electronics

MDPI

Article

Lightweight Segmentation Method for Wood Panel Images Based on Improved DeepLabV3+

Xiangwei Mou [1,2], Hongyang Chen [3], Xinye Yu [1,2], Lintao Chen [1,2], Zhujing Peng [1,2] and Rijun Wang [1,2,*]

[1] School of Teachers College for Vocational and Technical Education, Guangxi Normal University, Guilin 541004, China; xwmou@mailbox.gxnu.edu.cn (X.M.); xy-yu@mailbox.gxnu.edu.cn (X.Y.); chenlintao@mailbox.gxnu.edu.cn (L.C.); pengzhujing@mailbox.gxnu.edu.cn (Z.P.)
[2] Engineering Research Center of Agricultural and Forestry Intelligent Equipment Technology, Education Department of Guangxi Zhuang Autonomous Region, Guangxi Normal University, Guilin 541004, China
[3] School of Electrical and Electronic Engineering, Guilin Institute of Information Technology, Guilin 541004, China; chy495817@gmail.com
* Correspondence: rijunwang@mailbox.gxnu.edu.cn

Abstract: Accurate and efficient pixel-wise segmentation of wood panels is crucial for enabling machine vision technologies to optimize the sawing process. Traditional image segmentation algorithms often struggle with robustness and accuracy in complex industrial environments. To address these challenges, this paper proposes an improved DeepLabV3+-based segmentation algorithm for wood panel images. The model incorporates a lightweight MobileNetV3 backbone to enhance feature extraction, reducing the number of parameters and computational complexity while minimizing any trade-off in segmentation accuracy, thereby increasing the model's processing speed. Additionally, the introduction of a coordinate attention (CA) mechanism allows the model to better capture fine details and local features of the wood panels while suppressing interference from complex backgrounds. A novel feature fusion mechanism is also employed, combining shallow and deep network features to enhance the model's ability to capture edges and textures, leading to improved feature fusion across scales and boosting segmentation accuracy. The experimental results demonstrate that the improved DeepLabV3+ model not only achieves superior segmentation performance across various wood panel types but also significantly increases segmentation speed. Specifically, the model improves the mean intersection over union (MIoU) by 1.05% and boosts the processing speed by 59.2%, achieving a processing time of 0.184 s per image.

Keywords: DeepLabV3+; feature fusion; wood panel image segmentation; MobileNetV3; coordinate attention mechanism

check for updates

Citation: Mou, X.; Chen, H.; Yu, X.; Chen, L.; Peng, Z.; Wang, R. Lightweight Segmentation Method for Wood Panel Images Based on Improved DeepLabV3+. *Electronics* **2024**, *13*, 4658. https://doi.org/10.3390/electronics13234658

Academic Editor: Beiwen Li

Received: 18 October 2024
Revised: 18 November 2024
Accepted: 19 November 2024
Published: 26 November 2024

Copyright: © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

1. Introduction

Wood is a valuable renewable resource widely used in furniture, construction materials, and decoration, among other fields [1,2]. Wood processing is a prerequisite for its application, typically involving several steps such as sawing, drying, machining, gluing, and surface treatment [3–5]. In the sawing process, bark removal from wood panels is a critical step that directly affects the quality of subsequent processing. Insufficient bark removal can leave residual bark, negatively impacting the appearance, durability, and stability of the wood panels. Conversely, excessive bark removal leads to the significant waste of wood resources. The key to ensuring the quality of bark removal lies in accurately identifying and segmenting the bark regions on the wood panels.

Initially, the task of identifying and segmenting bark was primarily reliant on skilled technicians using visual inspection. These technicians require extensive training and substantial practical experience to perform the task effectively. However, due to the influence of subjective factors, this method struggles to guarantee accuracy and has low efficiency. The development and application of machine vision and image processing technologies have

provided a practical solution to this challenge [6]. Consequently, bark removal methods based on machine vision, combined with traditional image segmentation techniques, have emerged as a viable alternative to the manual identification and segmentation of bark [7]. These methods not only eliminate the influence of subjective factors but also significantly improve both accuracy and efficiency [8]. The main approach in these methods involves capturing images of wood panels using industrial cameras, followed by traditional image segmentation techniques. By leveraging features such as the color, texture, edges, and regions of the bark, bark identification and segmentation are carried out using methods such as threshold segmentation [9,10], edge detection [11,12], region-based segmentation [13,14], and morphological operations [15,16]. While traditional image segmentation techniques have the advantages of simplicity and a fast processing speed, they are sensitive to changes in complex environments, resulting in limited robustness and insufficient segmentation accuracy in such scenarios [17,18].

Deep learning-based image segmentation methods have made substantial advancements, offering more robust solutions for handling complex images and diverse target features [19]. These methods leverage pixel-level image annotations, enabling network models to automatically learn high-level features and semantic information from images, thereby achieving more precise and refined segmentation results, even in challenging scenes [20]. With the growing prominence of deep learning and advances in computational power, deep learning-based segmentation techniques are gaining widespread attention in both research and industrial applications [21].

In complex industrial scenarios, Yang [22] proposed the RCEAU-Net model to address challenges related to real-time performance and feature extraction in laser beam target images within remote visual positioning systems used in tunnel excavation equipment, especially in the harsh conditions of coal mines. This model not only ensured the real-time segmentation of laser beams but also achieved improvements in accuracy by 0.19%, precision by 2.53%, recall by 22.01%, and intersection over union (IoU) by 8.48%. These enhancements meet the stringent requirements for multi-laser beam feature segmentation and extraction in complex coal mining environments. Similarly, Ju [23] improved the FCN semantic segmentation network by incorporating multi-scale convolutional operations, achieving F1 and IoU scores of over 87% and 85%, respectively, on a real-world industrial smoke emission image dataset. Compared with the above typical object segmentation tasks, the wood board image segmentation task presents unique challenges. The surface of the wood has a complex and intricate grain, with subtle differences in color and grain between the bark and the wood area. In addition, it is susceptible to changes in lighting and background noise. These factors make it more difficult to achieve efficient and accurate segmentation of the board image. As deep learning-based image segmentation algorithms continue to evolve, their applications in the wood processing industry are expanding [24,25]. They have been successfully implemented in tasks like wood defect detection and log end face segmentation, thereby accelerating the digital and intelligent transformation of the wood processing industry. He et al. [26] employed a hybrid fully convolutional neural network (FCN) to train a VGG16 model on a large dataset, utilizing transfer learning to enable automatic defect localization and segmentation on a smaller wood defect dataset. Similarly, Hu et al. [27] leveraged a progressive adversarial generative network to generate synthetic wood defect images, thereby expanding the dataset. They then applied the Mask R-CNN model for defect detection and segmentation, demonstrating the superior capabilities of deep learning in wood defect identification. Furthermore, Tao et al. [28] introduced a Patch-U-Net semantic segmentation method to address sample imbalance in remote sensing tree species classification, achieving pixel accuracy (PA), mean IoU, and frequency-weighted IoU (FWIoU) scores of 80.33%, 57.46%, and 67.37%, respectively. Compared to traditional image segmentation methods, deep learning-based models offer significant improvements in performance, robustness, and resilience to environmental interference [29,30]. Jia et al. [31] proposed a deep learning-based rice pest and disease detection model that integrates the MobileNetV3 backbone network with the CA mecha-

nism, achieving real-time detection performance without sacrificing detection accuracy. He et al. [32] introduce a novel semantic segmentation model that integrates a lightweight backbone network, a coordinate attention module, and a pooling fusion module, specifically designed for lightweight building extraction and adaptive spatial contour restoration. The evaluation results demonstrate that our model not only significantly enhances computational efficiency but also ensures high accuracy in extracting buildings from high-resolution images. These models provide an optimal solution for segmenting pixel regions of wood panels in complex industrial environments. Drawing inspiration from the aforementioned research, especially the successful integration of MobileNet with the coordinate attention (CA) method, which has achieved a commendable balance between accuracy and speed in visual recognition tasks, this context considers the visual requirements in wood panel processing and production that necessitate a dual focus on segmentation accuracy and speed. This study proposes an improved DeepLabV3+ network model. By refining the feature extraction and fusion components of DeepLabV3+, we enhance the model's capabilities in feature extraction and fusion, thereby improving both the speed and accuracy of wood image segmentation. The contributions of this study are as follows:

1. An improved DeepLabV3+ network model is proposed to address the challenges of segmentation accuracy and detection performance in wood image segmentation tasks, particularly those related to large model parameters and high computational complexity.
2. A multi-level upsampling feature fusion mechanism combined with a coordinate attention mechanism is proposed to enhance feature extraction and integration capabilities, which can more effectively capture key details and local features in wood images, and contribute to the improvement of segmentation results.
3. The proposed model is well balanced between segmentation speed and accuracy through architectural modifications that make it suitable for practical applications in complex industrial environments.

2. Development of the Improved DeepLabV3+ Model Architecture

2.1. Overall Model Architecture

DeepLabV3+ [33], as the latest semantic segmentation model in the DeepLab series, employs an encoder–decoder architecture that enhances the precision of edge segmentation, which is critical for tasks requiring accurate contour recognition. Additionally, the model's Atrous Spatial Pyramid Pooling (ASPP) [34] module incorporates multiple dilated convolutions with varying dilation rates, augmenting its ability to capture features at different scales. This capability allows DeepLabV3+ to comprehend a broader range of contextual information and improves its detail-capturing performance. Given these characteristics, DeepLabV3+ is an ideal candidate for wooden board image segmentation tasks. The structure of DeepLabV3+ is shown in Figure 1.

However, the diversity of wooden board shapes and the complex backgrounds present in real industrial scenarios lead to challenges such as low segmentation accuracy and a slow segmentation speed when employing the original DeepLabV3+ network model for wooden board area detection and segmentation. To address these issues, this study proposes a wood panel region recognition and segmentation model based on an improved DeepLabV3+. Retaining the ASPP module (as shown in Figure 2), we first incorporate a lightweight MobileNetV3 [35] backbone network to enhance the feature extraction network. This modification aims to reduce model parameters and computational complexity while minimally sacrificing segmentation accuracy, thereby improving the model's segmentation speed. Simultaneously, the introduction of a channel attention (CA) mechanism enables the model to better capture the details and local features of the wooden board area, effectively suppressing the interference from complex backgrounds. Finally, a feature fusion mechanism is designed to enhance the network model's ability to capture shallow features, such as the edges and textures of wooden board images, by fusing features from both shallow and deep networks. This approach strengthens the fusion effect of features at various scales,

ultimately improving the segmentation accuracy of the network model. The architecture of the improved DeepLabV3+ network model is depicted in Figure 3.

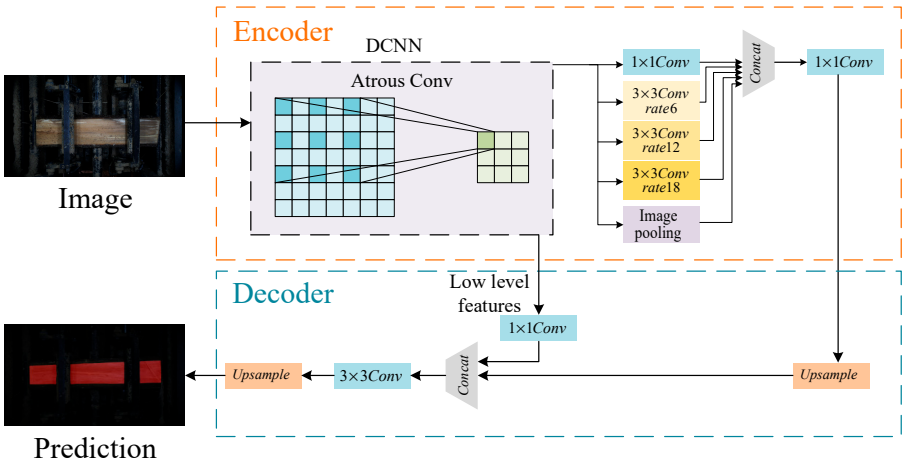


Figure 1. Structure of DeepLabV3+ model.

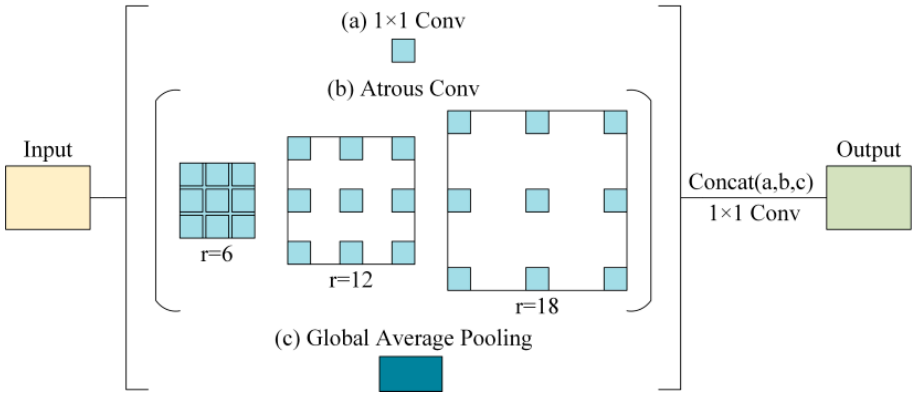


Figure 2. ASPP module structure.

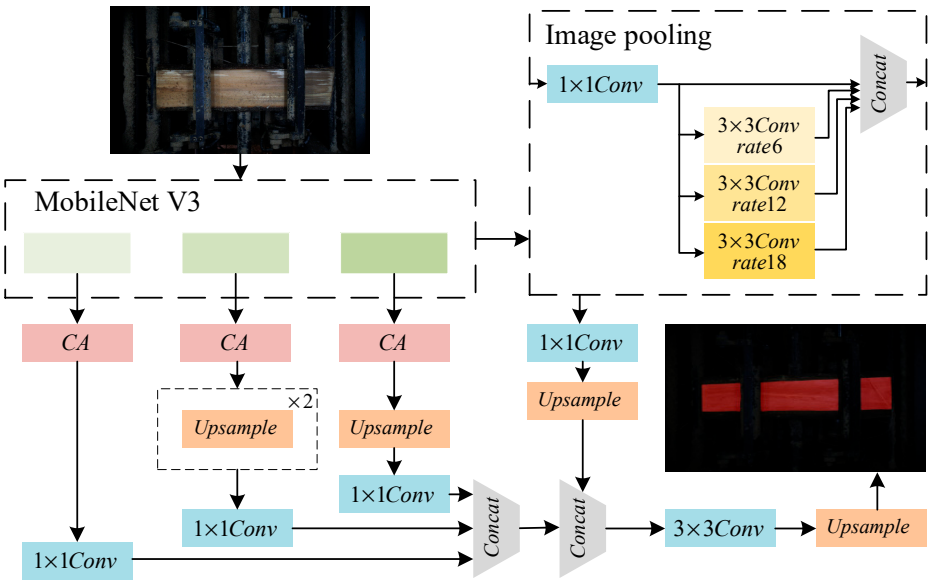


Figure 3. Architecture of improved DeepLabV3+ model.

2.2. Feature Extraction Based MobileNetV3

In the original DeepLabV3+, the Xception backbone network was employed as the feature extraction network, demonstrating commendable segmentation performance in practical applications. However, the model's real-time segmentation performance was constrained by its large number of parameters and complex computations. To enhance the real-time segmentation speed of the network model, this study proposes an improvement by utilizing the lightweight convolutional neural network MobileNetV3 as the backbone network for DeepLabV3+, thereby reducing the model's parameter count and computational complexity.

As the latest advancement in the MobileNet series, MobileNetV3 retains the advantages of its predecessors while introducing several enhancements. Notably, MobileNetV3 incorporates lightweight depthwise separable convolutions and an improved bottleneck structure, further decreasing the number of parameters and computational demands. These optimizations allow MobileNetV3 to achieve a lightweight model without compromising on accuracy and efficiency [36]. Consequently, this study adopts MobileNetV3 as the enhanced feature extraction network for DeepLabV3+.

To preserve more underlying image features, the downsampling factor of the network was adjusted to retain only four downsampling operations. The parameter settings of MobileNetV3 are presented in Table 1.

Table 1. Parameter settings for MobileNetV3.

Input	Operator	Exp Size	#out	SE	NL	s
512 × 512 × 3	Conv2d, 3 × 3	-	16	-	HS	2
256 × 256 × 16	bneck, 3 × 3	16	16	√	RE	2
128 × 128 × 16	bneck, 3 × 3	72	24	-	RE	2
64 × 64 × 24	bneck, 3 × 3	88	24	-	RE	1
64 × 64 × 24	bneck, 5 × 5	96	40	√	HS	2
32 × 32 × 40	bneck, 5 × 5	240	40	√	HS	1
32 × 32 × 40	bneck, 5 × 5	240	40	√	HS	1
32 × 32 × 40	bneck, 5 × 5	120	48	√	HS	1
32 × 32 × 48	bneck, 5 × 5	144	48	√	HS	1
32 × 32 × 48	bneck, 5 × 5	288	96	√	HS	1
32 × 32 × 96	bneck, 5 × 5	576	96	√	HS	1
32 × 32 × 96	bneck, 5 × 5	576	96	√	HS	1
32 × 32 × 96	bneck, 5 × 5	96	576	√	HS	1
32 × 32 × 576	Conv2d	-	320	-	-	1

Note: SE denotes whether there is a Squeeze-And-Excite (SE) in that block (√ means there is a SE in the module, - means there is not.). NL denotes the type of nonlinearity used. Here, HS denotes h-swish and RE denotes ReLU. s denotes stride.

2.3. Integration of the Coordinate Attention Mechanism

The coordinate attention (CA) mechanism [37] can enhance the feature expression ability of the model at low computational cost, thereby improving the segmentation accuracy of the semantic segmentation of wooden boards. The structure of the CA mechanism is shown in Figure 4.

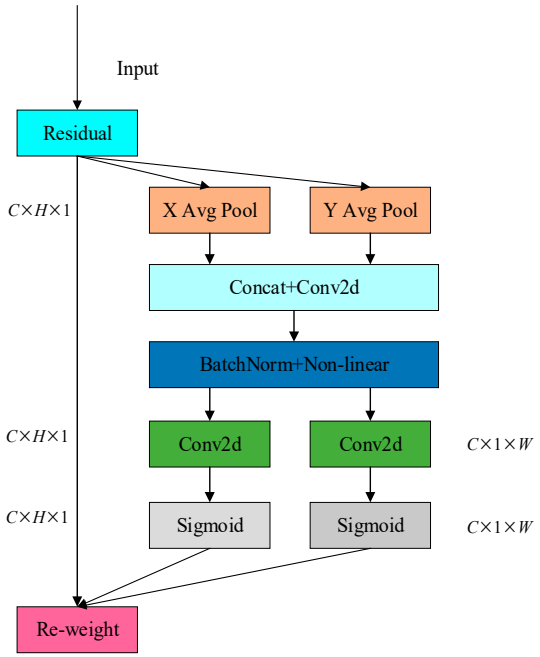


Figure 4. Structure of CA mechanism.

The CA mechanism avoids the complete compression of spatial information into channels by decomposing global pooling operations, thereby preserving crucial positional information. The key steps are as follows:

- Global Pooling Decomposition: Instead of traditional global pooling, the CA mechanism performs pooling operations separately in both the horizontal and vertical directions. This generates feature maps with dimensions of $1 \times H \times C$ and $W \times 1 \times C$, respectively. The calculation formula is as follows:

$$F_{\mathrm{h}}(i,k) = \frac{1}{W}\sum_{j=1}^{W} \mathrm{X}(i,j,k) \tag{1}$$

$$F_{\mathrm{w}}(j,k) = \frac{1}{H}\sum_{\mathrm{i}=1}^{H} \mathrm{X}(i,j,k) \tag{2}$$

- Feature Map Concatenation and Dimensionality Reduction: The feature maps from the horizontal and vertical pooling operations are concatenated, resulting in a feature map with dimensions of $1 \times (H + W) \times C$. Subsequently, a 1×1 convolution is applied to reduce the number of channels to C/r. After batch normalization and activation, a new feature map, denoted as F1, with dimensions of $1 \times (H + W) \times C/r$ is obtained.
- Feature Map Splitting and Attention Weight Calculation: The feature map F1 is split into two separate feature maps, F_{h1} and F_{w1}, corresponding to the height and width directions, respectively. These feature maps are then processed through activation functions to generate attention weights, denoted as A_{h} for the height direction and A_w for the width direction. The calculation formula is as follows:

$$A_{\mathrm{h}}(i,k) = \sigma(\mathrm{Conv}_{1\times1}(F_{\mathrm{h1}})) \tag{3}$$

$$A_{\mathrm{w}}(j,k) = \sigma(\mathrm{Conv}_{1\times1}(F_{\mathrm{w1}})) \tag{4}$$

- Weighted Calculation: The attention weights A_{h} and A_{w} are applied to the original feature map. The final feature map with attention weights is obtained through element-wise multiplication, enhancing the network's focus on critical regions.

By following these steps, the CA mechanism effectively preserves and leverages spatial positional information. This significantly enhances the segmentation accuracy in the semantic segmentation of wooden boards, while maintaining a low computational overhead.

2.4. Feature Fusion for Integrating Multi-Scale Spatial Information

The original DeepLabV3+ model employs the ASPP (Atrous Spatial Pyramid Pooling) module and a decoder structure for feature fusion, enabling it to capture contextual information at different scales and achieve multi-scale feature fusion [38]. Although this approach has yielded significant results in various tasks, it faces challenges when applied to specific tasks, such as the semantic segmentation of wooden boards. On one hand, the complex and varied surface textures of wooden boards necessitate a model with enhanced feature fusion capabilities. The original DeepLabV3+ model struggles to fully capture and leverage these fine-grained features, leading to inaccuracies in the segmentation of certain details. On the other hand, as the depth of the network increases, the spatial information in the feature map tends to diminish, posing a critical issue for semantic segmentation tasks that require the precise delineation of wooden board boundaries. Although the ASPP module can capture contextual information at multiple scales through varying dilation rates, it may not always fully recover lost spatial information in certain scenarios [39].

To address the aforementioned issues, we propose an improved feature fusion method for the DeepLabV3+ model, aimed at more effectively integrating multi-scale spatial information. First, feature extraction is conducted on the downsampled feature maps at scales of 4×, 8×, and 16× within the backbone network, yielding three feature maps of varying scales. For the feature maps downsampled by a factor of 4, we directly fuse them with the output of the ASPP module.

Before feature fusion, the CA mechanism is applied to the feature maps at each scale to obtain weighted feature maps. Subsequently, these three weighted feature maps are processed through a 1×1 convolution to adjust the channel dimensions, followed by a concatenation operation to produce the final fused feature map. This approach enables the improved DeepLabV3+ network model to effectively leverage feature information across different scales while preserving spatial information, thereby enhancing the accuracy of semantic segmentation for wood panels. Figure 5a,b illustrate the feature fusion structures before and after the improvements, respectively.

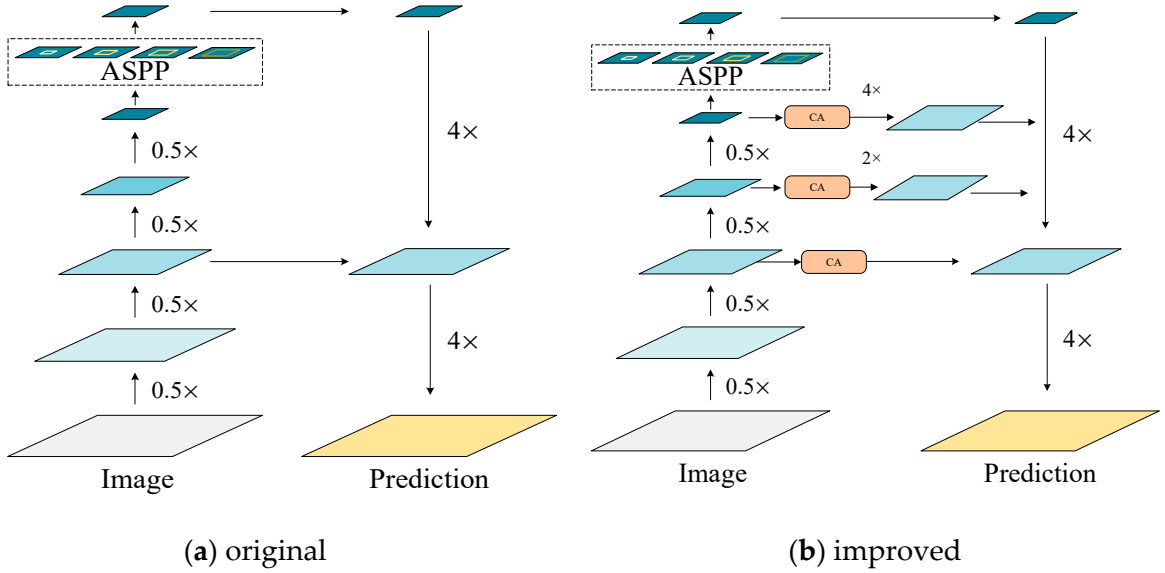


Figure 5. Structure of feature fusion module.

3. Experimental Results and Analysis

3.1. Overview of the Semantic Segmentation Dataset

The semantic segmentation dataset used in this study is sourced from the literature [40], which provides wood panel images captured by a customized visual acquisition device, as depicted in Figure 6. The dataset comprises a total of 3573 images of wood panels. For model training and validation, the dataset was randomly divided into a training set and a validation set in a ratio of 8:2, resulting in 2858 images in the training set and 715 images in the validation set.

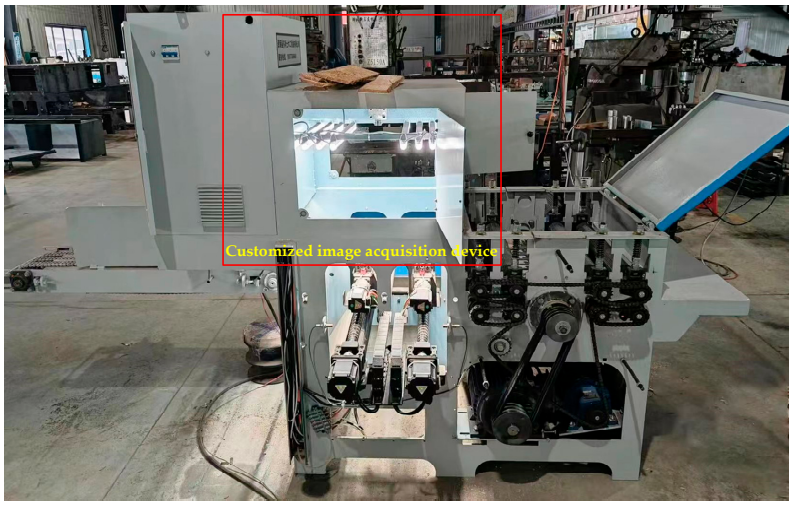


Figure 6. Customized image acquisition device [40].

3.2. Experimental Environment and Network Model Training

In this study, the DeepLabV3+ model was trained and optimized using Python 3.8 within the PyCharm integrated development environment. The hardware configuration included a 12th generation Intel Core i5-12490F processor and an NVIDIA GeForce RTX 3080 GPU. The experiments were conducted using the PyTorch 1.12.1 deep learning framework, with CUDA 11.3 serving as the backend for computation acceleration. The initial learning rate was set to 0.01, with a momentum value of 0.937 and a weight decay coefficient of 0.0005. The stochastic gradient descent (SGD) optimizer was employed for training. The batch size during training was set to 8, with the model being trained for 300 iterations. All input images were uniformly resized to a resolution of 512 × 512 pixels. Transfer learning techniques were utilized by loading pre-trained weights to initialize the DeepLabV3+ model. Figure 7 illustrates the loss function throughout the model training process.

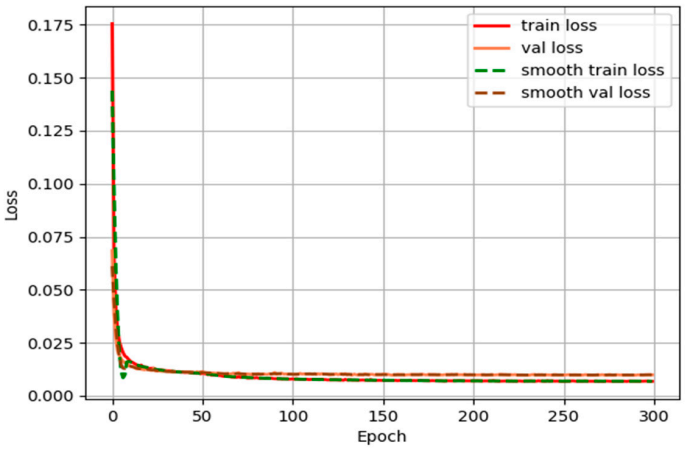


Figure 7. Model training loss function curves.

3.3. Evaluation Indicators

The segmentation of the wood panel region is framed as a binary classification task, wherein the objective is to categorize each pixel of the wood panel image into one of two classes: the panel region and the background. In the context of binary classification, particularly at the pixel level, several evaluation indicators are commonly utilized to assess the performance of the segmentation model:

1. Mean Intersection and Union Ratio (MIoU):

This is a critical metric for assessing the performance of an image segmentation model. The intersection over union (IoU) intuitively quantifies the similarity between the predicted segmentation and the ground truth segmentation. It is computed as the ratio of the intersection of the predicted and true regions to the union of those regions. The formula for calculating the IoU is given by the following:

$$\text{IoU} = \frac{\text{Area of Overlap}}{\text{Area of Union}} = \frac{A_{pred} \cap A_{true}}{A_{pred} \cup A_{true}} \tag{5}$$

where A_{pred} represents the area of the predicted segmentation and A_{true} denotes the area of the true segmentation. The mean IoU (mIoU) is then calculated by averaging the IoUs across all classes in the dataset, providing a holistic view of the model's segmentation accuracy.

2. Precision:

Precision measures the proportion of pixels that are correctly identified as positive samples (i.e., belonging to the target category) among all pixels that the model has predicted as positive. It serves as an indicator of the accuracy of the model's predictions. The precision is calculated using the following formula:

$$\text{Precision} = \frac{TP}{TP + FP} \tag{6}$$

where TP represents the number of true positive pixels (correctly identified target pixels) and FP denotes the number of false positive pixels (incorrectly identified as target pixels). High precision indicates that the model has a low rate of false positives, thereby demonstrating reliability in its predictions.

3. Recall:

Recall assesses the model's ability to identify positive samples (i.e., pixels belonging to the target category). It is defined as the proportion of correctly predicted positive pixels relative to the total number of actual positive pixels in the dataset. The recall is calculated using the following formula:

$$\text{Recall} = \frac{TP}{TP + FN} \tag{7}$$

where FN denotes the number of false negative pixels (actual target pixels that were not identified by the model). A high recall indicates that the model effectively captures most of the relevant positive samples.

4. Segmentation speed:

The segmentation speed indicates the time required for the model to generate a complete segmentation result after receiving a new image input. It is measured in seconds per image (s/image). This metric is crucial for evaluating the model's real-time applicability in industrial settings, where quick processing times are essential for efficient operations.

5. Model Parameter Size:

The model parameter size directly indicates the memory consumption of the model and serves as a key reference metric for assessing the model's size and complexity. Smaller parameter sizes generally lead to lower memory requirements, which can enhance the model's efficiency and deployment feasibility in resource-constrained environments.

3.4. Comparison of Experimental Results

3.4.1. Ablation Experiment

To validate the impact of different enhancements on the performance of wooden panel semantic segmentation, four ablation experiments were designed under identical environmental settings and hyperparameters. The results of ablation experiments are shown in Table 2.

Table 2. Comparison results of ablation experiments.

Experimental Group	Improvement Method	Precision/%	Recall/%	MIoU/%	Parameter Size/M	Segmentation Speed/(s/Image)
A	DeeplabV3+	98.36	98.82	97.38	20.9	0.452
B	MobileNetV3	98.02	98.75	97.08	22.3	0.177
C	MobileNetV3+CA	98.85	99.37	97.94	22.3	0.180
D	Final	99.03	99.50	98.43	22.4	0.184

- Group A utilized the original DeepLabV3+ network model as the baseline for comparison.
- Group B replaced the backbone with MobileNetV3, which resulted in a significant reduction in the model's parameter count to 22.3 M and improved detection speed, with minimal impact on accuracy.
- Group C, building on Group B, introduced the CA mechanism. This modification resulted in a notable increase in recall and accuracy, and the MIoU improved by 0.86%, demonstrating the ability of CA to enhance feature extraction and mitigate background interference.
- Group D, further refining Group C by improving the feature fusion structure, achieved an accuracy, recall, and MIoU of 99.03%, 99.50%, and 98.43%, respectively. Additionally, model parameters were reduced by 89.3%, and the segmentation speed was increased by 59.2%.

In summary, the proposed improvements in the DeepLabV3+ model not only significantly enhance the segmentation accuracy of wood panels but also improve the real-time segmentation speed, reinforcing the robustness and efficiency of the visual detection system in industrial environments.

3.4.2. Comparison with Other Semantic Segmentation Algorithms

To comprehensively evaluate the performance of the improved DeepLabV3+ network model in wood panel semantic segmentation, we compared it against three other widely used semantic segmentation models: PSP-Net, U-Net, and FCN. For consistency and fairness, all models were trained under identical conditions, including a unified experimental platform, consistent parameter configurations, and the same dataset partitioning. Each model was trained for 300 iterations to ensure optimal performance. A comparison of the key performance metrics for the four models is presented in Table 3. The results show that the improved DeepLabV3+ network achieved superior performance, with an accuracy of 99.03%, a recall of 99.50%, and an MIoU of 98.43%. These results highlight the model's effectiveness in wood image segmentation tasks. Moreover, the model exhibits a rapid segmentation speed of 0.184 s per image, substantially reducing memory usage while meeting real-time processing requirements.

Table 3. Experimental results of different models.

Network Model	Precision/%	Recall/%	MIoU/%	Parameter Size/M	Segmentation Speed/(s/Image)
PSP-Net [20]	98.29	97.90	96.29	179.6	0.379
U-Net [41]	98.50	98.77	97.24	96.9	0.732
FCN [19]	98.39	98.72	96.94	107.8	0.675
Ours	99.03	99.50	98.43	22.4	0.184

3.4.3. Analysis of Segmentation Results

To evaluate the performance of the improved DeepLabV3+ model for wood board segmentation, we randomly selected 100 images from the test set. These images included both wood boards with rich color variations and those with more consistent colors, allowing for a comprehensive assessment of the model's segmentation capabilities. Compared to other common segmentation models such as U-Net and HR-Net, the improved DeepLabV3+ model maintained high segmentation accuracy even when processing wood images with significant internal color differences, demonstrating superior performance over the other models. This advantage is particularly crucial as wood boards can exhibit a wide range of colors and textures due to material properties and processing conditions.

Additionally, the improved DeepLabV3+ model performed excellently in handling complex backgrounds, such as images with mechanical components or sawdust interference. It was able to accurately identify the edges and contours of the wood boards and effectively separate them from the background. In contrast, U-Net and FCN struggled with these complex environments, often leading to background interference or missegmentation. Overall, the improved DeepLabV3+ model exhibited robust performance in segmenting wood boards with diverse colors and textures, and it also performed exceptionally well in complex industrial settings. Its high segmentation accuracy and strong adaptability highlight the model's practical application value and potential for real-world scenarios. Figure 8 shows the segmentation results of the model in real-world wood board scenes, providing a stark contrast to the performance of other models.

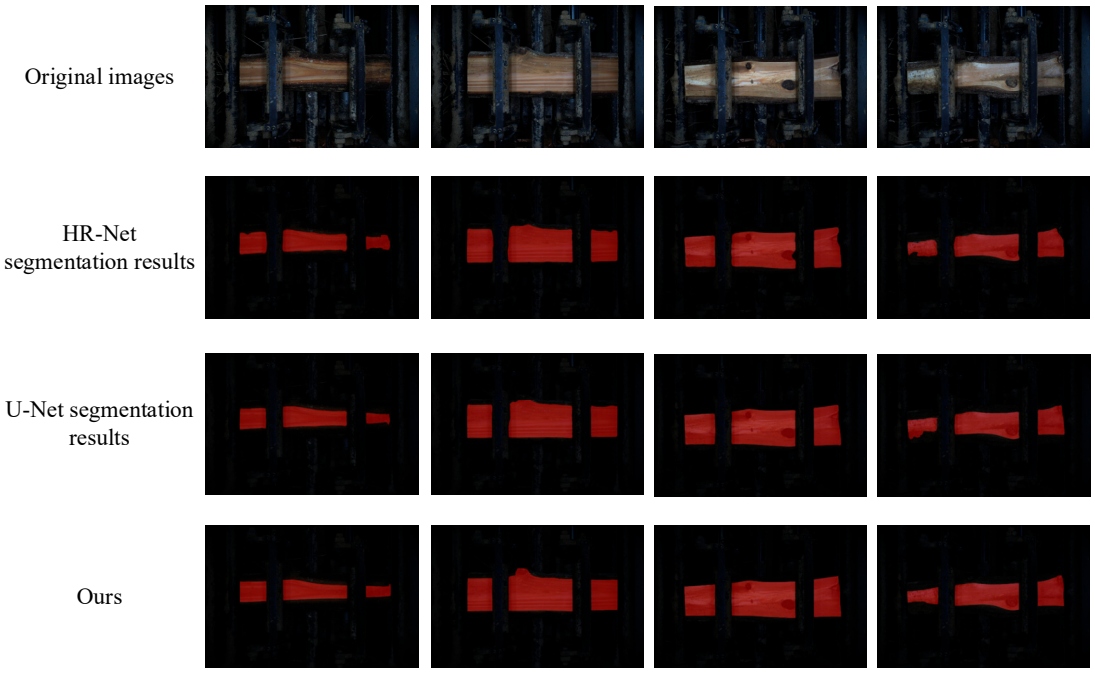


Figure 8. Comparison of segmentation results of different models.

4. Conclusions

To achieve an accurate and efficient segmentation of wood panel that meets the practical needs of wood sawing processes in complex industrial environments, this paper presents a lightweight segmentation method based on an improved DeepLabV3+ model. The enhancements focus on refining the feature extraction network and the feature fusion mechanism. The proposed method is validated using a dataset constructed from real-world wood sawing scenarios. Compared to the original model, the improved version demonstrates a 1.05% increase in the MIoU and a remarkable 59.2% enhancement in segmentation speed, processing each image in just 0.184 s. This model not only provides robust segmentation results for various types of wood panels but also significantly improves segmentation speed, effectively satisfying the dual requirements of real-time performance and accuracy for industrial applications. This study offers valuable technical support for the automation of wood processing and serves as a useful reference for future research in related fields.

Author Contributions: Conceptualization, X.M.; Methodology, X.M., H.C. and R.W.; Software, H.C. and X.Y.; Validation, Z.P.; Formal analysis, L.C.; Investigation, X.M., L.C. and Z.P.; Data curation, L.C., X.Y. and Z.P.; Writing—original draft, H.C.; Writing—review & editing, R.W.; Visualization, H.C. and X.Y.; Supervision, R.W.; Project administration, X.M.; Funding acquisition, R.W. All authors have read and agreed to the published version of the manuscript.

Funding: This study was co-supported by the Science and Technology Planning Project of Guangxi Province, China (No. 2022AC21012); the industry-university-research innovation fund projects of China University in 2021 (No. 2021ITA10018); the fund project of the Key Laboratory of AI and Information Processing (No. 2022GXZDSY101).

Data Availability Statement: Data are contained within the article.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Pulkki, R.; McDonald, G. Wood as a renewable resource: Societal benefits and sustainable development. *J. Sustain. For.* **2020**, *39*, 573–589.
2. Wiedenhoeft, A.C.; Miller, R.B. *Wood Handbook: Wood as an Engineering Material*; Forest Products Laboratory: Madison, WI, USA, 2019; pp. 1–5. ISBN 978-1422050068.
3. Tsoumis, G. *Science and Technology of Wood: Structure, Properties, Utilization*; Van Nostrand Reinhold: New York, NY, USA, 2020; pp. 255–275. ISBN 978-0442243181.
4. Zobel, B.J.; van Buijtenen, J.P. *Wood Variation: Its Causes and Control*; Springer: Berlin/Heidelberg, Germany, 2019; pp. 65–85. [CrossRef]
5. Simatupang, M.H.; Welling, J. Wood processing: An overview of the industry. *Wood Sci. Technol.* **2021**, *55*, 505–530. [CrossRef]
6. Zhang, X.; Wu, Q.; Zuo, W.; Lin, L. Machine vision in automated manufacturing: Current status and challenges. *IEEE Trans. Ind. Inform.* **2021**, *17*, 3480–3492.
7. Li, X.; Zhang, Y.; Wang, Y.; Wang, X. Wood defect detection based on image segmentation using machine vision. *J. Manuf. Process.* **2019**, *35*, 120–128. [CrossRef]
8. Sonka, M.; Hlavac, V.; Boyle, R. *Image Processing, Analysis, and Machine Vision*; Cengage Learning: Boston, MA, USA, 2014; pp. 650–660. ISBN 978-1133593607.
9. Singh, N.; Vatsa, M.; Singh, R.; Noore, A. Automated Face Recognition via a Lightweight Threshold Segmentation Algorithm. *IEEE Access* **2020**, *8*, 68105–68115.
10. Sun, J.; Yin, Y.; Luo, Z.; Zhang, G. A novel image thresholding method based on Gaussian mixture model. *Neurocomputing* **2021**, *452*, 535–544. [CrossRef]
11. Xie, S.; Tu, Z. Holistically-Nested Edge Detection. *Int. J. Comput. Vis.* **2017**, *125*, 3–18. [CrossRef]
12. Ma, H.; Qi, W.; Zhang, M.; Zhang, T.; Zhang, J. Canny edge detection enhancement by soft threshold and interpolation optimization. *Signal Process. Image Commun.* **2021**, *91*, 116039.
13. Fan, J.; Zeng, G.; Xia, X.; Wu, Y.; Zhu, Q. A New Robust Adaptive Region-Based Image Segmentation Model Using Local Statistical Information. *IEEE Trans. Image Process.* **2018**, *27*, 827–837.
14. Khan, S.; Siddiqui, M.F.; Ullah, A. An efficient region-based segmentation method using local binary pattern and morphological operations. *Multimed. Tools Appl.* **2021**, *80*, 18877–18897.
15. Gonzalez, J.; Sanchez-Azofeifa, G. Image Processing and Morphological Transformations for Remote Sensing Forest Applications. *Remote Sens.* **2020**, *12*, 1653. [CrossRef]

16. Mehta, A.; Manhas, J.; Kumar, R. Morphological segmentation of MRI data for object separation. *J. Appl. Res. Technol.* **2020**, *18*, 59–68.
17. Gonzalez, R.C.; Woods, R.E. *Digital Image Processing*; Pearson: London, UK, 2018; pp. 520–530. ISBN 978-0133356724.
18. Wang, S.; Fan, J.; Wang, L. Challenges in traditional image segmentation for industrial applications: A review. *Comput. Ind.* **2020**, *125*, 103328. [CrossRef]
19. Long, J.; Shelhamer, E.; Darrell, T. Fully Convolutional Networks for Semantic Segmentation. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Boston, MA, USA, 7–12 June 2015; pp. 3431–3440.
20. Zhao, H.; Shi, J.; Qi, X.; Wang, S.; Jia, J. Pyramid Scene Parsing Network. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Honolulu, HI, USA, 21–26 July 2017; pp. 6230–6239.
21. Lin, T.-Y.; Goyal, P.; Girshick, R.; He, K.; Dollár, P. Focal Loss for Dense Object Detection. *IEEE Trans. Pattern Anal. Mach. Intell.* **2020**, *42*, 318–327. [CrossRef]
22. Yang, W.; Wang, Y.; Zhang, X.; Zhu, L.; Ren, Z.; Ji, Y.; Li, L.; Xie, Y. RCEAU-Net: Cascade Multi-Scale Convolution and Attention-Mechanism-Based Network for Laser Beam Target Image Segmentation with Complex Background in Coal Mine. *Sensors* **2024**, *24*, 2552. [CrossRef] [PubMed] [PubMed Central]
23. Ju, A.; Wang, Z. A novel fully convolutional network based on marker-controlled watershed segmentation algorithm for industrial soot robot target segmentation. *Evol. Intell.* **2023**, *16*, 963–980. [CrossRef]
24. Wang, H.; Zhang, X.; Li, Y.; Zhang, Y. Applications of Deep Learning in the Wood Industry: A Review. *Forests* **2021**, *12*, 1341. [CrossRef]
25. Zhou, Y.; Wang, S.; Sun, Y.; Liu, L. Wood Defect Detection Based on Improved Deep Learning Algorithm. *Sensors* **2020**, *20*, 4891. [CrossRef]
26. He, T.; Liu, Y.; Xu, C.; Zhou, X.; Hu, Z.; Fan, J. A Fully Convolutional Neural Network for Wood Defect Location and Identification. *IEEE Access* **2019**, *7*, 123453–123462. [CrossRef]
27. Hu, K.; Wang, B.; Shen, Y.; Guan, J.; Cai, Y. Defect identification method for poplar veneer based on progressive growing generated adversarial network and MASK R-CNN Model. *BioResources* **2020**, *15*, 3041–3052. [CrossRef]
28. Tao, R.; Li, S.; Zhang, X.; Wang, W. Patch-U-Net: A Semantic Segmentation Method for Tree Species Classification Using Remote Sensing Data. *Remote Sens.* **2021**, *13*, 542.
29. Moru, D.K.; Borro, D. Analysis of Different Parameters of Influence in Industrial Cameras Calibration Processes. *Measurement* **2021**, *171*, 108750. [CrossRef]
30. Bürmen, M.; Pernuš, F.; Likar, B. LED Light Sources: A Survey of Quality-Affecting Factors and Methods for Their Assessment. *Meas. Sci. Technol.* **2008**, *19*, 122002. [CrossRef]
31. Jia, L.; Wang, T.; Chen, Y.; Zang, Y.; Li, X.; Shi, H.; Gao, L. MobileNet-CA-YOLO: An Improved YOLOv7 Based on the MobileNetV3 and Attention Mechanism for Rice Pests and Diseases Detection. *Agriculture* **2023**, *13*, 1285. [CrossRef]
32. He, J.; Cheng, Y.; Wang, W.; Ren, Z.; Zhang, C.; Zhang, W. A Lightweight Building Extraction Approach for Contour Recovery in Complex Urban Environments. *Remote Sens.* **2024**, *16*, 740. [CrossRef]
33. Peng, H.; Xue, C.; Shao, Y.; Chen, K.; Xiong, J.; Xie, Z.; Zhang, L. Semantic Segmentation of Litchi Branches Using DeepLabV3+ Model. *IEEE Access* **2020**, *8*, 164546–164555. [CrossRef]
34. Perez, L.; Wang, J. The Effectiveness of Data Augmentation in Image Classification Using Deep Learning. *arXiv* **2017**, arXiv:1712.04621. Available online: https://arxiv.org/abs/1712.04621 (accessed on 17 October 2024).
35. Howard, A.G.; Zhu, M.; Chen, B.; Kalenichenko, D.; Wang, W.; Weyand, T.; Andreetto, M.; Adam, H. MobileNets: Efficient Convolutional Neural Networks for Mobile Vision Applications. *arXiv* **2017**, arXiv:1704.04861. Available online: https://arxiv.org/abs/1704.04861 (accessed on 17 October 2024).
36. Howard, A.; Sandler, M.; Chen, B.; Wang, W.; Chen, L.-C.; Tan, M.; Chu, G.; Vasudevan, V.; Zhu, Y.; Pang, R.; et al. Searching for MobileNetV3. In Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV), Seoul, Republic of Korea, 27 October–2 November 2019; pp. 1314–1324. [CrossRef]
37. Sandler, M.; Howard, A.; Zhu, M.; Zhmoginov, A.; Chen, L.-C. MobileNetV2: Inverted Residuals and Linear Bottlenecks. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Salt Lake City, UT, USA, 18–22 June 2018; pp. 4510–4520.
38. Hou, Q.; Zhou, D.; Feng, J. Coordinate Attention for Efficient Mobile Network Design. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), Nashville, TN, USA, 19–25 June 2021; pp. 13713–13722.
39. Li, Y.; Geng, T.; Stein, S.; Li, A.; Yu, H. GAAF: Searching Activation Functions for Binary Neural Networks through Genetic Algorithm. *Tsinghua Sci. Technol.* **2022**, *28*, 207–220. [CrossRef]
40. Wang, R.; Zhang, G.; Liang, F.; Wang, B.; Mou, X.; Chen, Y.; Sun, P.; Wang, C. WPS-Dataset: A benchmark for wood plate segmentation in bark removal processing. *arXiv* **2024**, arXiv:2404.11051. [CrossRef]
41. Ronneberger, O.; Fischer, P.; Brox, T. U-Net: Convolutional Networks for Biomedical Image Segmentation. In Proceedings of the Medical Image Computing and Computer-Assisted Intervention (MICCAI), Munich, Germany, 5–9 October 2015; pp. 234–241.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.