forests

MDPI

Article

A Novel Wood Log Measurement Combined Mask R-CNN and Stereo Vision Camera

Chunjiang Yu [1], Yongke Sun [1,*], Yong Cao [2,*], Jie He [1], Yixing Fu [1] and Xiaotao Zhou [1]

[1] School of Big Data and Intelligent Engineering, Southwest Forestry University, Kunming 650224, China
[2] International Engineering and Technology Institute, Hongkong 999077, China
* Correspondence: sunyongke@swfu.edu.cn (Y.S.); cncaoyong@swfu.edu.cn (Y.C.)

Abstract: Wood logs need to be measured for size when passing through customs to verify their quantity and volume. Due to the large number of wood logs needs through customs, a fast and accurate measurement method is required. The traditional log measurement methods are inefficient, have significant errors in determining the long and short diameters of the wood, and are difficult to achieve fast measurements in complex wood stacking environments. We use a Mask R-CNN instance segmentation model to detect the contour of the wood log and employ a binocular stereo camera to measure the log diameter. A rotation search algorithm centered on the wood contour is proposed to find long and short diameters and to optimal log size according to the Chinese standard. The experiments show that the Mask R-CNN we trained obtains 0.796 average precision and 0.943 IOU_{mask}, and the recognition rate of wood log ends reaches 98.2%. The average error of the short diameter of the measurement results is 5.7 mm, the average error of the long diameter is 7.19 mm, and the average error of the diameter of the wood is 5.3 mm.

Keywords: wood log measurement; instance segmentation; Mask R-CNN; binocular stereo camera

check for updates

Citation: Yu, C.; Sun, Y.; Cao, Y.; He, J.; Fu, Y.; Zhou, X. A Novel Wood Log Measurement Combined Mask R-CNN and Stereo Vision Camera. *Forests* **2023**, *14*, 285. https://doi.org/10.3390/f14020285

Academic Editor: Angela Lo Monaco

Received: 27 December 2022
Revised: 26 January 2023
Accepted: 30 January 2023
Published: 1 February 2023

Copyright: © 2023 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 log is a bulk commodity and billions of cubic meters of them are consumed in the world every year [1]. There are problems of lack of the actual quantity and small diameter of the wood log than order when imported through customs, which causes economic losses. Therefore, verifying the measurement of the wood log is necessary [2,3]. The traditional wood log measurement is still widely used, it requires the cooperation of two to three people to measure wood, two of them for the measure at each side of the log and the other for recording. This method is not only labor consumption but also inefficient [4]. Moreover, the fatigue level of the operator often influences the accuracy of measurement [5,6].

To overcome the disadvantages of manual measurement, researchers have proposed automated measurement methods, which employ photoelectric, optical, or laser to measure the log size and calculate the volume. For example, a conveyor machine is employed to transfer the logs and measure their size automatically when the wood log pass through the optical instruments [7–12]. Computed Tomography(CT) scanning is also a common technique used in wood measurement. Longo used CT to scan fresh Douglas-fir (*Pseudotsuga menziesii* (Mirb.) Franco) logs, then detected and measured knots using the images obtained from these scans, and obtained good localization accuracy and measurement precision [13]. Gergel used CT techniques to scan the wood log of three species of wood: oak, beech, and spruce, reconstruct their three-dimensional structures, and then calculated their volumes using four methods, and find that a method named STN 480009 has the best performance that the error of less than 0.01 m^3 [14]. Although these methods achieve satisfactory detection accuracy, it is hard to be used widely because the measurement equipment is too large to be portable.

For the purpose to improve the portability and efficiency of the measurement, researchers attempted to utilize computer vision to automatically detect wood and measure

its size. For instance, Yan presented a method using a mobile phone with a rangefinder to measure the wood log and the measurement accuracy reached 98.2% [15]. Kruglov employed a camera to acquire the wood image, used an image segmentation method to detect the 3D structural information to measure the size of the wood log, and the error was only 4.8% compared to manual measurement [16]. Yurii developed a conveyor-tracking system that extracted wood images and measured the wood size from video sequences, and decreased the minimum mean square error that only 0.045 ± 0.041 [17]. These image-based methods are more portable than labor measurement, achieved high detection accuracy, and improved the efficiency of the measurement. However, it is struggled to measure the wood log in a complex environment because it is hard to extract good wood image features from the complex background.

The convolutional network-based methods have a good ability to extract high-quality object features in complex environments and have been widely used in various scenarios in the forest. For example, Ting used a DCNN model to detect the wood defects of *Pinus*, and the accuracy reached 99.13% [18]. Gao used a TL-ResNet34 model that integrated the Reset-34 model and the transfer learning method to classify seven different wood knots, and the accuracy reached 98.69% [19]. Hao proposed a method based on the single shot multi-box detector (SSD) [20] model to detect the log end of the wood in a natural scene and obtained 94.87% accuracy and 91.34% recall [21]. Samdangdech used a fully convolutional network(FCN) [22] to segmented the wood log of *Eucalyptus*, and achieved an average accuracy of 94.45% in log segmentation and 2.71% of false negative [23]. LIN combined the YOLOv3-tiny model [24] and Hough [25] transform method to detect the wood cross-section and calculate its area. The positive detection rate as high as 98.79% and the negative detection rate was only 0.602% [26]. Wang transformed the point cloud map of rubber trees obtained from mobile LiDAR scanning into a depth map, used Faster R-CNN to segment them and detect the rubber trees, and obtained a segmentation accuracy of 98% [27]. Luo proposed an improved Faster R-CNN algorithm for tree detection in mining areas and obtained 89.89% AP and 91.61% accuracy in tree detection [28]. Lin proposed an improved YOLOv4-Tiny network and K-median clustering algorithm to detect bundled log ends, and precision, recall, and the F1-score reached 93.97%, 95.34%, and 0.95 respectively on the test set [29]. Fang used YOLOv5 to detect surface knots on sawn timbers and obtained f-scores of 91.7% and 97.7% on the two datasets, respectively [30]. These methods have a good quality of detecting the wood defect and log end. However, These one-stage detection methods based on bounding boxes cannot accurately obtain the precise contours of logs, and the detection method based on Faster R-CNN still uses ROI Pooling, which leads to the loss of translational invariance of the subsequent network features and affects the final localization accuracy of wood contour [31].

Mask region-based convolutional neural network (Mask R-CNN) [32] is a classic instance segmentation network, which combines object detection and semantic segmentation methods. Recent research shows that this method has been widely used in many field and remote sensing scenarios and has obtained good performance. Ling used CutPaste-Based Self-Supervised learning module to improved Mask R-CNN to detect ships in remote sensing images, compared with the original model, the mAP improved by 17.8%, and the detection accuracy of small target objects improved by 22.8% [33]. Zhang combined the Sobel edge detection algorithm with Mask R-CNN for the segmentation of buildings in high spatial resolution remote sensing images, and the average value of IOU (intersection over union) for the proposed method was 88.7% and the average value of Kappa was 87.8% [34]. Liu proposed an improved Mask R-CNN to detect cracks in ground penetrating radar (GPR) images and measure their widths, and obtained a measurement error of 2.33%, a segmentation accuracy of 0.833, an F1 score of 0.822, and a mean intersection-union (mIoU) of 0.701 with a processing speed of 4.2 frames per second [35]. Zhou used genetic algorithm combined with the gradient descent method to optimize the parameters of the Mask R-CNN model to insulator fault-identification, the average accuracy of 98% and frames per second (FPS) of 5.75 [36]. Zhang proposed an improved Mask RCNN model

for segmentation and statistics of Unmanned Aerial Vehicle(UAV) image trees in mixed forests and obtained the highest overall accuracy of 90.13% with an average statistical error of 5.11% [37]. Hao trained a Mask R-CNN model to detect discontinuous tree crowns and height of *Cunninghamia lanceolata* in a plantation in China, in which, six different bands of LiDAR data were used to detect the tree, and the F1-score reached 84.68%, the Intersection over Union (IoU) of tree crown reached 91.27% [38]. Hu added a multiscale receptive field block module on the Mask R-CNN model to monitor *pine* diseases in forests. The precision, recall, and F1-score increased by 22.4%, 3.5% and 14.4% respectively [39]. Li used a cycle generative adversarial network (GAN) [40] to augment the wood defect dataset and constructed a layered deformable Mask R-CNN model to detect and segment the knots, cracks, and worms in *Betula davuric*, *Pinus* and *Populus L* species. The detection and segmentation precision reached 84.4% and 82.8%, respectively [41]. Shi integrated a glance network and a multiple channel mask R-CNN model to detect the wood veneer defects, with an accuracy of 98.7% and the precision of 95.31% [42]. Although the backgrounds, lights, and contrasts are different in these experiments, the methods based on Mask R-CNN still have good performance in segmenting subjects from wood images. It indicates that the Mask R-CNN is a feasible method to detect the wood contour in a complex environment.

Depth cameras based on binocular stereo vision have been widely used to get sizes from the image and measure various objects. For example, Chen used the Oriented FAST and Rotated BRIEF (ORB) [43] feature point detection method to detect the log edges and proposed an algorithm based on binocular vision to measure the log end of the wood, the error is less than 2 mm [44]. Zheng used a binocular stereo camera to measure the diameter and length of the vegetables, and the mean absolute percentage error (MAPE) less than 8% [45]. Suo utilized binocular cameras to estimate the fish's length, and the mean of the error was only 5.58% [46]. Solak proposed a new triangulation method based on binocular cameras, which adding in a look-up table and curve-fitting to calculate the distance between a robot and an object. An average accuracy of 97.69% and average accuracy rates of 98.24% for Manhattan and 98.03% for Euclidean distances were obtained in the experiments [47]. Huang proposed an improved semi-global matching (SGM) algorithm based on least-squares fitting interpolation to obtain the disparity map of binocular cameras. It was then applied to obstacle distance measurements on high-voltage transmission lines, and a measurement error of less than 5% was obtained between 0.5 m and 5.0 m [48]. Adil proposed a Python-based algorithm to find the parameters of the binocular camera, create disparity maps, and use these maps for distance measurements. Experiments showed that 99.83% measurement accuracy was obtained at 100 cm, with a processing time of less than 0.355 s [49]. These studies explored an approach to measuring objects using a binocular stereo camera and supported a feasible way to measure the wood log using the image detection method with a binocular stereo camera.

Based on the successful results of previous studies, we propose an improved automatic wood measurement model combining Mask R-CNN and a binocular stereo camera to achieve fast and accurate wood measurement tasks in complex environments. The specific contributions are as follows:

1. We proposed a log diameter detection method conforming to Chinese wood measurement standards that takes the circumcircle center of the wood contour as the contour center and uses a rotational search to obtain the long and short diameters of the wood contour. The method can quickly obtain the long and short diameters from irregular wood cross-sections with improved accuracy.
2. We proposed a novel log diameter measurement method that uses a Mask R-CNN instance segmentation model and binocular camera to automatically extract wood log contours and calculate the real wood log size, improving the measurement efficiency.

The robustness of the method is enhanced with better segmentation against complex backgrounds, overlapping and shading of wood, and uneven ground color, enabling fast and accurate calculation of wood diameter. This study provides a new idea to achieve fast measurement of wood cross-sections in complex environments.

2. Materials

The wood images contained two parts and were collected from two different sources. The first part of images comes from a public dataset [50] and download from https://deepai.org/publication/the-hawkwood-database (accessed on 4 May 2022). A series augmentation methods, such as affine transformations, perspective transformations, contrast changes, Gaussian noise, dropout of regions, hue/saturation changes, cropping/padding, blurring are used to augment the dataset [51]. Each image contains multiple logs, as shown in Figure 1, there are 1967 images in this part, including 19,436 logs.

The first part was used to train log end detection model, and divided into three subsets, training dataset, validation dataset and test dataset with ration of 8:1:1. Finally, the training dataset has 1573 images, the validation dataset has 197 images and test dataset has 197 images.

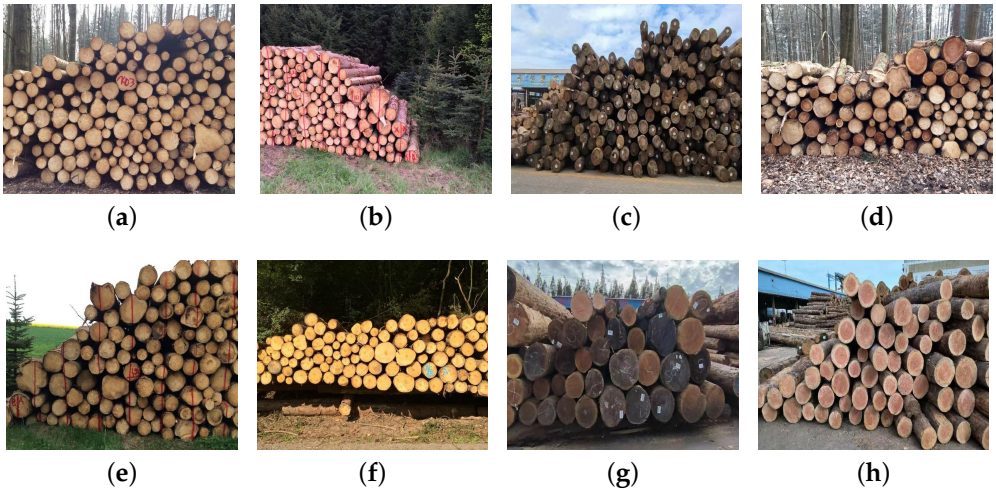


Figure 1. Wood log images in different environment. (**a**,**e**). Different backgrounds.(**b**,**f**). Different shooting angles. (**c**,**g**). Different color textures. (**d**,**h**). Different lighting.

The second part of the images was acquired from wood samples in the wood specimen laboratory of Southwest Forest University. The majority of the tree species is *eucalyptus*, as Figure 2 shows. According to the Chinese wood log measurement standard, we manually measured the short diameter and long diameter for every samples [52]. These samples are used to evaluate the accuracy of our proposed measurement method. The Chinese measurement standard is different from other measurement standards in other countries, which requires to measure the short diameter first and measure the long diameter later.

A binocular stereo camera, named ZED, as shown in Figure 3, manufactured by Stereolabs company, is employed to capture wood images and distance from the camera to the wood log. A Linux system with GPU RTX 3090 was used to train the Mask R-CNN model to detect the wood log contours.



(**a**) (**b**)

Figure 2. Wood stacks used in measurements. (**a**). Wood stacks indoors. (**b**). Wood stacks outdoors.



(**a**) (**b**)

Figure 3. (**a**). Long and short diameter measurement schematic, d_1 is the short diameter, and d_2 is the long diameter. (**b**). ZED camera.

3. Methods

Our method measures the wood log size with three processes, as Figure 4 shows. First, employing a stereo camera to capture the wood log images, and detecting the contours of the log end with a trained Mask R-CNN model. This Mask R-CNN model is trained using the image dataset. Second, detecting the positions of short diameter and long diameter using our proposed method. Third, Calculating the length of the short and long diameters using the depth and difference of the images from the binocular camera.

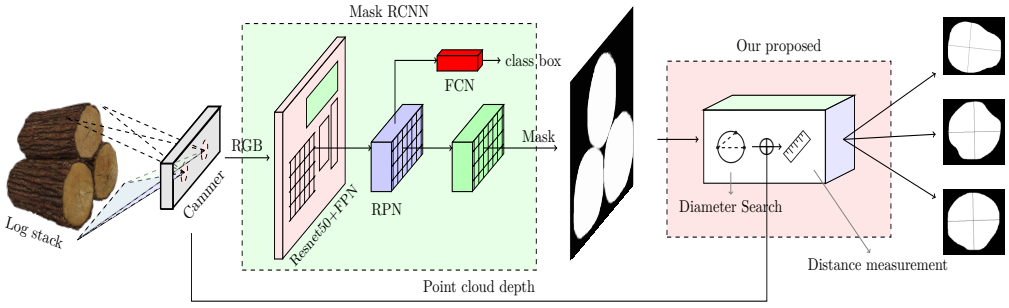


Figure 4. Processes of the wood measurement.

3.1. Mask R-CNN Algorithm

Mask R-CNN is a mask-based method that added a mask prediction branch to Faster R-CNN and achieved contour segmentation. It uses ResNet50 [53] to extract a features map from the input image, and feature pyramid network (FPN) [54] to build a pyramid of features in different sizes, and retrieves pre-selection objects from each level of the pyramid. Then, the RPN network receives the extracted feature maps from FPN and outputs the ROI, i.e., the candidate regions of objects. The ROI Algin layer employs bilinear interpolation

to complete pixel alignment of the ROI.Finally, these ROIs are sent to three branches for category recognition, bounding box regression, and mask construction.

When a wood log image is processed through the Mask R-CNN model layers, the Resnet layers are responsible for extracting the image features, the FPN layers selects the contour features in the different pyramids, and the RPN layer is used to choose candidate areas of the log end. The final extracted features will be fed into three branches, two of which perform classification and regression, and the third branch generates a binary mask image of the object using FCN. The output of the mask R-CNN is the contours of each log end.

3.2. Diameter Search Algorithm

We propose a method to detect the short diameter and the long diameter. The wood contour is shown in Figure 5, defining the point dataset of the wood contour as $P = \{p_1, p_2, \ldots, p_n\}$, and getting the center $o(o_x, o_y)$ of the circumcircle of the contour by Formulas (1) and (2).

$$(o_x, o_y) = (\frac{1}{2}(p_{ix} + p_{jx}), \frac{1}{2}(p_{iy} + p_{jy})) \tag{1}$$

$$(p_i, p_j) = max\{d(p_i, p_j), p_i, p_j \in P\} \tag{2}$$

where $d(p_i, p_j)$ is the Euclidean distance, and the line segment $p_i p_j$ is shown in Figure 5.

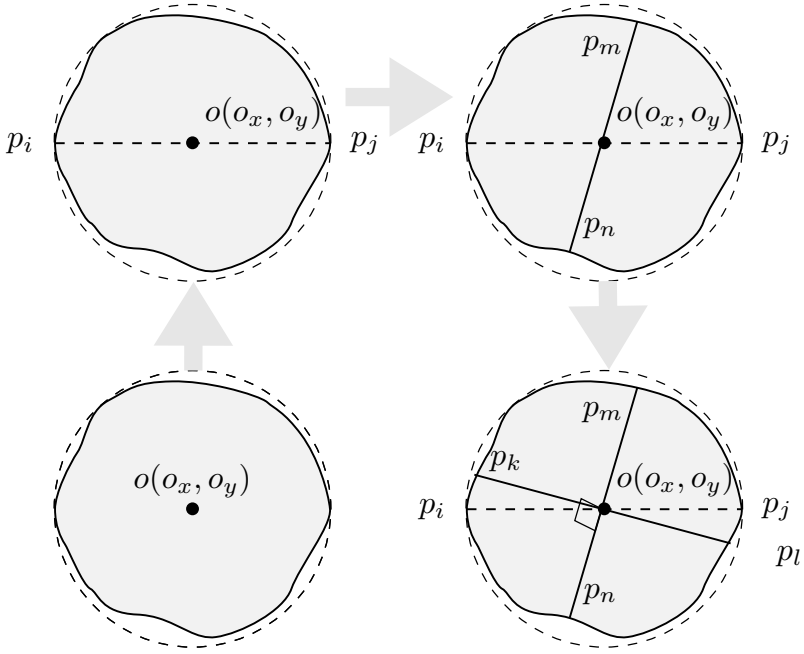


Figure 5. Long and short diameter search process.The line segment $p_i p_j$ is the diameter of the circumcircle of the contour. The line segment $p_m p_n$ is the short diameter of the wood which passes through the center $o(o_x, o_y)$ and has the minimum length. The line segment $p_k p_l$ is perpendicular to the straight line $p_m p_n$ which passes through the center is considered as the long diameter.

Define the linear equation that pass through the center $o(o_x, o_y)$ as the Formula (3).

$$y = k(x - o_x) + o_y \tag{3}$$

The line segment has two cross points p_m and p_n with the contour of the wood, and $p_m p_n$ is the short diameter, where p_m, p_n must statisfied Formula (4), and $p_m p_n$ is as shown in the Figure 5.

$$(p_m, p_n) = min\{d(p_m, p_n), p_m, p_n \in P\} \tag{4}$$

Then, the equation for a line perpendicular to the short diameter and passing through the center can be expressed as Formula (5).

$$y = -\frac{1}{k}(x - o_x) + o_y \tag{5}$$

Similarly, the straight line expressed by Formula (5) has two intersections p_k and p_l with the wood contour, and the long diameter is formed by the points p_k and p_l, where p_k, p_l must satisfied the Formula (6). The Figure 5 shows the straight line $p_k p_l$.

$$(p_k, p_l) = max\{d(p_k, p_l), p_k, p_l \in P\} \tag{6}$$

Through the above algorithm, we find the specific short diameter $p_m p_n$ and long diameter $p_k p_l$, and define them as d_s and d_l respectively. According to Chinese wood measurement standards, the wood's diameter d is determined by its long and short diameters. The specific calculation is shown in the Formula (7).

$$d = \begin{cases} d_s, & (d_l - d_s) \leq 2 \text{ cm} \\ (d_l + d_s)/2, & (d_l - d_s) > 2 \text{ cm} \end{cases} \tag{7}$$

The specific process is as follows:

1. Obtain the binary mask images of wood by feeding the wood image into the trained Mask R-CNN model;
2. Fit the circumcircle of the wood contour to get the rotation center;
3. Calculate the next rotation point every five degrees in circumcircle successively by using a point on the wood's contour as the beginning point;
4. Connect each point to the center point to generate a straight line, and extend the line to the intersect of the wood contour other side as another point. There is a line segment between the point pair.
5. Compare the pixel length of each line segment and select the shortest one as the short diameter of the wood contour.
6. Calculate the length of the line segment perpendicular to the short diameter which passes through the wood contour center. If the pixel length is not an integer, the neighbor points are used to calculate the length and the maximum length line segment is treated as the long diameter.
7. Output the pixel coordinates of the long diamter and short diamter.

The specific pseudocodes of the diameter search algorithm is shown as Algorithm 1.

3.3. Distance Measurement Algorithm

The binocular stereo camera is similar as the human eyes. It is different from the depth camera based on TOF and structured light principles, and it does not actively project light sources to the outside. Instead, it solely relies on two captured images obtained by the left and right cameras to calculate the depth. Therefore, it is sometimes referred to as a passive binocular depth camera.

To build the camera imaging geometry model, we usually need to calibrate the camera [55] to get the internal and external parameters and complete the correction of the camera before using the binocular camera. Since this study uses the ZED series camera, which has been calibrated when it come out, we can skip this step.

In the diameter search algorithm, we get the pixel coordinates of the two endpoints of the short diameter. To convert the pixel coordinates into world coordinates and calculate the actual distance between two points, we must get the depth information from the camera and combine it with different coordinate systems to establish an affine model. Finally, use the model to complete the transformation of coordinates and obtain the actual distance.

Algorithm 1 Short diameter finding.

Require: P, P *is an array of wood mask contours*
Ensure: short diameter
1: $P = ModelPredictor(wood_image).mask$
2: $center = minEnclosingCircle(P)$
3:
4: $points \leftarrow []$
5: $startPoint = P[0]$
6: **if** $(nextPoint - startPoint)'s$ *degrees* > 5 **then**
7: *add nextPoint to points*
8: **end if**
9:
10: **function** GETSHORTDIAMETER($Poins, P$)
11: **for** *point in points* : **do**
12: 1. *Calculate the slope*
13: 2. *Find a reverse extension cord*
14: 3. *Get diameter*
15: **end for**
16: **return** diameter of the shortest distance
17: **end function**

The principle of the binocular camera to obtain depth information is shown in the Figure 6. Where B is the distance between the two cameras, o_r and o_t are the optical centers of the two cameras respectively, f is the focal length, p is a real-world point, p' and p'' are two points on the imaging plane of the camera, and x_r, x_t are the horizontal coordinates of the imaging points p', p''.

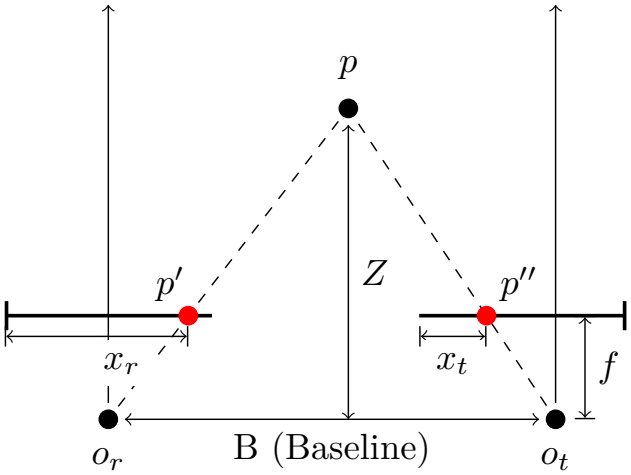


Figure 6. Binocular stereo vision system.

Assume that the left and right cameras are parallel, so the y-value at point p' and p'' are the same, and the distance of $p'p''$ can be calculate as the Formula (8).

$$dis = B - (x_r - x_t) \tag{8}$$

From the Figure 6, we can establish the similar triangle relationship between the $pp'p''$ and po_ro_t, and it is shown as the Formula (9).

$$\frac{B - (x_r - x_t)}{B} = \frac{Z - f}{Z} \tag{9}$$

Then, we can further derive a Formula (10) to calculate the depth Z.

$$Z = \frac{fB}{x_r - x_t} \tag{10}$$

In the Formula (10), the focal length f and the baseline B can be obtained from the parameters calibrated by the camera.The $x_r - x_t$ is called disparity, which can be obtained by stereo matching.

Now that we have obtained the depth information Z. Suppose we have a point m in the imaging plane, then there must exist a point A in the world coordinate system corresponding to point m, this is as shown in the Figure 7. Therefore, we can further construct similar triangle models between camera coordinates, imaging coordinates, and point A.

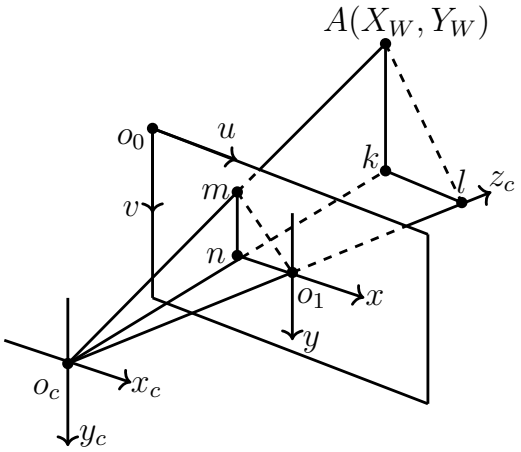


Figure 7. Similar Triangle Model.

Assuming the pixel coordinate of point m is (u_1, v_1), we should first convert it to image coordinates by using the following Formula (11).

$$\begin{cases} u = \frac{x}{d_x} + u_0 \\ v = \frac{y}{d_y} + v_0 \end{cases} \tag{11}$$

Where the d_x,d_y are the ratio of pixels to actual size, then the image coordinates $((u_1 - u_0) * d_x, (v_1 - v_0) * d_y)$ of the point m can be gotten.

In Figure 7, we can establish the similar triangle relationship between mo_cn and Ao_ck, no_co_1 and ko_cl,mo_co_1 and Ao_cl, and get the Formula (12).

$$\frac{o_co_1}{o_cl} = \frac{no_1}{kl}, \frac{o_co_1}{o_cl} = \frac{o_cm}{o_cA} = \frac{mn}{Ak} \tag{12}$$

Then, we can derive the Formula (13).

$$\frac{f}{Z} = \frac{(u_1 - u_0) * d_x}{X_W}, \frac{f}{Z} = \frac{(v_1 - v_0) * d_y}{X_Y} \tag{13}$$

Define f_x, f_y are the focal length in pixels, and d_x,d_y are the ratio of pixels to actual size, then the corresponding relationship between f_x, f_y and d_x,d_y are shown in the Formula (14).

$$f_x = \frac{f}{d_x}, f_y = \frac{f}{d_y} \tag{14}$$

Combining Formulas (13) and (14), we can derive the world coordinates of point A as shown in Formula (15).

$$\begin{cases} X_W = Z * \frac{u_1 - u_0}{f_x} \\ Y_W = Z * \frac{v_1 - v_0}{f_y} \end{cases} \tag{15}$$

Assuming that point A is an endpoint of the wood contour short diameter, the short diameter must exist another endpoint B, as shown in the Figure 8. Similarly, point B has a corresponding point n in the imaging plane, and we can hypothesize that the image coordinate of point n is (u_2, v_2). Then, we can also derive the world coordinate of point B by using Formula (15). Finally, we use the world coordinates of point A and B to calculate their Euclidean distances by Formula (16) as the actual value of short diameter.

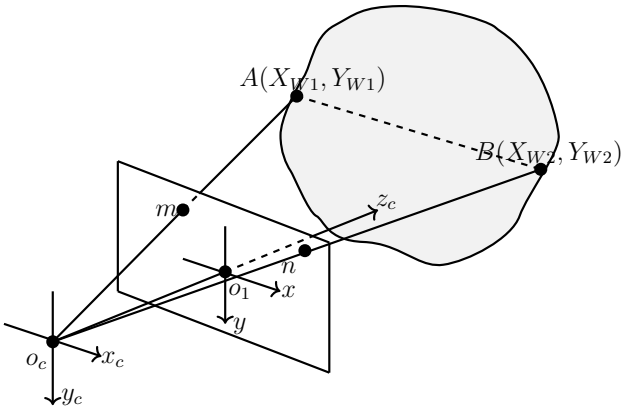


Figure 8. Wood contour in the coordinates.

$$l_{AB} = Z * \sqrt{\frac{(u_1 - u_2)^2}{f_x^2} + \frac{(v_1 - v_2)^2}{f_y^2}} \tag{16}$$

Therefore, in this way we can also calculate the actual distance of the long diameter by obtaining the pixel coordinates of the two endpoints of the long diameter and applying Formula (16).

3.4. Evaluate

In this paper, we evaluate the performance of our proposed wood measurement method in two aspects.

3.4.1. Mask R-CNN Model Evaluation

To evaluate the accuracy of our trained Mask R-CNN model for wood contour recognition and segmentation, we use average precision (AP) and recall rate as the evaluation metrics.

Define True Positive (TP), the number of samples inferred to be positive, and in fact the number of positive samples; False Positive (FP), the number of samples that are inferred to be positive, but are in fact negative; False Negative (FN), the number of samples that are inferred as negative but are in fact positive. Consequently, the following Formulas (17) and (18) can be used to calculate precision and recall rate:

$$precision = \frac{TP}{TP + FP} \tag{17}$$

$$recall = \frac{TP}{TP + FN} \tag{18}$$

Then, we take the recall rate and precision under different confidence levels and establish different recall thresholds, choose the maximum precision value under the associated recall threshold, draw the P-R curve, and calculate the area contained by the curve, which

is the AP value. Another metric called mAP is the mean of AP under multiple categories. However, in this paper, just one type of wood needs to be predicted by Mask R-CNN so that the AP metric can satisfy our needs.

Although the AP metric can evaluate the model performance, it is more suited to evaluating the categorization, and the real segmentation quality of the mask cannot be examined. Hence the IoU_{mask} is introduced to evaluate the segmentation quality of the mask, i.e., and the quality of the segmented wood contour. Formula (19) shows how IoU_{mask} is calculated.

$$IoU_{mask} = \frac{area(P) \cap area(G)}{area(P) \cup area(G)} \tag{19}$$

In the Formula (19), P is the area of the wood contour output by the model, G is the area of the wood contour manually labeled. This formula calculates the ratio of the overlap of the two areas to the total area of both. So, the IoU_{mask} value can precisely evaluate the segmentation accuracy of the model on the wood contour.

3.4.2. Long and Short Diameter Measurement Comparison in Actual

We define the automatically measured short diameter as S_a, the long diameter as L_a, the manually measured short diameter as S_m, and the long diameter as L_m, with the number of measurement samples as n. The average error of short diameter is defined as error1, and the average error of long diameter is defined as error2, as shown in Equations (20) and (21).

$$error1 = \frac{1}{N}\sum_{i=1}^{N}(S_a - S_m) \tag{20}$$

$$error2 = \frac{1}{N}\sum_{i=1}^{N}(L_a - L_m) \tag{21}$$

In addition, we calculated the standard deviation and Root Mean Square Error(RMSE) separately, and the equations are shown in (22) and (23).

$$\sigma = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(x_i - \mu)^2} \tag{22}$$

where N is the number of samples, x_i is the specific measurement, μ is the mean value of short diamter or long diameter.

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(Y_i - Y_j)^2} \tag{23}$$

where Y_i is the manually measured value, and Y_j is the result of the automatic measured value by our method.

4. Results and Discussions

4.1. YOLO vs. Mask R-CNN

We execute the experiment with the YOLOv5 [56], YOLOv6 [57], and YOLOv7 [58] object detection models respectively. When training the model, we set the super-parameters uniformly, the training epochs are 300, the batch size is 16, the optimizer is SGD, the initial learning rate is 0.01, and the CosineLRScheduler learning rate scheduling strategy is used. The Table 1 shows the scores of the evaluation metrics obtained from the different models. The experimental results indicate that, for the same training configurations, YOLOv5 obtained the best AP metric of 0.903 and AP50 metric of 0.990, and compared to YOLOv6, YOLOv7, YOLOv5 has better detection performance for wood cross-sections.

Table 1. Detection performance of different YOLO frameworks.

Model	AP	AP50
YOLOv5	**0.903**	0.990
YOLOv6	0.683	0.760
YOLOv7	0.867	**0.993**

Therefore, we use the model trained by YOLOv5 to detect the bounding box of the wood, for each bounding box, we apply the Hough transform to perform a circular fit on the wood contour, and the fit results are shown in the Figure 9, the YOLOv5 detect results are shown in the fourth column of Figure 10.

The Figure 9 shows that the Hough transform cannot fit the wood contour well to the circle, and there are many wrong fittings. For example, some fittings are too large or too small, and even some nonsensical fittings that do not correspond to wood cross-sections appear. If the result of the Hough transform is used as the wood contour, then the final measurement is bound to have a large error. The reason for the poor effect of the Hough transform is that the Hough transform requires more parameters to be configured, and different parameters need to be configured for different targets. When performing circle transformations on multiple targets, it is difficult to perform better circular transformations on multiple targets with the same set of parameter settings.



Figure 9. Results of Hough transform and some misfitting.

Even though the YOLOv5 can recognize the object quickly and with a high accuracy, it can only find the rectangle of each object and need other algorithms to find a more accurate contour. So, we choose the Mask R-CNN instance segmentation framework to execute a precise search for wood contours.

4.2. Segmentation and Detection Results for Wood

Table 2 shows the experimental results of the Mask R-CNN model with the different backbone networks and whether to use data augmentation. To verify the performance of the model, we compute the AP segment metrics and IoU_{mask} scores of the four models on the test set.

By Comparing ResNet50 and ResNet101 as the backbone network of the Mask R-CNN, we find that the APs and APl metrics obtained using ResNet101 as the backbone network are higher than those obtained by ResNet50. However, on the contrary, the APm metric shows a decline, indicating that the detection ability of ResNet50 for medmedium-sizedium objects is more potent than that of ResNet101, which may be the ResNet101 network is deeper than ResNet50 and the features extracted for medmedium-sizedium objects are worse. By comparing the models trained with augmented and unaugmented dataset, we find that the evaluation metrics obtained have improved after using the augmented dataset. In particular, using ResNet101 as the backbone network combined with augmented dataset,

the APs and APl metrics are improved by 1%, and received the highest IoU_{mask} score of 0.943. This shows that the model trained with resnet101 and the enhanced dataset has the best segmentation effect on wood cross-sections.

Table 2. AP segment performance.

Methods	AP	APs	APm	APl	IoU_{mask}
Resnet50 Ori	0.790	0.705	0.699	0.822	0.938
Resnet101 Ori	0.788	0.711	0.675	0.827	0.937
Resnet50 Aug	0.793	0.701	**0.702**	0.823	0.943
Resnet101 Aug	**0.796**	**0.723**	0.652	**0.847**	**0.943**

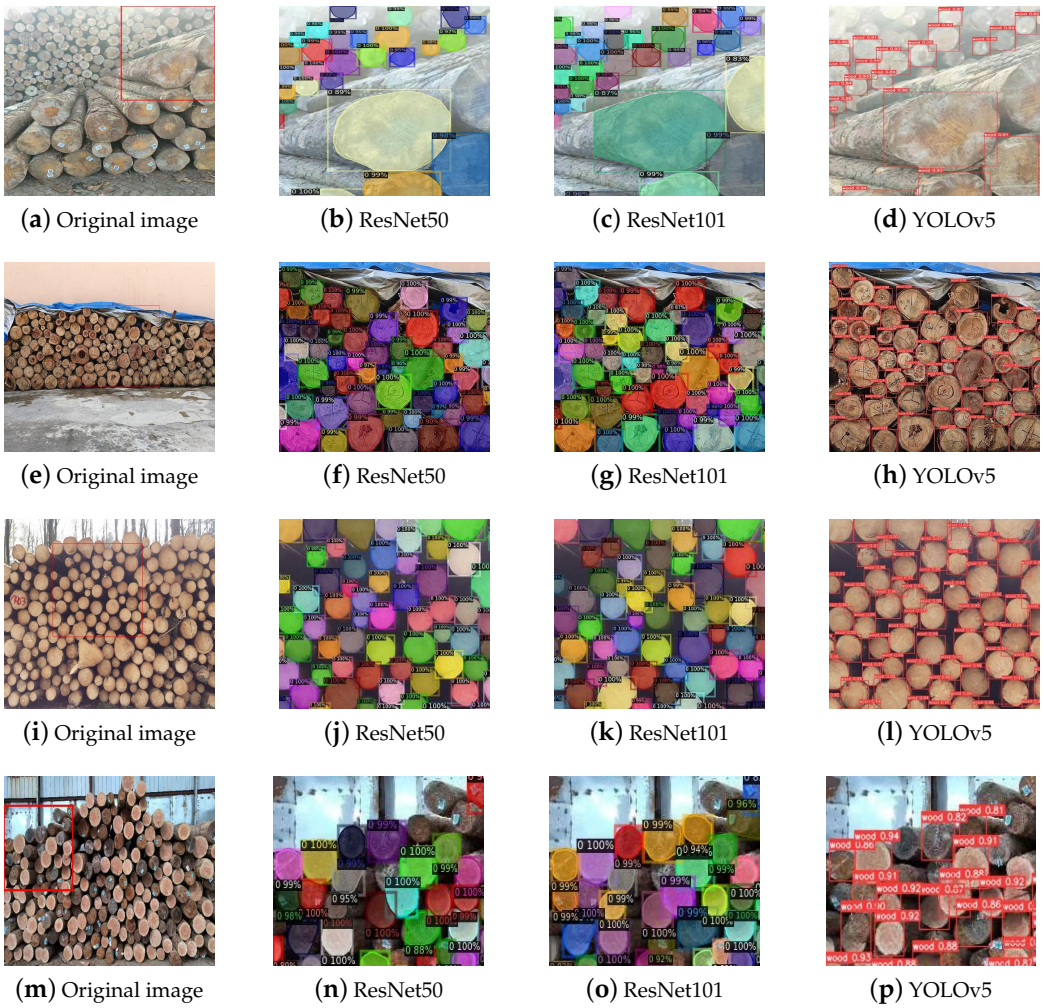
Ori: Model training using the original dataset, Aug: Model training using the augmented dataset. APs, APm, APl, respectively, indicate the AP value of different wood cross-section size, APs (area < 32 pixel × 32 pixel), APm (area < 32 pixel × 32 pixel < area < 96 pixel × 96 pixel), APl (area > 96 pixel × 96 pixel).

We simultaneously feed new wood cross-section images in different scenes to three models: ResNet50 trained with an augmented dataset, ResNe101 trained with an augmented dataset, and YOLOv5, and count the number of woods detected by different models.Finally, the Table 3 gives detection results and the detection effects of the three models are shown in Figure 10. The model trained using the augmented dataset combined with resnet101 as the backbone network identified the most logs, and the wood detection rate reached 98.2%. In contrast, using resnet50 as the backbone network identified fewer logs, and the wood detection rate reached 96.1%. It shows that the effect of using resnet101 as the backbone network is better. Moreover, when we use the YOLOv5 model, the detection performance is the worst, and the wood detection rate is only 90%. By comparing the four images, we found that the detection accuracy of images e, i, and m is similar, while the detection accuracy difference of image a increases, which is where the main difference between the YOLO framework and the Mask R-CNN framework appears. We found that the background of image a is more blurred, and the color difference is the largest of the four images compared to the original dataset, which indicates that the generalization performance of Mask R-CNN is better. The above findings also indicate that Mask R-CNN model can better detect wood in images.

In conclusion, data augmentation positively impacts the performance of the model, and the deeper resnet101 backbone network is also positively impacted in the identification and segmentation of wood cross-sections. Therefore, we adopt the model trained on resnet101 and the augmented dataset as the final detection and segmentation model.

Table 3. Different model detect and segment result.

Images	Total	ResNet50	ResNet101	YOLOv5
Image a	147	135	142	112
Image e	140	137	139	138
Image i	202	197	199	201
Image m	183	177	180	155
Total	672	646	660	606
Detection rate		96.1%	98.2%	90.1%



(**a**) Original image (**b**) ResNet50 (**c**) ResNet101 (**d**) YOLOv5

(**e**) Original image (**f**) ResNet50 (**g**) ResNet101 (**h**) YOLOv5

(**i**) Original image (**j**) ResNet50 (**k**) ResNet101 (**l**) YOLOv5

(**m**) Original image (**n**) ResNet50 (**o**) ResNet101 (**p**) YOLOv5

Figure 10. Different detect images.ResNet50, ResNet101 denote the different backbone networks used by Mask R-CNN.

4.3. Analyse of Diameter Search

In previous studies, ellipse fitting or circle fitting is usually performed on a wood contour, and then the long and short diameters of the fitted ellipse or diameter of the circle are used as the long and short diameters of the wood cross-section [26,44]. Because the cross-section of the wood does not always appear oval or round, this can produce a large error.

Therefore, we proposed a new long-short path search algorithm. We fitted the incircle and circumcircle circles for the wood contour separately and use their centers as the center of the wood. Thus, our algorithm has two different long and short diameter center search methods.Then we compare the long and short diameters obtained by different rotation centers. In most cases, the method of long and short diameter search and computation using the circumcenter are superior to the method by using the inscribed center.

We plot the long and short diameters search and computation results by the different rotation centers, as shown in Figure 11. The binary images in the third and fourth columns in Figure 11 are predicted by Mask R-CNN. We extract the contour information of the wood in the binary images by opencv and apply our long and short diameter search algorithm to find the long and short diameters. We discover that when the wood cross-section is reasonably regular, the difference between the long and short diameters determined by different rotation centers is extremely modest. However, when the cross-section has an irregular form, the difference is substantial, and Figure 12 shows this phenomenon.

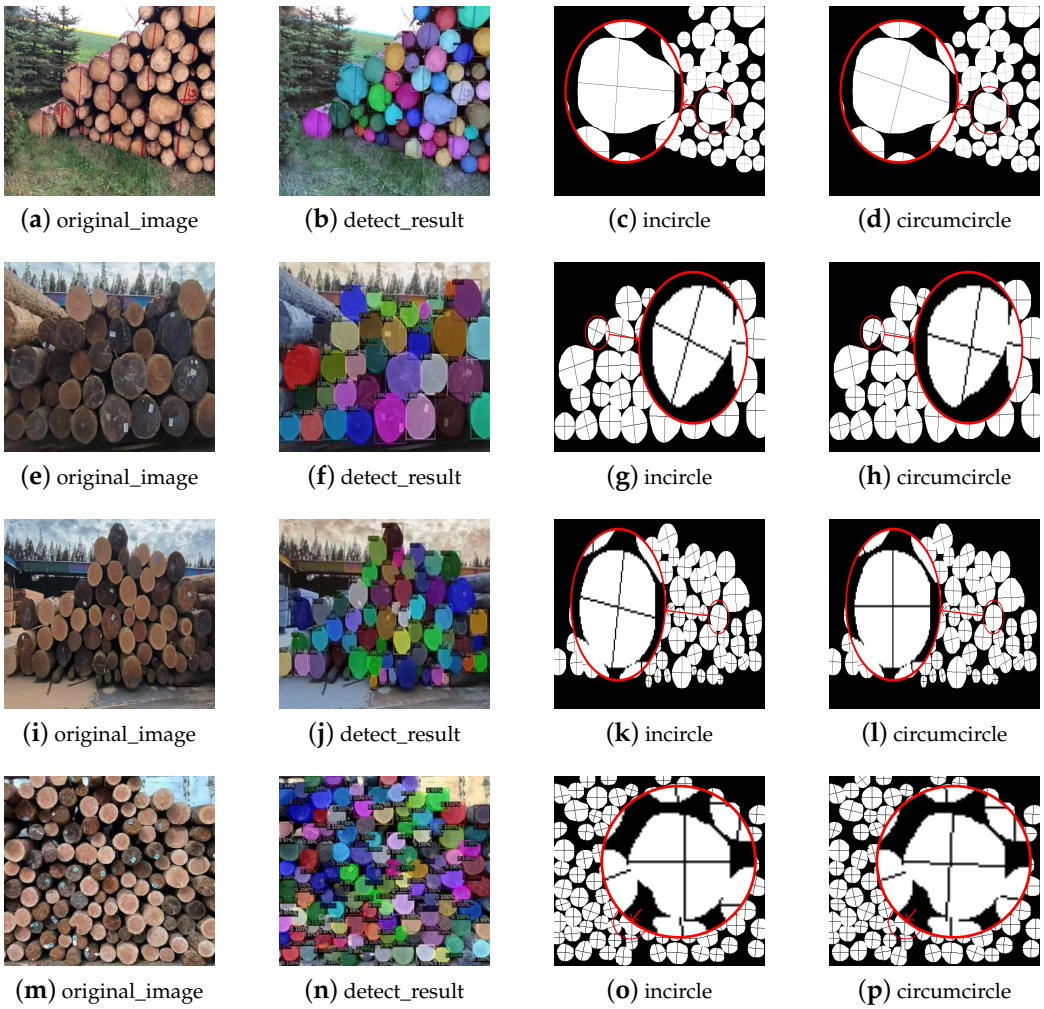


Figure 11. Different ratio search. Incircle:Using the inscribed center as the ratio search enter. Circumcircle: Using the circumcenter as the ratio search enter.

Figure 12a–c shows the long and short diameters found using the incircle center, while Figure 12d–f shows the long and short diameters found using the circumcenter. We found that in these irregular wood contours, the short diameter determined by the incircle center has a large error, which leads to a subsequent error in the long diameter, while the long and short diameters determined by the circumcenter are more accurate. The long and short diameters calculate by circumcenter as the center are more in line with the manual measurement habits. Therefore, the circumcenter can be utilized as the circle center in subsequent measurement to determine the long and short diameters.

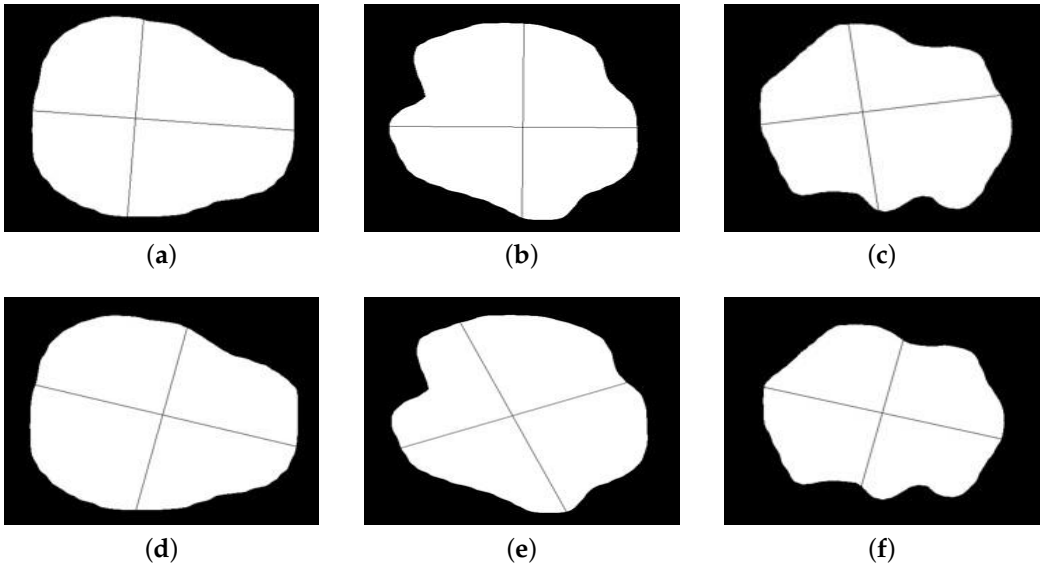


Figure 12. Long and short diameters of irregular timber sections. (**a**–**c**) calculate by incircle as the center. (**d**–**f**) calculate by circumcenter as the center.

4.4. Analyse of Actual Measurement of Wood

We place the cameras parallel to the log stacks at 1.5 m, 2 m and 2.5 m, respectively. Then combine the Mask R-CNN model with an improved search algorithm for automatic measurement, and the measured results are shown in the Table 4. We measure 59 wood cross-sections and calculate the long and short diameter error, standard deviation, and root mean square error (RMSE) respectively by using the manually measured data as the benchmark.

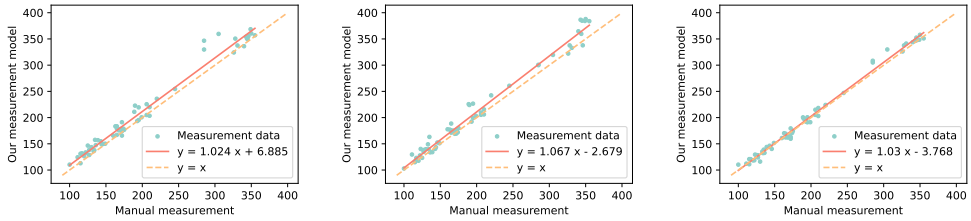
Table 4. Measured Result.

Distance	Samples	Error1	Error2	Standard Deviation1	Standard Deviation2	RMSE1	RMSE2
1.5 m	59	12.32	15.11	12.87	14.92	17.3	20.5
2 m	59	12.01	14.79	12.86	15.92	16.77	19.9
2.5 m	59	5.70	7.19	7.30	8.47	7.67	9.1

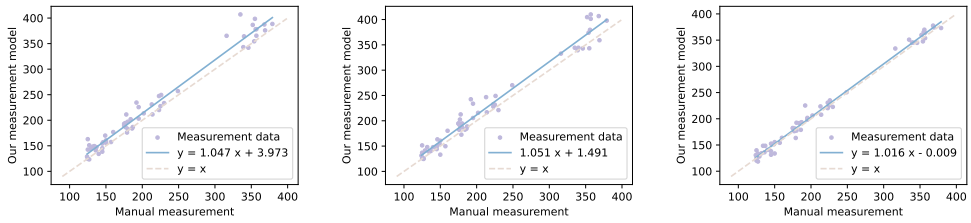
Error1, Error2: The average error of the short or long diameter is measured automatically. RMSE1, RMSE2: The mean square error of the measured short diameter or long diameter.

The errors of the short and long diameters measured by the camera at different distances is compared. The results show that at a distance of 2.5 m, the average error between the manual measurement and the automatical measurement reached the minimum, with a difference of 5.7 mm for the short diameters and 7.19 mm for the long diameters. When the distance of the camera from the log stack is 1.5 m, the average error reaches a maximum of 12.32 mm for the short diameter and 15.11 mm for the long diameter.The standard deviation and root mean square error also shows the above trend, with the best results obtained when the camera was located at 2.5 m.

To show more visually the difference between the data measured by the model and the manually measured data, we use the manually measured values as the horizontal coordinate and the automatically measured values as the vertical coordinate to draw a scatter plot of the measured values and calculate a linear regression on them, and this is as shown in the Figure 13. The closer the two measurements are, the closer the linear regression of the scatter is to the line y = x. As shown in the Figure 13c,f, the closer the two lines (Solid and dotted line) are at 2.5 m, indicating that the automatic measurement values of the short and long diameters are closer to the manual measurement at this time.



(**a**) short diameter 1.5 m (**b**) short diameter 2.0 m (**c**) short diameter 2.5 m

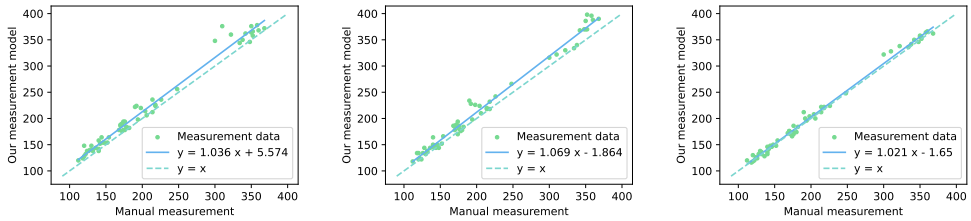


(**d**) short diameter 1.5 m (**e**) short diameter 2.0 m (**f**) short diameter 2.5 m

Figure 13. Linear regression of the length of short and long diameter between automatic measurement and manual measurement at different distance (1.5 m, 2 m, 2.5 m).

In Figure 13, we find by comparing the different images that the main reason for the larger error when the camera is located at 1.5 and 2 m is the larger measurement error in the 30 to 40 cm wood cross-section. In addition, the average error of short diameter is smaller than the average error of long diameter. These phenomena may occur because, when the camera is closer, the wood in the image is larger. Since the long diameter is a straight line taken directly perpendicular to the short diameter, when there is a small deviation in the short diameter, the deviation in the long diameter will be larger due to the larger wood cross-section, and thus the error increases. These problems are mitigated when the camera is located at 2.5 m, and thus the error is smaller.

According to Chinese Measurement Standard, We calculated the wood cross-sectional diameter using Equation (7), and similarly, we calculated the manually measured diameter and the diameter measured by our model separately and plotted the fit and calculated the error between the two sets of measurements. The Figure 14 and the Table 5 show the result, We find that after rounding the measurement results according to the standard, most metrics have decreased compared to the previous ones.



(**a**) diameter 1.5 m (**b**) diameter 2 m (**c**) diameter 2.5 m

Figure 14. Linear regression of the diameter between automatic measurement and manual measurement at different distance (1.5 m, 2 m, 2.5 m).

Table 5. Measured result after rounding.

Distance	Samples	Error	Standard Deviation	RMSE
1.5 m	59	13.18	11.89	17.75
2 m	59	12.85	11.30	17.11
2.5 m	59	5.32	4.68	7.08

5. Conclusions

In this paper, we proposed a wood measurement method that conforms to Chinese measurement standards. It uses the Mask R-CNN instance segmentation model to detect the contour of the wood log ends and employs a binocular stereo camera to calculate the size of the log ends. A rotation-based diameter search algorithm was proposed to detect the long and short diameters of the log end. The experiments show that the Mask R-CNN model has a good performance, the wood log detection rate reached 98.2%, the IOU_{mask} is 0.943 and the average precision of the contour is 0.796. Compared to manual measurement, the error of the short diameter was 5.7 mm, the error of the long diameter was 7.19 mm, and the average error of the diameter of the wood is 5.3 mm. It indicates that our proposed method has a good ability for detecting the log end and measuring the diameter according to Chinese standards.

Author Contributions: Conceptualization, C.Y. and Y.S.; methodology, Y.S.; software, C.Y.; validation, C.Y., Y.S. and J.H.; formal analysis, C.Y. and Y.F.; investigation, Y.C.; resources, Y.S.; data curation, Y.S.; writing—original draft preparation, C.Y.; writing—review and editing, Y.S. and Y.C. and Y.F.; visualization, X.Z.; supervision, J.H.; project administration, Y.S.; funding acquisition, Y.C. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by National Natural Science Foundation of China (61962055), Scientific Research Project of Southwest Forestry University (110222011) and Scientific Research Fund of Yunnan Provincial Department of Education (2023J0699).

Data Availability Statement: Part 1: The raw dataset used in this study can be found in the deepai available at https://deepai.org/publication/the-hawkwood-database (accessed on 4 May 2022). Part 2: If you need this part of experimental data, you can send an email to sunyongke@swfu.edu.cn to obtain it.

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

References

1. Food and Agriculture Organization of the United Nations. Forest Product Statistics. Available online: https://www.fao.org/forestry/statistics/80938/en/ (accessed on 20 May 2022).
2. Yu, H.; Tian, M.; Shi, Y.; Cheng, J.; Zhang, Z. The Measuring Methods of Dependence on Foreign Trade of China's Wooden Forest Products and the Estimating after Measuring. *Sci. Silvae Sin.* **2018**, *54*, 16. [CrossRef]
3. Pásztory, Z.; Heinzmann, B.; Barbu, M. Manual and Automatic Volume Measuring Methods for Industrial Timber. In Proceedings of the IOP Conference Series: Earth and Environmental Science, Da Nang, Vietnam, 25–27 February 2018; IOP Publishing: Bristol, UK, 2018; Volume 159, p. 012019. [CrossRef]
4. Borz, S.A.; Proto, A.R. Application and accuracy of smart technologies for measurements of roundwood: Evaluation of time consumption and efficiency. *Comput. Electron. Agric.* **2022**, *197*, 106990. [CrossRef]
5. Cai, Q.; Cheng, S.; Zhao, R. Interaction Relationship between Central Forestry Investment and Forestry Economic Growth in China. *Sci. Silvae Sin.* **2015**, *51*, 126–133.
6. de Miguel-Díez, F.; Reder, S.; Wallor, E.; Bahr, H.; Blasko, L.; Mund, J.P.; Cremer, T. Further application of using a personal laser scanner and simultaneous localization and mapping technology to estimate the log's volume and its comparison with traditional methods. *Int. J. Appl. Earth Obs. Geoinf.* **2022**, *109*, 102779. [CrossRef]
7. McGee, A.L.; Browning, R.A.; Yock, L.M. Log Centering Apparatus and Method Using Transmitted Light and Reference Edge Log Scanner. US Patent 4,197,888, 15 April 1980.
8. Dai, L.; Gu, J.G. The method of minimal square circularity and its application in log centering. *Wood Process. Mach.* **2002**, *13*, 16–18. [CrossRef]
9. Ramoser, H.; Cambrini, L.; Rötzer, H. Real-time 3D wood panel surface measurement using laser triangulation and low-cost hardware. In Proceedings of the Machine Vision Applications in Industrial Inspection XIV, International Society for Optics and Photonics, San Jose, CA, USA, 9 February 2006; Volume 6070, p. 60700C. [CrossRef]
10. Brüchert, F.; Baumgartner, R.; Sauter, U. Ring width detection for industrial purposes-use of CT and discrete scanning technology on fresh roundwood. In Proceedings of the Conference COST E, Delft, The Netherlands, 29–30 October 2008; Volume 53, pp. 29–30.
11. Chun-feng, W. Research of Measuring Standing Volume Based on Laser Scanning System. Master's Thesis, Northeast Forestry University, Harbin, China, 2009.
12. Zi-yan, H. A study on Wood 3d Profile Measurement System Based on Fringe Projection. Ph.D. Thesis, Beijing Forestry University, Beijing, China, 2011.

13. Longo, B.L.; Brüchert, F.; Becker, G.; Sauter, U.H. Validation of a CT knot detection algorithm on fresh Douglas-fir (Pseudotsuga menziesii (Mirb.) Franco) logs. *Ann. For. Sci.* **2019**, *76*, 1–16. [CrossRef]
14. Gergel', T.; Bucha, T.; Gracovskỳ, R.; Chamula, M.; Gejdoš, M.; Veverka, P. Computed Tomography as a Tool for Quantification and Classification of Roundwood—Case Study. *Forests* **2022**, *13*, 1042. [CrossRef]
15. Feifan, Y.; Zhongke, F.; Tao, L.; Mingming, W. Research on the Method of Log Scaling by Mobile Phone Combined With Rangefinder. *For. Resour. Manag.* **2016**, *6*, 120. [CrossRef]
16. Kruglov, A.V. The algorithm of the roundwood volume measurement via photogrammetry. In Proceedings of the 2016 International Conference on Digital Image Computing: Techniques and Applications (DICTA), Gold Coast, QLD, Australia, 30 November–2 December 2016; pp. 1–5. [CrossRef]
17. Chiryshev, Y.V.; Kruglov, A.V.; Atamanova, A.S.; Zavada, S.G. Detection and dimension of moving objects using single camera applied to the round timber measurement. In Proceedings of the 2017 Federated Conference on Computer Science and Information Systems (FedCSIS), Prague, Czech Republic, 3–6 September 2017; pp. 49–56. [CrossRef]
18. He, T.; Liu, Y.; Yu, Y.; Zhao, Q.; Hu, Z. Application of deep convolutional neural network on feature extraction and detection of wood defects. *Measurement* **2020**, *152*, 107357. [CrossRef]
19. Gao, M.; Qi, D.; Mu, H.; Chen, J. A Transfer Residual Neural Network Based on ResNet-34 for Detection of Wood Knot Defects. *Forests* **2021**, *12*, 212. [CrossRef]
20. Liu, W.; Anguelov, D.; Erhan, D.; Szegedy, C.; Reed, S.; Fu, C.Y.; Berg, A.C. Ssd: Single shot multibox detector. In *Lecture Notes in Computer Science, Proceedings of the European Conference on Computer Vision, Amsterdam, The Netherlands, 11–14 October 2016*; Springer: Cham, Switzerland, 2016; pp. 21–37. ._2. [CrossRef]
21. Tang, H.; Wang, K.; Gu, J.; Li, X.; Jian, W. Application of SSD framework model in detection of logs end. *J. Phys. Conf. Ser.* **2020**, *1486*, 072051. [CrossRef]
22. Long, J.; Shelhamer, E.; Darrell, T. Fully convolutional networks for semantic segmentation. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Boston, MA, USA, 7–12 June 2015; pp. 3431–3440.
23. Samdangdech, N.; Phiphobmongkol, S. Log-End Cut-Area Detection in Images Taken from Rear End of Eucalyptus Timber Trucks. In Proceedings of the 2018 15th International Joint Conference on Computer Science and Software Engineering (JCSSE), Nakhonpathom, Thailand, 11–13 July 2018; pp. 1–6. [CrossRef]
24. Adarsh, P.; Rathi, P.; Kumar, M. YOLO v3-Tiny: Object Detection and Recognition using one stage improved model. In Proceedings of the 2020 6th International Conference on Advanced Computing and Communication Systems (ICACCS), Coimbatore, India, 6–7 March 2020; pp. 687–694. [CrossRef]
25. Ballard, D.H. Generalizing the Hough transform to detect arbitrary shapes. *Pattern Recognit.* **1981**, *13*, 111–122. [CrossRef]
26. Haiyao, L.; Honglu, Z.; Zecan, Y.; Mengting, L. An equal length log volume inspection system using deep-learing and Hough transformation. *J. For. Eng.* **2021**, *006*, 136–142. [CrossRef]
27. Wang, J.; Chen, X.; Cao, L.; An, F.; Chen, B.; Xue, L.; Yun, T. Individual rubber tree segmentation based on ground-based LiDAR data and faster R-CNN of deep learning. *Forests* **2019**, *10*, 793. [CrossRef]
28. Luo, M.; Tian, Y.; Zhang, S.; Huang, L.; Wang, H.; Liu, Z.; Yang, L. Individual Tree Detection in Coal Mine Afforestation Area Based on Improved Faster RCNN in UAV RGB Images. *Remote Sens.* **2022**, *14*, 5545. [CrossRef]
29. Lin, Y.; Cai, R.; Lin, P.; Cheng, S. A detection approach for bundled log ends using K-median clustering and improved YOLOv4-Tiny network. *Comput. Electron. Agric.* **2022**, *194*, 106700. [CrossRef]
30. Fang, Y.; Guo, X.; Chen, K.; Zhou, Z.; Ye, Q. Accurate and Automated Detection of Surface Knots on Sawn Timbers Using YOLO-V5 Model. *BioResources* **2021**, *16*, 5390–5406. [CrossRef]
31. Zou, Z.; Shi, Z.; Guo, Y.; Ye, J. Object detection in 20 years: A survey. *arXiv* **2019**, arXiv:1905.05055.
32. He, K.; Gkioxari, G.; Dollár, P.; Girshick, R. Mask r-cnn. In Proceedings of the IEEE International Conference on Computer Vision, Venice, Italy, 22–29 October 2017; pp. 2961–2969.
33. Jian, L.; Pu, Z.; Zhu, L.; Yao, T.; Liang, X. SS R-CNN: Self-Supervised Learning Improving Mask R-CNN for Ship Detection in Remote Sensing Images. *Remote Sens.* **2022**, *14*, 4383. [CrossRef]
34. Zhang, L.; Wu, J.; Fan, Y.; Gao, H.; Shao, Y. An Efficient Building Extraction Method from High Spatial Resolution Remote Sensing Images Based on Improved Mask R-CNN. *Sensors* **2020**, *20*, 1465. [CrossRef]
35. Liu, Z.; Yeoh, J.K.; Gu, X.; Dong, Q.; Chen, Y.; Wu, W.; Wang, L.; Wang, D. Automatic pixel-level detection of vertical cracks in asphalt pavement based on GPR investigation and improved mask R-CNN. *Autom. Constr.* **2023**, *146*, 104689. [CrossRef]
36. Zhou, M.; Wang, J.; Li, B. ARG-Mask RCNN: An Infrared Insulator Fault-Detection Network Based on Improved Mask RCNN. *Sensors* **2022**, *22*, 4720. [CrossRef] [PubMed]
37. Zhang, C.; Zhou, J.; Wang, H.; Tan, T.; Cui, M.; Huang, Z.; Wang, P.; Zhang, L. Multi-Species Individual Tree Segmentation and Identification Based on Improved Mask R-CNN and UAV Imagery in Mixed Forests. *Remote Sens.* **2022**, *14*, 874. [CrossRef]
38. Hao, Z.; Lin, L.; Post, C.J.; Mikhailova, E.A.; Li, M.; Chen, Y.; Yu, K.; Liu, J. Automated tree-crown and height detection in a young forest plantation using mask region-based convolutional neural network (Mask R-CNN). *ISPRS J. Photogramm. Remote Sens.* **2021**, *178*, 112–123. [CrossRef]
39. Hu, G.; Wang, T.; Wan, M.; Bao, W.; Zeng, W. UAV remote sensing monitoring of pine forest diseases based on improved Mask R-CNN. *Int. J. Remote Sens.* **2022**, *43*, 1274–1305. [CrossRef]

40. Zhu, J.Y.; Park, T.; Isola, P.; Efros, A.A. Unpaired Image-To-Image Translation Using Cycle-Consistent Adversarial Networks. In Proceedings of the IEEE International Conference on Computer Vision (ICCV), Venice, Italy, 22–29 October 2017; pp. 2223–2232. [CrossRef]
41. Li, D.; Xie, W.; Wang, B.; Zhong, W.; Wang, H. Data Augmentation and Layered Deformable Mask R-CNN-Based Detection of Wood Defects. *IEEE Access* **2021**, *9*, 108162–108174. [CrossRef]
42. Shi, J.; Li, Z.; Zhu, T.; Wang, D.; Ni, C. Defect Detection of Industry Wood Veneer Based on NAS and Multi-Channel Mask R-CNN. *Sensors* **2020**, *20*, 4398. [CrossRef] [PubMed]
43. Rublee, E.; Rabaud, V.; Konolige, K.; Bradski, G. ORB: An efficient alternative to SIFT or SURF. In Proceedings of the 2011 International Conference on Computer Vision, Barcelona, Spain, 6–13 November 2011; pp. 2564–2571. [CrossRef]
44. Guanghua, C.; Qiang, Z.; Meiqian, C.; Jianwei, L.; Yonghuai, Y. Rapid detection algorithms for log diameter classes based on binocular vision. *J. Beijing Jiaotong Univ.* **2018**, *42*, 9. [CrossRef]
45. Zheng, B.; Sun, G.; Meng, Z.; Nan, R. Vegetable Size Measurement Based on Stereo Camera and Keypoints Detection. *Sensors* **2022**, *22*, 1617. [CrossRef]
46. Suo, F.; Huang, K.; Ling, G.; Li, Y.; Xiang, J. Fish keypoints detection for ecology monitoring based on underwater visual intelligence. In Proceedings of the 2020 16th International Conference on Control, Automation, Robotics and Vision (ICARCV), Shenzhen, China, 13–15 December 2020; pp. 542–547. [CrossRef]
47. Solak, S.; Bolat, E.D. A new hybrid stereovision-based distance-estimation approach for mobile robot platforms. *Comput. Electr. Eng.* **2018**, *67*, 672–689. [CrossRef]
48. Huang, L.; Wu, G.; Liu, J.; Yang, S.; Cao, Q.; Ding, W.; Tang, W. Obstacle distance measurement based on binocular vision for high-voltage transmission lines using a cable inspection robot. *Sci. Prog.* **2020**, *103*, 0036850420936910. [CrossRef] [PubMed]
49. Adil, E.; Mikou, M.; Mouhsen, A. A novel algorithm for distance measurement using stereo camera. *CAAI Trans. Intell. Technol.* **2022**, *7*, 177–186. [CrossRef]
50. Schraml, R.; Uhl, A. Similarity Based Cross-Section Segmentation in Rough Log End Images. In *Artificial Intelligence Applications and Innovations, Proceedings of the 10th IFIP WG 12.5 International Conference, AIAI 2014, Rhodes, Greece, 19–21 September 2014*; Iliadis, L., Maglogiannis, I., Papadopoulos, H., Eds.; Springer: Berlin/Heidelberg, Germany, 2014; pp. 614–623.
51. Jung, A.B.; Wada, K.; Crall, J.; Tanaka, S.; Graving, J.; Reinders, C.; Yadav, S.; Banerjee, J.; Vecsei, G.; Kraft, A.; et al. Imgaug. 2020. Available online: https://github.com/aleju/imgaug (accessed on 1 February 2020).
52. *Log Inspection: GB/T 144—2013[S]*; General Administration of Quality Supervision, Inspection and Quarantine of the People's Republic of China. Standards Press of China: Beijing, China, 2013.
53. He, K.; Zhang, X.; Ren, S.; Sun, J. Deep residual learning for image recognition. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Las Vegas, NV, USA, 26 June–1 July 2016; pp. 770–778.
54. Lin, T.Y.; Dollár, P.; Girshick, R.; He, K.; Hariharan, B.; Belongie, S. Feature pyramid networks for object detection. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Honolulu, HI, USA, 21–26 July 2017; pp. 2117–2125.
55. Zhang, Z. Flexible camera calibration by viewing a plane from unknown orientations. In Proceedings of the Seventh IEEE International Conference on Computer Vision, Kerkyra, Greece, 20–27 September 1999; Volume 1, pp. 666–673. [CrossRef]
56. Liu, Z.; Gu, X.; Chen, J.; Wang, D.; Chen, Y.; Wang, L. Automatic recognition of pavement cracks from combined GPR B-scan and C-scan images using multiscale feature fusion deep neural networks. *Autom. Constr.* **2023**, *146*, 104698. [CrossRef]
57. Li, C.; Li, L.; Jiang, H.; Weng, K.; Geng, Y.; Li, L.; Ke, Z.; Li, Q.; Cheng, M.; Nie, W.; et al. YOLOv6: A Single-Stage Object Detection Framework for Industrial Applications. *arXiv* **2022**, arXiv:2209.02976.
58. Wang, C.Y.; Bochkovskiy, A.; Liao, H.Y.M. YOLOv7: Trainable bag-of-freebies sets new state-of-the-art for real-time object detectors. *arXiv* **2022**, arXiv:2207.02696.

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.