment on hardware devices. It weighs substantially less, with reductions of 62.6%, 26.4%, 61%, 88.6%, 73.3%, and 75.5% compared to YOLOv7s, YOLOv6s, YOLOXs, YOLOv3s, SSD, and Faster R-CNN, respectively. Additionally, it presents numerous benefits over models like YOLOv7s in terms of parameter efficiency. Although the FPS value is slightly lower than that of techniques such as YOLOv5s, it is still within acceptable bounds and can fulfill the demands of wood surface flaw detection in real-time.

G. Test Results

To provide a clearer comparison between the old model

and the upgraded model, several samples were tested to assess the impact of detection.

Fig. 10 illustrates that the upgraded GH-YOLOv8s model exhibits higher confidence levels for each defect target compared to the original model. This increased confidence translates into a more remarkable probability of accurately detecting such defects, showcasing the GH-YOLOv8s model's enhanced ability to precisely locate target objects in the scene.

Fig.11 demonstrates that, when utilizing the same IoU threshold and NMS parameters, the original YOLOv8s model fails to accurately eliminate overlapping frames. In contrast, the optimized GH-YOLOv8s model incorporates a structure that is better suited for detecting wood defects, resulting in improved detection performance.

Furthermore, Fig.12 reveals that the YOLOv8s model shows some missing detections, particularly for defect targets such as the Live_Knot and Crack. However, the GH-YOLOv8s algorithm significantly addresses these issues. It not only accurately detects and distinguishes five different types of defects but also improves the detection confidence for smaller defect targets. Therefore, the GH-YOLOv8s algorithm has better performance in detection.

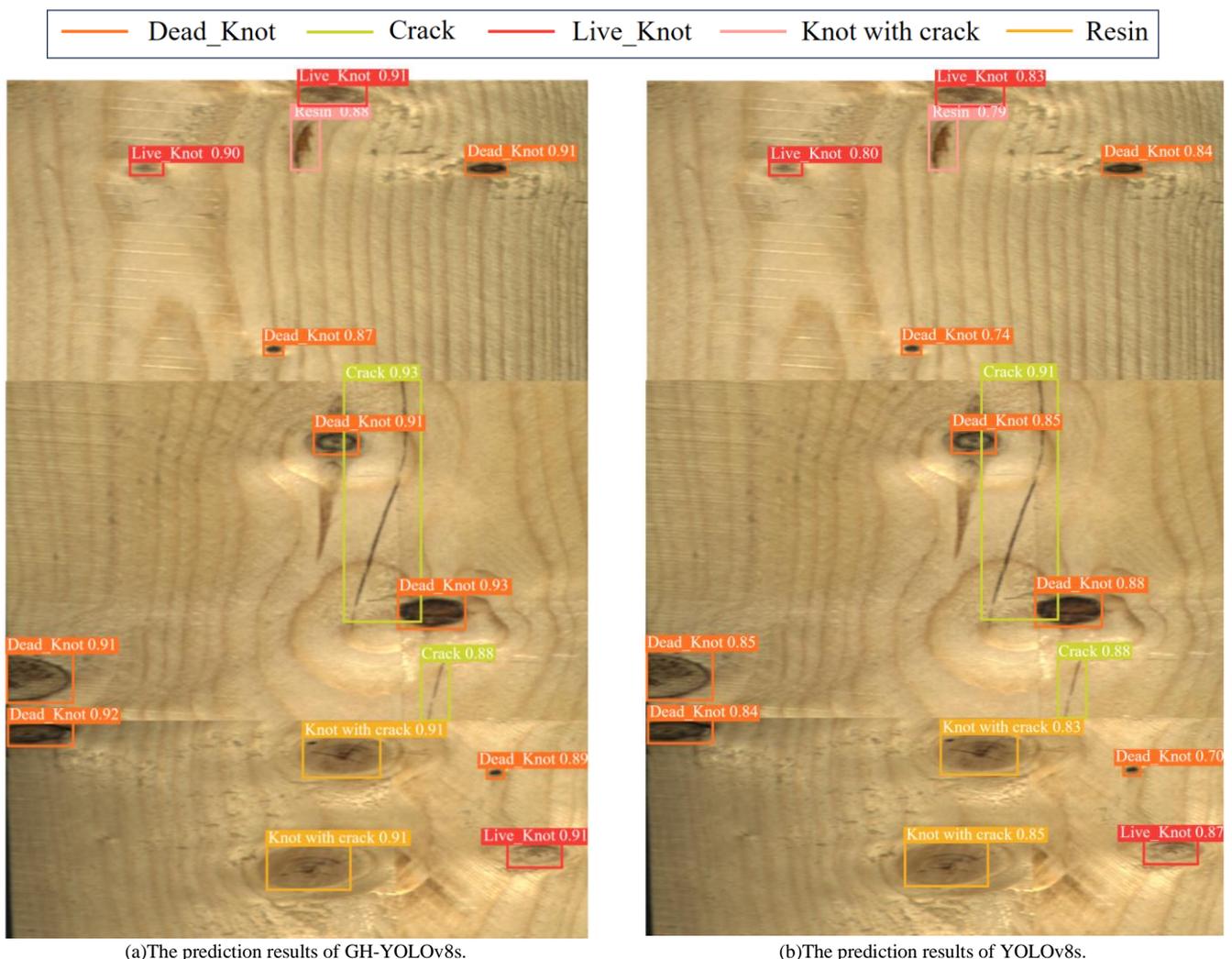
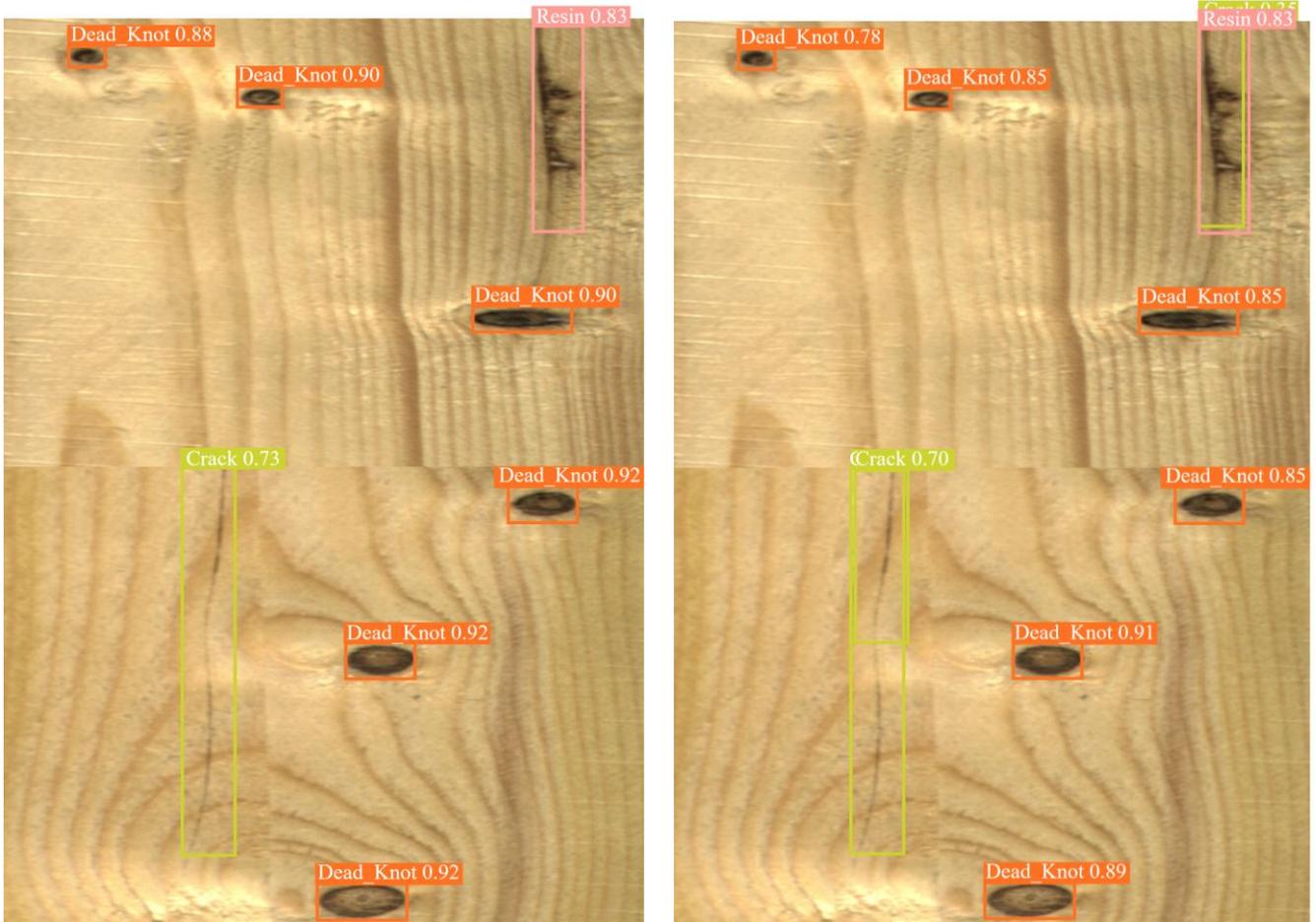


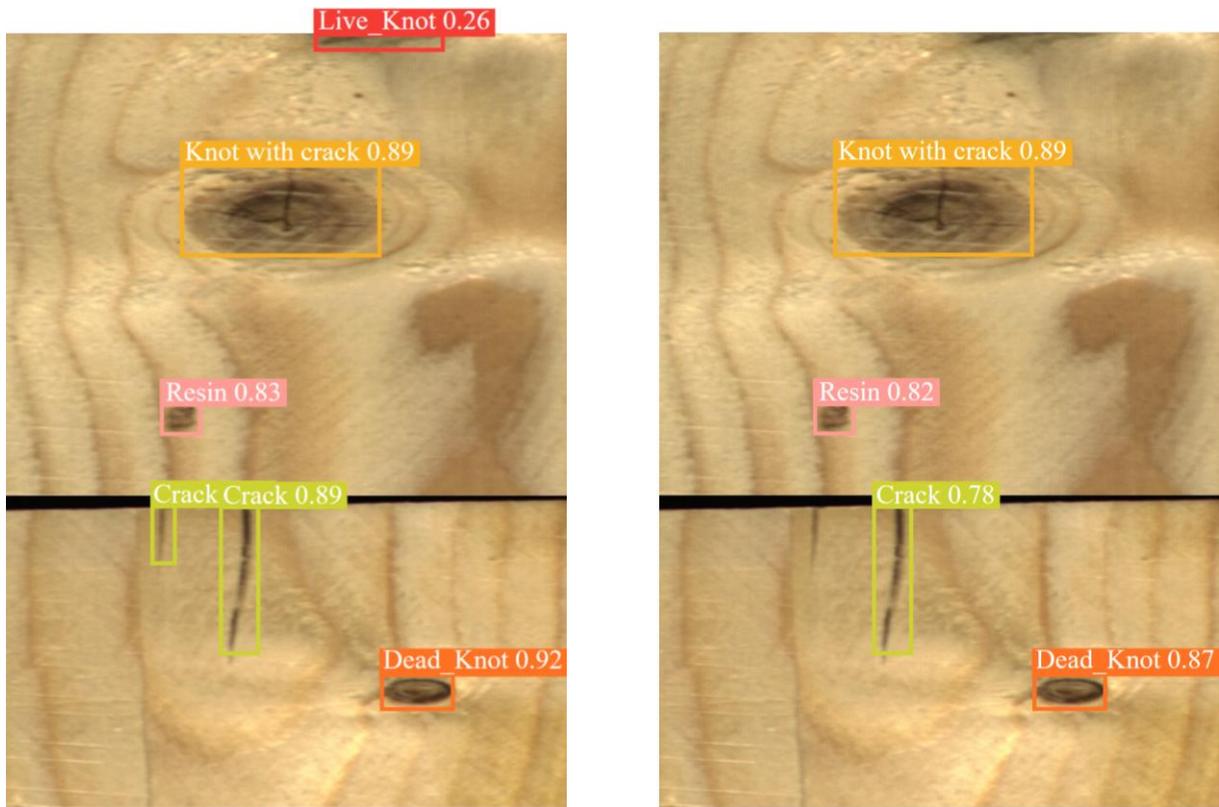
Fig.10. Confidence score comparison results.



(a)The prediction results of GH-YOLOv8s.

(b)The prediction results of YOLOv8s.

Fig.11. Detection results under the same parameters.



(a)The prediction results of GH-YOLOv8s.

(b)The prediction results of YOLOv8s.

Fig.12. Whether to miss the target comparison

V. CONCLUSION

This work focuses on the identification of wood surface flaws using the YOLOv8s model for detecting wood surface defects. Several enhancements have been incorporated to improve the model's performance. Firstly, the hybrid attention mechanism has been introduced into the Backbone network to gather feature information, enabling the model to prioritize the target of interest. Additionally, the expansion convolution structure has been added to expand the receptive field and enhance the feature extraction capabilities of the Neck network. To mitigate the increased parameter count resulting from these enhancements, ghost convolution has been employed to reduce the model's size. Experimental findings demonstrate that the GH-YOLOv8s model achieves higher accuracy and robustness in defect detection, with an impressive mAP of 98.4% and a model detection speed of 163.9 FPS. Additionally, GH-YOLOv8s surpasses other commonly used detection algorithms, confirming the effectiveness of the proposed enhancements. These findings provide strong support for practical applications and have the potential to enhance the level of automation in wood processing enterprises. In future work, we aim to expand the dataset to include a wider range of defect types, including smaller targets. We will refine and upgrade the YOLOv8 model to improve its flexibility for wood defect identification tasks. We also intend to increase the model's applicability in real-world scenarios such as industrial automation and wood quality monitoring, thus further advancing its practical utility.

REFERENCES

- [1] Food and Agriculture Organization of the United Nations, State of the World's Forests 2022, Available: <http://www.fao.org>.
- [2] Weiwei Song, Tianyi Chen, Zhenghua Gu, Wen Gai, Weikai Huang and Bin Wang, "Wood materials defects detection using image block percentile color histogram and eigenvector texture feature", in First International Conference on Information Sciences, Machinery, Materials and Energy, Atlantis Press, pp779-783, 2015.
- [3] W. Polzleitner, "Defect detection on wooden surface using gabor filters with evolutionary algorithm design", in IJCNN'01. International Joint Conference on Neural Networks. Proceedings (Cat. No. 01CH37222), IEEE, 07 August 2002, Washington, DC, USA, pp750-755, 2001.
- [4] Irene Y. H. Gu, Henrik Andersson and Raúl Vicen, "Automatic classification of wood defects using support vector machines", in International Conference on Computer Vision and Graphics, Springer, Vol. 5337, pp356-367, 2008.
- [5] Yaxin Tu, Zhigang Ling, Siyu Guo and He Wen, "An accurate and real-time surface defects detection method for sawn lumber", IEEE Transactions on Instrumentation and Measurement, Vol. 70, pp1-11, 2020.
- [6] Joseph Redmon and Ali Farhadi, "Yolov3: An incremental improvement", arXiv:1804.02767, 2018.
- [7] Shifeng Zhang, Longyin Wen, Xiao Bian, Zhen Lei and Stan Z. Li, "Single-shot refinement neural network for object detection". CVPR 2018: 4203-4212.
- [8] Fenglong Ding, Zilong Zhuang, Ying Liu, Dong Jiang, Xiaohan Yan and Zhengguang Wang, "Detecting defects on solid wood panels based on an improved SSD algorithm", Sensors, Vol. 20, No. 18, pp.5315, 2020.
- [9] Gao Huang, Zhuang Liu, Laurens van der Maaten and Kilian Q. Weinberger, "Densely connected convolutional networks". CVPR 2017: 4700-4708.
- [10] Wei Liu, Dragomir Anguelov, Dumitru Erhan, Christian Szegedy, Scott Reed, Cheng-Yang Fu and Alexander C. Berg, "Ssd: Single shot multibox detector", in Computer Vision—ECCV 2016: 14th European Conference, Amsterdam, The Netherlands, October 11–14, 2016, Proceedings, Part I 14, Springer, pp21-37, 2016.
- [11] Yiming Fang, Xianxin Guo, Kun Chen, Zhu Zhou and Qing Ye, "Accurate and Automated Detection of Surface Knots on Sawn Timbers Using YOLO-V5 Model.", BioResources, Vol. 16, No. 3, pp 5390-5406. 2021.
- [12] Jocher Glenn, Chaurasia Ayush, Stoken Alex and others, "ultralytics/yolov5: v7. 0-yolov5 sota realtime instance segmentation", Zenodo, 2022.
- [13] Kai Han, Yunhe Wang, Qi Tian, Jianyuan Guo, Chunjing Xu and Chang Xu, "Ghostnet: More features from cheap operations". CVPR 2020: 1580-1589.
- [14] Hegui Zhu, Huimin Zeng, Jinhai Liu and Xiangde Zhang, "Logish: A new nonlinear nonmonotonic activation function for convolutional neural network", Neurocomputing, Vol. 458, pp490-499, 2021.
- [15] Junfeng Li and Yuanxun Yang, "HM-YOLOv5: A fast and accurate network for defect detection of hot-pressed light guide plates", Engineering Applications of Artificial Intelligence, Vol. 117, Part A, 105529, 2023.
- [16] Francois Chollet, "Xception: Deep learning with depthwise separable convolutions". CVPR 2017: 1251-1258.
- [17] ultralytics, Available: <https://github.com/ultralytics/ultralytics>.
- [18] Zheng Ge, Songtao Liu, Feng Wang, Zeming Li and Jian Sun, "Yolox: Exceeding yolo series in 2021", arXiv:2107.08430, 2021.
- [19] Kaiming He, Xiangyu Zhang, Shaoqing Ren and Jian Sun, "Spatial pyramid pooling in deep convolutional networks for visual recognition", IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 37, No. 9, pp1904-1916, 2015.
- [20] Tsung-Yi Lin, Piotr Dollar, Ross Girshick, Kaiming He, Bharath Hariharan and Serge Belongie, "Feature pyramid networks for object detection". CVPR 2017: 2117-2125.
- [21] Shu Liu, Lu Qi, Haifang Qin, Jianping Shi and Jiaya Jia, "Path aggregation network for instance segmentation". CVPR 2018: 8759-8768.
- [22] Jie Hu, Li Shen and Gang Sun, "Squeeze-and-excitation networks". CVPR 2018: 7132-7141.
- [23] Qilong Wang, Banggu Wu, Pengfei Zhu, Peihua Li, Wangmeng Zuo and Qinghua Hu, "ECA-Net: Efficient channel attention for deep convolutional neural networks". CVPR 2020: 11534-11542.
- [24] Sanghyun Woo, Jongchan Park, Joon-Young Lee and In So Kweon, "Cbam: Convolutional block attention module". ECCV 2018: 3-19.
- [25] Pavel Kodytek, Alexandra Bodzas and Petr Bilik, "A large-scale image dataset of wood surface defects for automated vision-based quality control processes", F1000Research, Vol. 10, No.581, 2021.

Shunyong Zhou (1975-), male, master, associate professor. Main research directions: signal and information processing, image processing, circuit design.

Hao Zhu (1998-), male, master student. Main research directions: computer vision, image processing, object detection.

Xue Liu (1998-), female, master student. The main research directions: image processing, object detection.

Qin Hu (2000-), female, master student. The main research direction: signal processing.

Huan Lu (1999-), female, master student. The main research direction: signal processing.

Ziyang Peng (2000-), male, master student. The main research direction: image processing.