IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received 28 November 2023, accepted 30 December 2023, date of publication 8 January 2024, date of current version 23 January 2024.

Digital Object Identifier 10.1109/ACCESS.2024.3351482

APPLIED RESEARCH

Automatic Sorting System for Wood Diameter Grades

JIAXIN LIU[1,2], LINGHUA KONG[1,3], JISHI ZHENG[1,4], ZHIGANG DING[1,3], LI FENG[5], AND YANLU LV[1,3]

[1]Digital Fujian Industrial Manufacture IoT Laboratory, Fuzhou 350118, China
[2]School of Management, Fujian University of Technology, Fuzhou 350118, China
[3]School of Mechanical and Automotive Engineering, Fujian University of Technology, Fuzhou 350118, China
[4]School of Transportation, Fujian University of Technology, Fuzhou 350118, China
[5]School of Mechanical and Electronic Engineering, East China University of Technology, Nanchang 330013, China

Corresponding author: Jiaxin Liu (1760795628@qq.com)

This work was supported in part by the School-Enterprise Cooperation Project of Fujian Jinsen Forestry Company Ltd., under Grant GY-H-20154, in part by the Forestry Technology Project of Fujian Province under Grant 2021FKJO6, and in part by the Nature Foundation of Fujian Science and Technology Department under Grant 2018JO1619.

ABSTRACT Wood is one of the energy sources for sustainable development, which is closely related to our life. In the wood production process, it will be sorted according to different diameter grade is an essential link, at present the wood sorting work mainly rely on manual completion, not only low efficiency, high cost, and cannot meet the automated era of production needs. In this paper, a wood diameter level automatic sorting system is designed, first of all, using computer vision technology to measure the diameter of the wood section in real time, then the diameter data is transmitted to the PLC control system, and finally the PLC control system pushes out the wood in the specified position, realizing the automated on-line sorting of wood. In addition, a wood diameter measurement method based on deep learning combined with binocular vision is proposed, which can quickly and accurately recognize the diameter of each piece of wood, and provides a solid foundation for the efficient performance of the automatic wood diameter grade sorting system. The experimental results show that the measurement method proposed in this paper has high robustness and the maximum error is within 5mm, which meets the accuracy requirements needed for sorting, and the sorting system can sort 10 pieces of wood per minute on average, which greatly improves the efficiency of wood sorting and reduces the labor workload and cost, and the research design is of great significance to the wood production and manufacturing industry.

INDEX TERMS YOLOv5, wood sorting, PLC control system, diameter measurement, machine vision.

I. INTRODUCTION

As one of the sustainable forestry resources, the value of wood is not only reflected in people's life and production, but also closely related to the protection of the ecological environment. With the dual-carbon issue, countries have begun the strategic deployment of low-carbon emission reduction [1], and the protection of forest resources, rational exploitation and utilization is an important initiative to achieve a virtuous carbon cycle. China's forestry resources are relatively scarce, and the per capita possession of wood reserves is only 12% of the global level [2], in order to reduce the phenomenon of irrational exploitation and production of wood, the Chinese government encourages the transition of forestry production to information and intelligent direction. The wood production process includes harvesting, transportation and sorting, and sorting wood by different diameter grades not only gives full play to the utilization value of wood, but also reduces unnecessary waste. All along, China's wood production and processing enterprises rely on manual completion of the wood diameter level of the sorting work, as shown in Figure 1, the sorters use calipers or tape measure to measure each piece of wood and mark the size of the diameter of the section, measurement is completed and then sorted out the

The associate editor coordinating the review of this manuscript and approving it for publication was Yongjie Li.

FIGURE 1. Manual sorting.

wood according to the different diameter level, the sorting not only consumes a lot of manpower, high production costs, but also low efficiency, long work cycle [3], unable to meet the automation of the production requirements of the era. Therefore, in order to change the way of wood sorting, it is of high value and significance to research an automated wood sorting system with high precision and speed.

Earlier, the use of light source sensors such as grating and infrared to detect wood was proposed [4], [5], [6], but due to the presence of irregular shapes of wood such as ellipse, semi-ellipse and rectangle, the linear propagation characteristics of light resulted in low accuracy of the measured wood diameter. With the development of computer vision, digital image processing, deep learning, industrial automation and other intelligent manufacturing-related technologies, the industrial automated sorting system based on visual inspection technology has gradually been applied in various industries, but its application in the field of forestry production and processing has not yet been realized. At present, there have been a number of studies related to the detection of wood section and volume by vision technology, which are mainly categorized into two directions: traditional digital image processing and deep learning. Digital image processing is the conversion of image signals into digital signals and then transforming, coding, segmenting and classifying the images according to the relevant theories of mathematical research [7]. Budiman et al. [8] fixed the Raspberry Pi equipped with a USB camera on one end of the iron rod, to measure the diameter of the wood, it is necessary to use the other end of the iron rod to place on the wood section, the image is captured by the camera, and the results are obtained by the operations of grayscale transformation, threshold segmentation, contour finding, and circle fitting, etc. The robustness of this detection method is low, and it is easy to be interfered with by the intensity of light, and it can't be applied to the outdoor. Yella and Dougherty [9] proposed a wood section segmentation method combining geometric features in HSV color space, and achieved counting and classification of whole truck wood by Circular Hough Transform (CHT) algorithm, which is less interfered by light, but it is not good at detecting soil and sewage-soaked wood sections. Yan et al. [10] used infrared light as a light source and installed a filter in front of the camera lens, so that only the wood section area exists on the picture captured by the camera, and the diameter can be calculated by ellipse fitting without the need of image segmentation, which is a fast and accurate method, but is only applicable to indoor environments. Chen et al. [11] proposed a fast detection algorithm for wood diameter grade based on binocular vision, which segmented the wood section area from the picture by dynamic threshold segmentation algorithm, and then used the least squares method to fit the wood contour to get the diameter-related parameters, which can measure the wood diameter grade at any distance, but the segmentation of the wood contour area is not accurate. In the literatures [12], [13], and [14], Kruglov proposed three methods for wood testing: Firstly, the 3D modeling technology was used to obtain the 3D information of wood piles and combined with the mean clustering algorithm, Delaunay triangulation algorithm and watershed algorithm to carry out the wood volume measurement experiments, respectively, and the final measurement results were compared with the manual measurement results, with an average error of 9.2%. Secondly, an improved radially symmetric target detection algorithm is obtained by the combination of the Stoer-Wagner algorithm and the watershed algorithm, which can be adjusted to the contour of the detected wood section and accurately segment the target region. Thirdly, a circular target automatic detection method is proposed, but the method is not effective due to the irregularity of the shape of the wood section. Deep learning is a multilevel network algorithm that mimics the way information is transmitted between neurons in the human brain, and has efficient image recognition, classification, and segmentation capabilities, and its application to wood detection and segmentation yields better results and higher robustness than traditional digital image processing techniques. Samdangdech and Phiphobmongkol [15] proposed SSD model combined with full convolutional network semantic segmentation model to segment the section contour of the whole car wood, firstly, the SSD model is used to detect the coordinates of the wood region in the image, and then the wood region picture is obtained after cropping, and then the wood region picture is inputted to the FCN model to segment the section contour of each wood, and finally the number of wood roots is obtained by counting the contours, and the accuracy of this method is verified to be 94.45%, but it is not effective for small diameter wood and obscured wood. Finally, the number of roots is obtained by counting the contours. The accuracy of this method is verified to be 94.45%, but it is ineffective in segmenting small-diameter wood and obscured wood. Tang et al. [16] proposed a wood detection method based on a single-shot multibox detector, which is able to extract target features at different scales by automatically learning the relevant features from the labeled information of the wood section in the image, and

the detection accuracy reaches 94.87%, which solves the problem of wood detection in complex backgrounds. Yu et al. [17] proposed the fusion of BiFPN and YOLOv5s for wood section detection, adding a small target feature extraction network layer, which improves the detection accuracy and checking rate of the small targets in wood section, and can quickly detect the wood section area. Lin et al. [18] used YOLOv3-tiny to detect the wood section area to get the corresponding coordinate information, and then used Hough transform for circle detection to get the wood diameter. Yang et al. [19] proposed a wood section detection method based on Mask R-CNN, which is robust and capable of accurately segmenting the wood section region in different environments, with a true detection rate of 97.989% on its self-defined dataset.

From the results of the above research, it can be seen that the use of deep learning to detect wood is more reliable than the traditional image processing methods, due to the strong anti-interference ability, applicable to a variety of complex environments, deep learning technology as the basis for the design of a wood diameter level of the automatic sorting system can meet the needs of the work of outdoor complex environments. However, the existing wood detection methods and deep learning networks are unable to simultaneously meet the application requirements of wood diameter-level sorting systems in terms of accuracy and speed. For example, the wood contour detection method based on Hough's circle transform cannot accurately detect the non-circular wood; the target detection model with simple network structure can quickly identify the wood position coordinates, but cannot further segment the wood contour; the instance segmentation model can accurately segment the wood contour region, but its network structure is complex, parameter count is large, and reasoning is slow, which is not applicable to embedded wood sorting systems with low arithmetic power. In addition, the distance from the wood section to the camera lens during the sorting process is not fixed, and in order to ensure the accuracy of wood diameter measurement, it is necessary to obtain the depth information of each wood in real time. According to the above problems, this paper proposes a wood diameter measurement method based on deep learning combined with binocular vision, which improves the speed and accuracy of wood diameter measurement, and designs a wood diameter level automatic sorting system, which utilizes the vision measurement system to obtain the wood diameter data and transmits it to the PLC control system, realizing the function of automatic sorting of wood diameter level, which effectively solves the problems of low efficiency of manual sorting, high cost, and inability to realize the digitalization of the wood industry.

II. SORTING SYSTEM OVERALL

The structure of the wood sorting system is shown in Figure 2, which is mainly composed of a visual inspection system and a PLC control system. First, the ZED binocular camera is used to capture the picture of the wood section, then the picture is inputted to the Jetson nx platform for inference and calculation to get the diameter data, and finally the diameter data is transmitted to the PLC. PLC is the core module for the motion control of the sorting system, which can not only control the motor to transfer the wood at different speeds, but also control the action of the cylinder through the solenoid valve to push out the wood at the specified position, so as to realize the automatic sorting of wood. In the sorting process, the photoelectric sensor is equivalent to the locator of the wood, no matter the wood arrives at the visual detection position or the sorting position, it is by triggering the photoelectric sensor that the PLC sends out the visual detection signal and the sorting signal.

A. SORTING SYSTEM HARDWARE DESIGN

1) WOOD TRANSFER MODULE

In the wood sorting work, first of all, the robotic arm grabs the wood and puts it in the stacking area of the sorting system, and then the transmission module transmits the piles of wood to the visual detection position in sequence, in order to ensure the stability and efficiency of the wood transmission, this paper designs the wood transmission structure shown in Fig. 3, which mainly consists of the stacking area, the wood flattening device, the single separator device, and the feeding device. Pushed by the first conveyor belt, the wood in the stacking area moves forward, and the wood laying device can rotate at a constant speed to lay the wood in order on the storage rack, and when there is wood on the single divider, the first conveyor belt stops, and when the feeding device takes the wood away, the first conveyor belt continues to run until the next wood reaches the single divider. With the cooperation of the mechanical devices of the wood transfer module, the wood can be stably transported to the visual inspection position in sequence, realizing the basic work of automated wood sorting.

2) VISUAL INSPECTION MODULE

Due to the complexity of the application environment of the automatic sorting system of wood diameter grade, in order to meet the performance requirements of wood inspection, the visual inspection system shown in Figure 4 is designed. It mainly consists of binocular camera and embedded image processing platform. The binocular camera adopts Stereo-Labs' ZED 2i binocular camera, which can realize the work of aberration correction and calibration through the SDK (Software Development Kit) and API (Application Programming Interface) provided by the official government, which can eliminate the tedious steps brought by the traditional calibration method, and can quickly carry out the calibration and aberration correction during the actual development process, and it has the advantages of high efficiency and high precision. In addition, the camera software carries a stereo matching algorithm optimized in combination with a deep learning algorithm, which can use the API to directly obtain the depth information of each pixel point in the

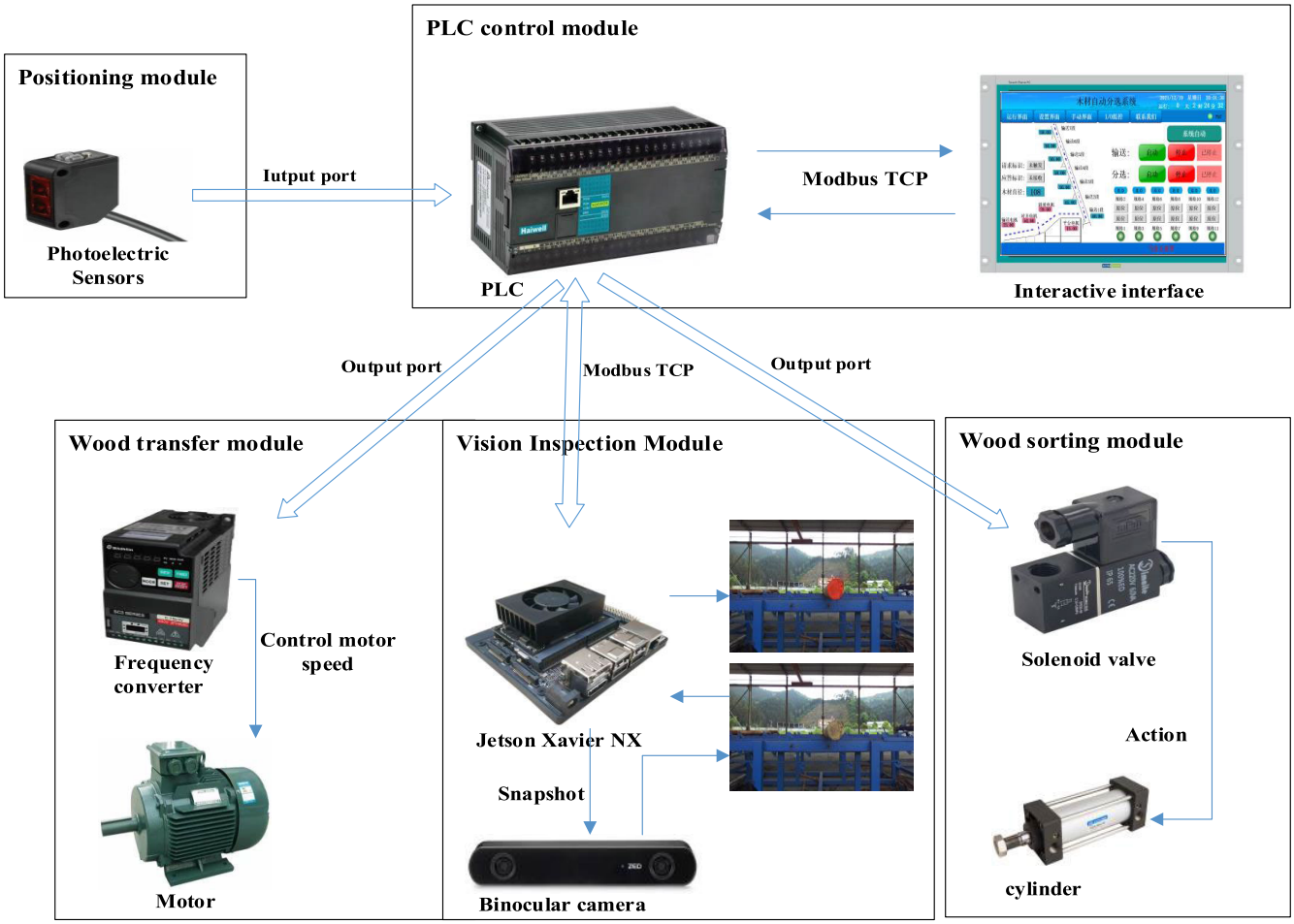

FIGURE 2. Sorting system architecture diagram.

image, and the error is controlled at the millimeter level, which ensures the accuracy required for wood diameter measurement. The embedded image processing platform uses NVIDIA's Jetson nx development board, a high-performance, low-power AI edge computing platform with NVIDIA's Volta architecture GPUs and six-core Carmel ARM CPUs, which integrates GPUs, CPUs, memory, storage, and a variety of interfaces, greatly improving development efficiency. The Jetson nx development board also has a well-established software ecosystem that supports the SDK deployment conditions for the ZED 2i binocular camera and depth acquisition requirements in various modes. This paper utilizes the ZED 2i binocular camera and the Jetson nx development board to build a wood diameter inspection system, which not only realizes the requirements of high accuracy, high robustness and high efficiency, but also has a simple structure and high flexibility.

3) WOOD SORTING MODULE

The visually inspected wood will be transferred to the wood sorting module, where the PLC controller needs to control the cylinders to push out the different diameter grades of

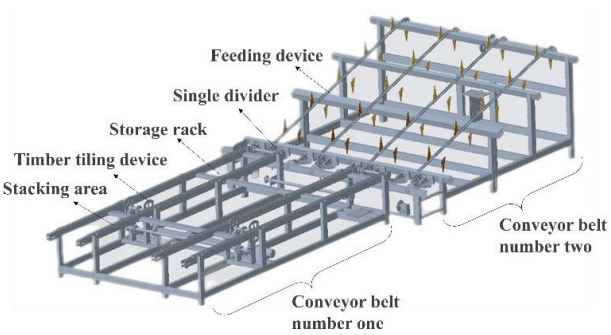

FIGURE 3. Structural diagram of the wood transfer module.

wood in a defined position. In order to realize the sorting action, this paper designs the wood sorting module shown in Fig. 5, which is mainly composed of cylinder, push arm and conveyor belt, and the whole sorting line consists of six same sorting modules. When the wood arrives at the specified sorting module, the photoelectric sensor detects the presence of wood on the sorting module, and the PLC controller controls the cylinder through the solenoid valve to lift the push arm to push out the wood and complete the wood sorting.

FIGURE 4. Visual inspection system structure diagram.

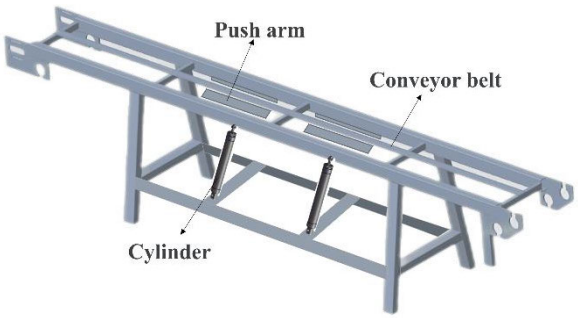

FIGURE 5. Wood sorting module structure.

B. SORTING SYSTEM SOFTWARE PROCESS DESIGN

The software flow of the wood diameter level automatic sorting system is shown in Fig. 6, when the wood is transmitted to the visual detection position, the photoelectric sensor is triggered, and the visual detection photoelectric signal bit is changed from 0 to 1, the ZED camera starts to capture a frame of the wood section picture and saves it to a local folder, and after that, the deep learning model splits out the contour of the wood section in the picture, and then carries out ellipse fitting to get the parameters related to the diameter of the wood, and finally, combines with the depth information obtained by the ZED camera, the calculation is done to get the wood diameter data. Wood diameter data through Modbus TCP communication to the PLC controller data storage queue, the data storage queue implementation of the principle of first-in-first-out, when the wood arrives at the corresponding sorting position, triggered by photoelectric sensors, the PLC controller detects the sorting photoelectric signals, control solenoid valve rotation, so that the cylinder action will be pushed out of the wood, to complete the wood sorting.

III. WOOD DIAMETER MEASUREMENT BASED ON DEEP LEARNING COMBINED WITH BINOCULAR VISION

Fast and accurate measurement of the diameter of wood is the key technology of the automatic sorting system of wood diameter level, according to the existing wood detection methods exist in the low robustness, high accuracy and deep learning model network complex and parametric number of problems, an improved YOLOv5-seg model is proposed for the segmentation of the contour region of the wood section. The segmented wood section profile is ellipse-fitted and then combined with the depth information obtained from the binocular camera to calculate the diameter, which is not only highly accurate and robust, but also can be deployed in the low arithmetic power of the wood diameter level of the automatic sorting system.

A. YOLOV5N-SEG ALGORITHM IMPROVEMENT AND OPTIMIZATION

1) E-SN MODULE

The C3 module in the YOLOv5n-Seg network exists in the feature extraction and feature fusion network, which mainly

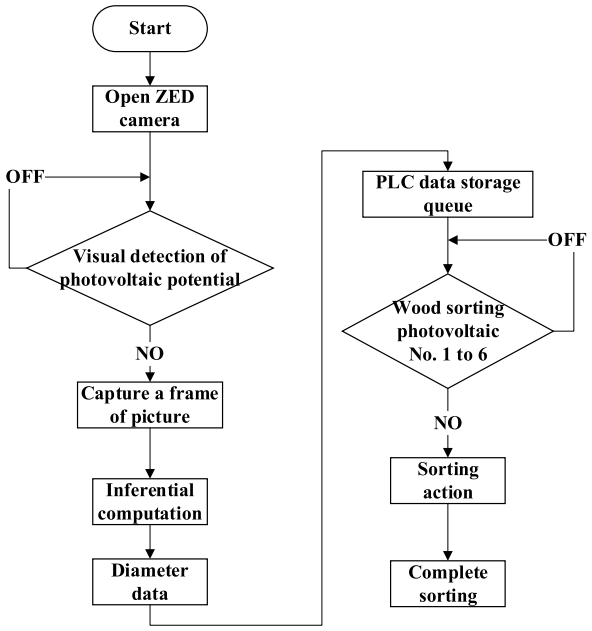

FIGURE 6. Sorting system software flowchart.

consists of CBS and different number of Bottlenecks connected in a cross-layer manner.Although it improves the feature representation capability, it also increases the computational complexity and memory occupation. The existence of jump connections in the Bottleneck of the C3 module in the feature extraction network enables internal transfer of information across layers so that the higher layers can use the lower layer information, thus mitigating the problem of gradient vanishing and gradient explosion and enhancing the flow of information in the network, at the same time this introduces additional parameters due to the need for element level summation operations, making the problem of parameter redundancy more pronounced, and increasing the parameter count of the model. In general, the C3 module has better complex feature extraction capability with a large number of parameters, while the wood section segmentation task has a single objective and needs to be applied in low-calculation sorting equipment, so designing a lightweight module to replace the C3 module has become the preferred method of improvement in this study.

ShuffleNetV2 [20] is a lightweight and efficient neural network model whose base module consists of three main components: channel separation (Channal Split), depth separable convolution (DWConv) and channel shuffling (Channal Shuffle). E-SN module consists of the basic module of ShuffleNetV2 network fused with SE channel attention mechanism, the network structure is shown in Figure 7, K1, K2 are the number of channels of input and output feature maps, respectively, and Channal Split operation is carried out first to separate the feature maps in the channel dimension to get the two groups of feature maps with the number of channels of K1/2, so as to carry out the processing in different ways, which not only helps to enhance the feature diversity and expressive ability, but also effectively reduces the amount of computation of the model and the number of parameters. One group first performs deep convolution and point-by-point convolution operation, and then obtains the new feature map after SE channel attention weight optimization, deep convolution is a lightweight convolution operation, the main role is to extract feature information in the spatial dimension, a channel corresponds to a convolution kernel, which ensures the independence of the channel, and point-by-point convolution carries out feature fusion and information extraction in the channel dimension, and connects each channel in series, which ensures the richness of the output feature information, and the two-step operation makes the depth-separable convolution to get the same feature extraction effect as the ordinary convolution at a small computational cost, which is greatly reduced compared with the number of the ordinary convolution parameter counts. The new feature map is superimposed and fused with another group to form a feature map with the number of channels as K2, and then the channel mixing is performed to rearrange the feature channels and combine them according to a certain law in order to improve the interactivity of the feature information, so as to enhance the ability of the network to extract and transmit feature information.

The structure of SE module is shown in Fig. 8, which mainly contains two operations Squeeze and Excitation [21], where Squeeze is used to solve the problem of exploiting channel dependency by globally pooling the feature maps, so that the spatial information of each channel is compressed into the corresponding value, which is equivalent to a value with a global sensory field of the corresponding channel, whose expression is shown in Eq. (1). Excitation is used to fuse the information between the channels, keep the channels with large feature information and eliminate the channels with small amount of information, the related operation is completed by the two fully-connected layers, and the feature information is finally superimposed to get the required feature information through the weight allocation, and its expression is shown in Eq. (2).

$$Z = F_{\mathrm{sq}}(U) = \frac{1}{H \times W} \sum_{\mathrm{i}=1}^{H} \sum_{\mathrm{j}=1}^{W} U(i,j) \tag{1}$$

$$S = F_{ex}(Z, \omega) = \sigma(g(Z, \omega)) = \sigma(\omega_2 \sigma(\omega_1 Z)) \tag{2}$$

In formula (1), $U(i, j)$ represents a single-channel feature map of size $i \times j$, and H and W are the height and width of the feature map respectively. In formula (2), ω_1 and ω_2 represent different weight values.

2) IMPROVEMENTS IN CONVOLUTIONAL LAYERS

In order to solve the problem that ordinary convolution generates a large number of redundant feature maps during feature extraction, which leads to consuming more computational memory and decreasing the inference speed, GhostConv, a base module in the Ghost [17] network proposed by Huawei's team, is introduced instead of ordinary convolution.

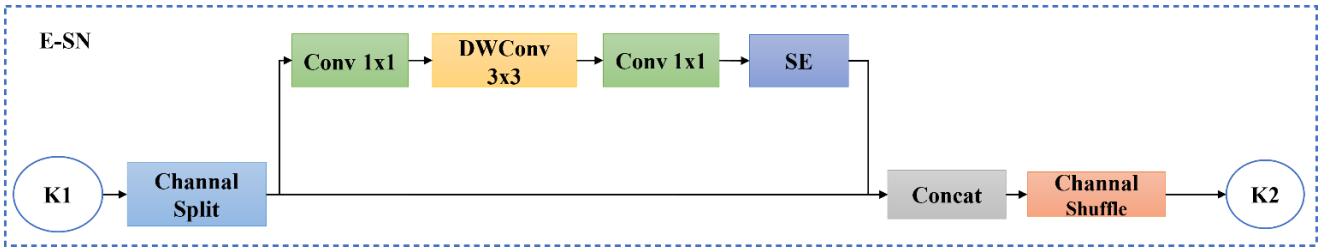

FIGURE 7. E-SN module structure diagram.

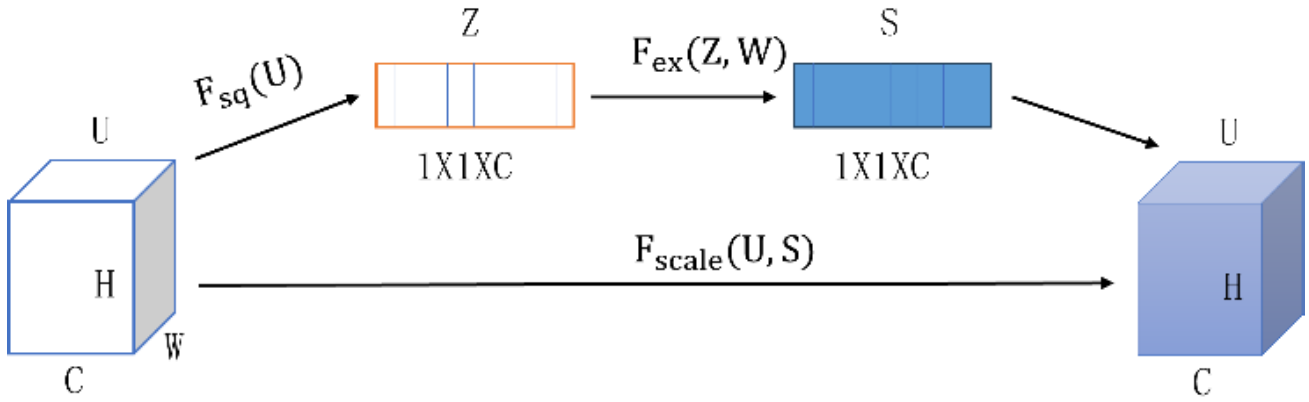

FIGURE 8. SE module structure diagram.

GhostConv can extract rich feature information in half of the feature maps, the key lies in the Cheap operation [18], which generates similar feature maps by linear transformation of the feature maps, which not only ensures the richness of the feature information but also greatly reduces the amount of computation. GhostConv structure is shown in Figure 9, C1, C2 for the feature map input and output channel number, first through the convolution kernel size $K \times K$, the number of C2/2 ordinary convolution layer to get the number of channels for the C2/2 feature map M, and then use Depthwise convolution to generate M's "phantom" feature map N, convolution kernel size for the Depthwise convolution of 5×5 has a larger sense of the field, so that the extracted features are more rich in information, and finally will be the feature map M and N for the output of Concat operation for the number of channels for the C2 feature map.

3) IMPROVEMENT OF THE LOSS FUNCTION

The Bounding Box loss function in the YOLOv5n-Seg model adopts the CIOU loss [19], which is composed of three parts: Classification loss, Localization loss, and Confidence loss, in which both Classification loss and Confidence loss are calculated using the binary cross-entropy function for each label loss, reducing the computational complexity. Confidence loss are calculated using the binary cross-entropy function for each label loss, which reduces the computational complexity, the formula of CIOU loss is shown below:

$$CIOU = IOU - \frac{\rho^2(b, b^{gt})}{c^2} - \alpha\upsilon \tag{3}$$

In Eq. (3) b and b^{gt} are the Euclidean distances between the centers of the prediction frame and the real frame, respectively, c is the diagonal distance between the prediction frame and the smallest outer frame of the real frame, and the formulas for α and υ are shown below:

$$\alpha = \frac{\upsilon}{(1 - IOU) + \upsilon} \tag{4}$$

$$\upsilon = \frac{4}{\pi^2}(\arctan\frac{w^{gt}}{h^{gt}} - \arctan\frac{w}{h})^2 \tag{5}$$

An excellent Bounding Box loss function must take into account the overlap area, the distance from the center point and the aspect ratio of the three parameters [20], CIOU loss will take these three parameters into account, not only can make the prediction box quickly close to the target box can also make the prediction box of the aspect ratio of the box quickly close to the target box, but when the prediction box does not match the real box will lead to slower convergence speed of the model and the accuracy decreases. In this regard, SIOU loss [21] is adopted as the model improved Bounding Box loss function, which introduces the vector angle parameter between the real box and the predicted box and redefines the calculation method of the loss, which consists of four parts in total: the vector angle loss between the real box and the predicted box, the centroid distance loss, overlapping area loss, and aspect ratio loss, of which Angle cost will make the prediction frame first move quickly to the nearest coordinate axis, and then along the coordinate axes X or Y axis constantly close to the real frame, which reduces the degrees of freedom and greatly improves the convergence

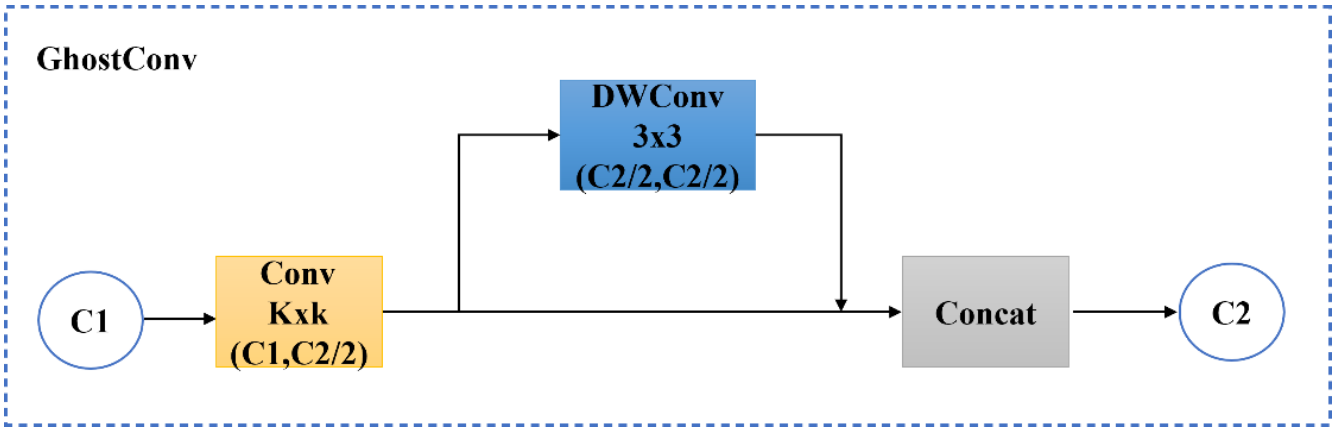

FIGURE 9. Ghostconv network structure diagram.

speed and efficiency of the model, the Angle cost is defined as shown in the following equation:

$$\Lambda = 1 - 2 * \sin(\arcsin(\frac{c_h}{\sigma}) - \frac{\pi}{4})^2 \tag{6}$$

Eq. (6) where c_h is the height difference between the center point of the real frame and the predicted frame, σ is the distance between the real frame and the center point of the predicted frame, $\arcsin(\frac{c_h}{\sigma})$ is the vector angle α, and the expressions for c_h and σ are shown in the following equation:

$$\sigma = \sqrt{(b_{c_x}^{gt} - b_{c_x})^2 + (b_{c_y}^{gt} - b_{c_y})^2} \tag{7}$$

$$c_h = \max(b_{c_y}^{gt}, b_{c_y}) - \min(b_{c_y}^{gt}, b_{c_y}) \tag{8}$$

In Eqs. (7) and (8) $(b_{c_x}^{gt}, b_{c_y}^{gt})$ are the coordinates of the center point of the real frame and (b_{c_x}, b_{c_y}) are the coordinates of the center point of the predicted frame. The vector angular loss is calculated as in Fig.10, when α is $\frac{\pi}{2}$ or 0, the angular loss is 0. During training when $\alpha < \frac{\pi}{4}$, the convergence process is minimized α and vice versa minimized β.

Based on the above definition of angular loss, SIOU loss is redefined for the three types of losses: centroid distance, overlap area, and aspect ratio, and the expressions are shown below:

Center distance loss definition expression:

$$\Delta = \sum_{t=x,y} (1 - e^{-\gamma\rho t}) \tag{9}$$

$$\rho_x = (\frac{b_{c_x}, gt - b_{c_x}}{c_w})^2 \tag{10}$$

$$\rho_y = (\frac{b_{c_y}, gt - b_{cy}}{c_h})^2 \tag{11}$$

$$\gamma = 2 - \Lambda \tag{12}$$

When α tends to 0, the percentage of center distance loss in SIOU loss decreases dramatically, and when α tends to $\frac{\pi}{4}$, the percentage of center distance loss in SIOU loss becomes larger, but the larger the angle the more complicated the loss calculation becomes, for which the parameter γ, which is prioritized over the value of distance as the angle increases, is added.

Aspect ratio loss definition expression:

$$\Omega = \sum_{t=w,h} (1 - e^{-wt})^{\theta} \tag{13}$$

$$w_w = \frac{|w - w^{gt}|}{\max(w, w^{gt})} \tag{14}$$

$$w_h = \frac{|h - h^{gt}|}{\max(h, h^{gt})} \tag{15}$$

(w, h) and (w^{gt}, h^{gt}) in Eqs.(14) and (15) are the width and height of the prediction and real frames, respectively, and the θ parameter acts on the attention of the aspect ratio loss, which is set to 4 in this paper in order to avoid too much attention to reduce the movement of the prediction frame.

In summary, the final defined expression for SIOU loss is shown in equation (16):

$$Loss_{SIOU} = 1 - IOU + \frac{\Omega + \Delta}{2} \tag{16}$$

B. IMPROVED YOLOv5n-SEG NETWORK MODELING

The structure of the improved YOLOv5n-seg network is shown in Fig. 6, except for the first convolutional layer in the feature extraction network, the other ordinary convolutional layers of the feature extraction network and feature fusion network all use GhostConv convolution, which is able to reduce the number of network parameters under the circumstance of guaranteeing the extraction of rich features. In addition, all the C3 modules in the model are changed to E-SN modules, which further reduces the network complexity and the number of parameters and improves the model accuracy, and the SIOU loss is used as the regression loss of the model, which increases the vector angle penalty term compared to the CIOU loss, which speeds up the model convergence and is conducive to the improvement of segmentation accuracy in the homemade dataset.

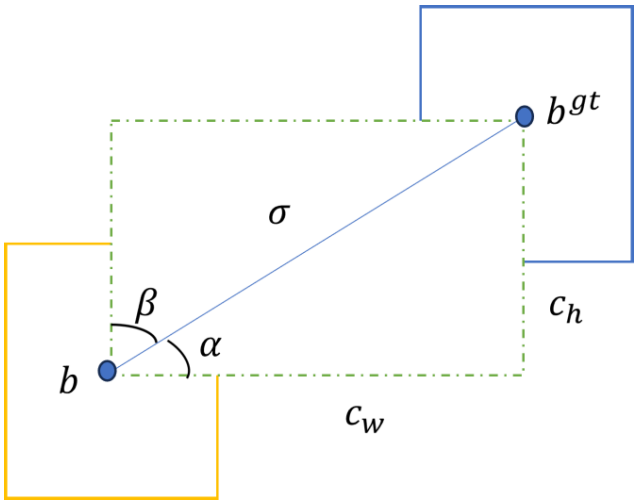

FIGURE 10. Vector angle loss calculation diagram.

C. DEPTH INFORMATION ACQUISITION AND WOOD DIAMETER CALCULATION

According to the actual engineering needs, this paper ranging experiments selected 1080P distortion-free USB binocular camera module, focal length of 6mm. The principle of binocular vision ranging is shown in Fig.12, assuming that W is a point on the wood section, points O_L and O_R are the optical centers of the left and right camera lenses, respectively, points W_L and W_R are the imaging points of W on the left and right camera sensors, f is the focal length of the left and right cameras, B is the center distance between the left and right cameras, and D is the distance from W to the camera.

Let the distance between points W_L and W_R be L, then:

$$L = B - (X_L - X_R) \tag{17}$$

It can be obtained from similar triangles:

$$\frac{L}{B} = \frac{D - f}{D} \tag{18}$$

Can be obtained from the above two equations:

$$D = \frac{B * f}{X_L - X_R} \tag{19}$$

where B and f can be obtained by camera calibration, $X_L - X_R$ is the parallax of W points imaged in the left and right cameras, and the parallax value can be obtained by calling the BM algorithm or the SGBM algorithm in the Opencv library [27] to get the distance from the point to the camera on the wood section.

Due to the actual manual diameter level measurement, the short diameter of the wood small head section is used as the measurement position, and the wood is sorted according to this measurement result. When the instance segmentation model outputs the wood section Mask, this paper adopts the least squares ellipse fitting algorithm to ellipse fit the contour of the Mask, and the corresponding ellipse center coordinates (x, y), the long axis r_a and the short axis r_b can be obtained, and the rotation angle of the ellipse center θ.

Assuming that the coordinates of the two endpoints of the short axis r_b of the ellipse in the world coordinate system are (x_1, y_1, z_1), (x_2, y_2, z_2), and that the points on the wood section are approximated to be on the same level, $z_1 = z_2$, the distance between the two endpoints can be found out according to the Euclidean distance formula, which is the true diameter of the wood section L:

$$L = \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2} \tag{20}$$

from the geometric relationship in binocular imaging:

$$\begin{cases} \frac{x_1}{D} = \frac{u_1}{f} \\ \frac{y_1}{D} = \frac{v_1}{f} \end{cases} \quad \begin{cases} \frac{x_2}{D} = \frac{u_2}{f} \\ \frac{y_2}{D} = \frac{v_2}{f} \end{cases} \tag{21}$$

(u_1, v_1) and (u_2, v_2) in Eq.(21) correspond to the pixel coordinates of short axis r_b in the image according to the elliptic law, respectively:

$$\begin{cases} (u_1, v_1) = (x + 0.5 r_b \sin\theta, y + 0.5 r_b \cos\theta) \\ (u_2, v_2) = (x - 0.5 r_b \sin\theta, y - 0.5 r_b \cos\theta) \end{cases} \tag{22}$$

Combining Eqs. (20), (21), and (22) yields a formula for calculating the section diameter of wood:

$$L = D\sqrt{\frac{(r_b \sin\theta)^2}{f} + \frac{(r_b \cos\theta)^2}{f}} \tag{23}$$

IV. EXPERIMENTAL DESIGN AND ANALYSIS

A. CONSTRUCTION OF THE DATASET

The dataset used in this paper comes from a wood diameter grade sorting yard of Fujian Jinlin Forestry Co. Firstly, 1280 × 720 resolution wood section pictures were collected with binocular camera under different natural environments, such as sunny, cloudy, and rainy days, and a total of 5,175 pictures were screened, and then a total of 5,775 pictures were collected. Next, labelme is used to generate the corresponding json file, and then a python script is used to convert the labeled json file into the txt file required by the YOLOv5n-Seg model. Finally, 5175 datasets are divided into training set (4140) and validation set (1035) by 4:1, and the sample dataset is shown in Fig. 13.

B. EXPERIMENTAL ENVIRONMENT AND PARAMETER SETTING

The operating system for the experiments in this paper is Windows 10, the processor is Intel Xeon W-2145, the graphics card is RTX2080ti, and the deep learning framework is pytorch1.12. The SGD optimizer is used for training and the relevant parameters are set as shown in Table 1.

C. MODEL EVALUATION AND PERFORMANCE ANALYSIS

1) EVALUATION INDEX

The purpose of the model improvement in this study is to ensure the detection accuracy while reducing the number of

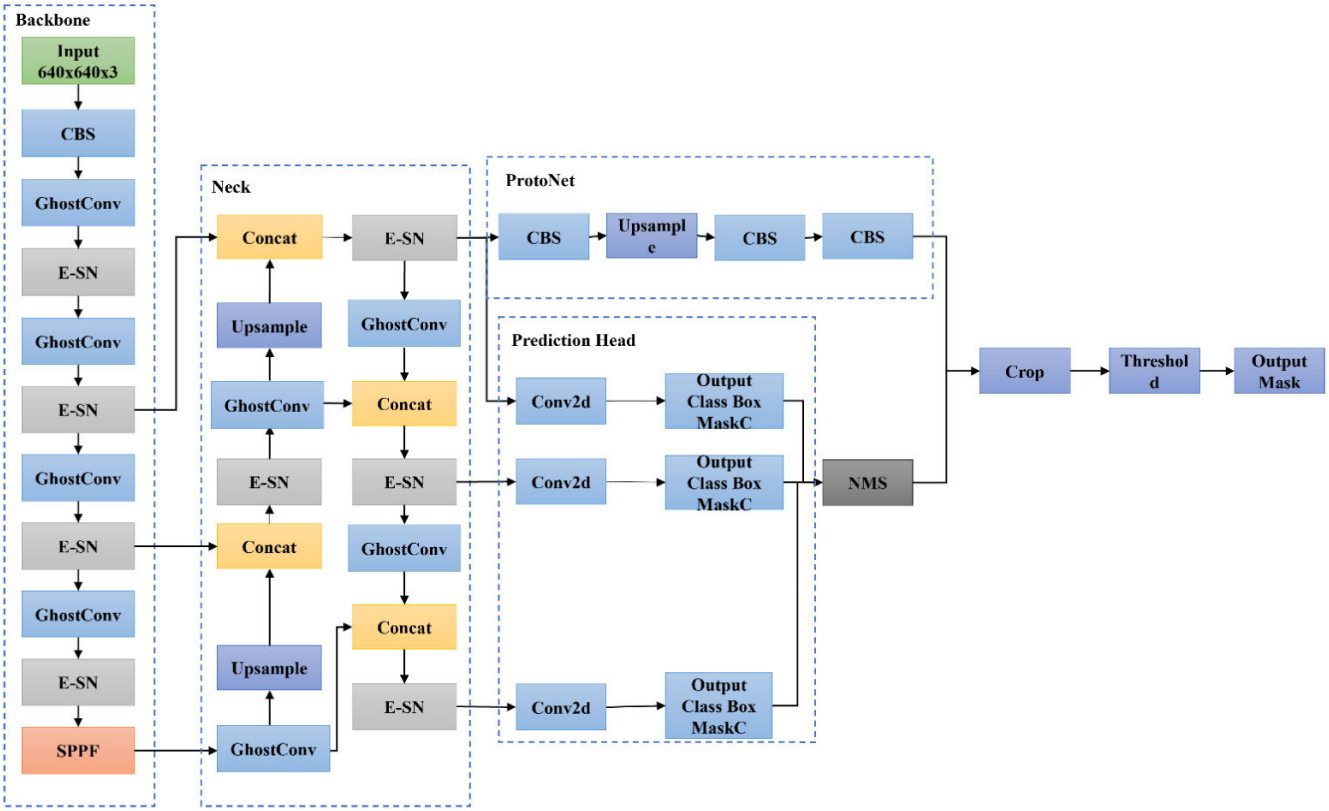

FIGURE 11. Improved YOLOv5n-seg network structure.

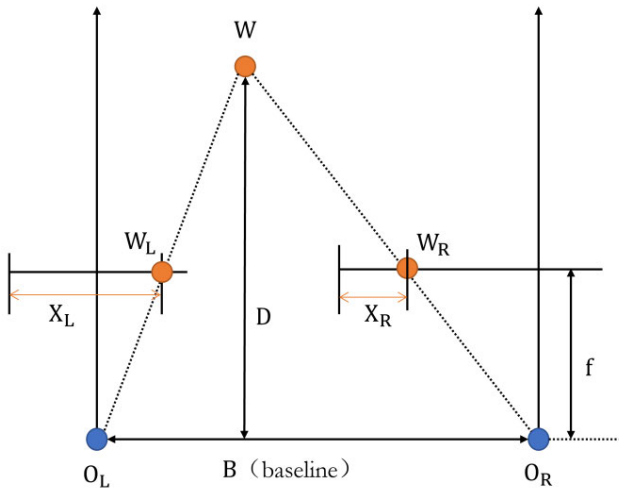

FIGURE 12. The principle of binocular ranging.

parameters and complexity of the model. In this regard, the performance of the improved model is verified using the mask average accuracy, the number of model parameters, and the amount of computation. Different from the commonly used evaluation metrics for target detection models, the threshold of Mask AP acts on the intersection and concurrency ratio of labeled Mask and predicted Mask,which can comprehensively evaluate the segmentation accuracy of instance segmentation models. Under the same threshold, the larger Mask AP value indicates the higher segmentation accuracy of the model, and the value of AP is selected when the threshold is 0.5 to 0.95, and one value is taken every 0.05, and a total of 10 values are averaged to obtain $AP_{0.5:0.95}$.Mask AP is calculated from the accuracy and recall of the mask segmentation, P as the vertical coordinate and R as the horizontal coordinate to draw the $P-R$ curve, the area enclosed by the curve and the horizontal and vertical coordinates is the AP value, the expression is shown below:

$$P = \frac{T_P}{T_P + F_P} \tag{24}$$

$$R = \frac{T_P}{T_P + F_N} \tag{25}$$

$$AP = \int_0^1 P(R) d_R \tag{26}$$

In Eqs. (24)and (25), T_P indicates the number of positive samples predicted to be positive, F_P indicates the number of negative samples predicted to be positive, and F_N indicates the number of negative samples predicted to be negative.

2) ABLATION EXPERIMENT

In order to verify the validity of each improvement point in the improved model of this paper, the AP value, the number of model parameters and the amount of computation, etc. are used as the evaluation indexes of the ablation experiments,

FIGURE 13. Dataset sample diagram.

and the improvement points are categorized into four groups of ablation experiments, namely, A, B, C, D, and E, based on YOLOv5n-Seg model, with "√" representing the adoption of the method in the experiments, and the results are shown in Table 2.

As can be seen from Table 2, Group A replaces all ordinary convolutional layers with Ghostconv convolutional layers in the baseline model, and the reduction in the amount of ordinary convolution leads to a reduction in the number of parameters but its accuracy decreases by 4.1 percentage points; Group B replaces all C3 modules with E-SN modules in the baseline model, and the number of parameters is significantly reduced and the accuracy decreases by only 0.7 points, which shows that the E-SN modules are able to extract rich feature information in a small number of feature maps; Group C replaces the CIOU loss function with the SIOU loss function in the baseline model, which improves its accuracy by 0.6 percentage points while the number of model parameters remains unchanged, so it can be seen that the addition of angular penalties to the SIOU also improves the accuracy a little. According to the results of the individual experiments of the improvement points of groups A, B and C, each improvement point plays a positive role in the accuracy improvement or the reduction of the amount of model parameters, and in order to further explore the extent of the impact of the superposition of the use of these improvement points on the model, they are divided into different teams to conduct the experiments. Group D replaces the ordinary convolution and C3 module with Ghostconv convolution and E-SN module in the benchmark model, the number of model parameters decreases significantly but the accuracy also decreases by 1.1 percentage points, in order to minimize the loss of accuracy, Group E replaces the loss function with SIOU based on Group D, and the resultant accuracy also rises by 0.3 percentage points compared with that of the benchmark model, and the amount of model parameters decreases by 61.1%. All in all, the lightweight and improved model meets the needs of deployment on embedded devices by reducing the number of parameters drastically while improving the accuracy slightly.

3) COMPARATIVE EXPERIMENTS OF DIFFERENT SEGMENTATION MODELS

In order to further verify the excellence of the lightweight improved algorithm, it is compared with the classic instance segmentation algorithms such as Mask R-CNN [28], Cascade Mask R-CNN [29], YOLACT [30] and SOLO [31] on the homemade dataset, and the results are shown in Table 3. This paper's algorithm is better than other instance segmentation models both in terms of the number of parameters, the amount of computation, and the accuracy, in which YOLACT as a single-stage instance segmentation model is better than the other two-stage models both in terms of the number of parameters and the speed, but compared with the improved algorithm of this study, the number of parameters is 12 times as much as that of this paper's algorithm, the amount of computation is 4 times as much as that of this paper's algorithm, and the accuracy is still 5.1 percentage points less. It can be seen that the improved algorithm in this paper has superior performance when applied to embedded wood sorting equipment, and the Mask segmentation effect of the improved model and each classical segmentation model is shown in Fig. 14. a, b, c, d, e, and f correspond to the effect diagrams of wood section separation for the improved models in this paper, Mask R-CNN, Cascade Mask R-CNN, YOLACT, YOLOv5, and SOLO, respectively.

TABLE 1. Model training related parameter settings.

parameter name	parameter value
image input size	640X640
batch size	16
learning rate	0.01
attenuation coefficient	0.0005
momentum	0.937
iterations	300

D. DIAMETER MEASUREMENT EXPERIMENTS AND ANALYSIS OF RESULTS

1) MEASURING EXPERIMENTAL PROGRAMS AND EVALUATING INDICATORS

According to the rules of sorting wood diameter level, the wood diameter of 8cm or less for a diameter level, the diameter of 8cm or more wood to 2cm for a diameter level amount of sequential increments into different diameter levels. In this experiment, a total of 12 diameter classes were set up from small to large, numbered A to L in order, and 10 pieces of wood were selected in each class, a total of 120 pieces of wood, the minimum diameter of wood is 3.5cm, the maximum is 29.8cm, and each piece of wood was manually measured in millimeters and labeled as a comparison of the measurement methods of this study. In order to verify the practicability of this method, a test was carried out on the wood diameter grade sorting equipment in a yard of Fujian Jinlin Forestry Co. 120 pieces of wood were sequentially transferred to the visual measurement point without interruption until all the measurements were finished, and the test scene is shown in Fig. 15.

The results of the experiment were evaluated using the following evaluation metrics:

1. Absolute error of wood diameter measurement: The absolute error was obtained by subtracting the measured

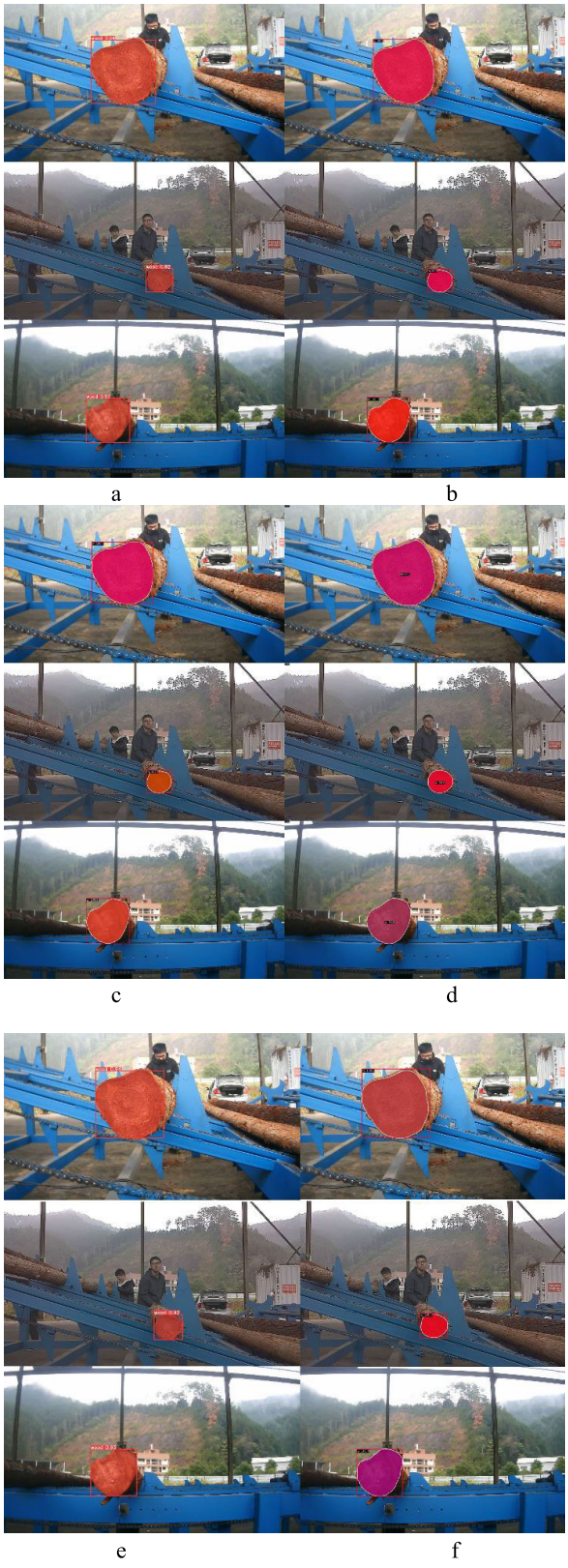

FIGURE 14. Comparison of mask results of each segmentation model.

values of each wood from the corresponding manual measurements and was used to evaluate the measurement accuracy of the method in this study.

TABLE 2. Comparison of ablation experiment results.

Model	Ghostconv	E-SN	SIOU	Parameters/10^6	FLOPs/10^9	AP@0.5:0.95/%
YOLOv5n-Seg				7.41	25.9	96.2
A	√			6.2	23.5	92.1
B		√		4.1	17.9	95.5
C			√	7.41	25.9	96.8
D	√	√		2.88	15.5	95.1
E	√	√	√	2.88	15.5	96.5

TABLE 3. Comparison of different segmentation models.

Model	Parameters/10^6	FLOPs/10^9	AP@0.5:0.95/%
Mask R-CNN	43.97	111	92.9
Cascade Mask R-CNN	77.33	1678	93.2
YOLACT	34.73	61.43	91.4
SOLO	36.3	101	89.7
YOLOv5	18.8	6.7	94.8
This article	2.88	15.5	96.5

FIGURE 15. Test scene diagram.

2. Mean absolute error of diameter measurement for each diameter class: The absolute errors of the measurements in each of the 12 diameter classes were averaged to evaluate the average accuracy of the wood diameter measurements in the different diameter classes.

3. 5 mm error ratio: The number of 120 wood measurements with an absolute error greater than or equal to 5 mm was counted, and their percentage of the total was calculated and used to evaluate the percentage of larger errors in the measurements.

4. Trail-level false detection rate: 120 pieces of wood were classified into different diameter grades according to the measured values, compared with the real labeled diameter grades, and the ratio of the number of false detections of the diameter grades was calculated, which was used to evaluate

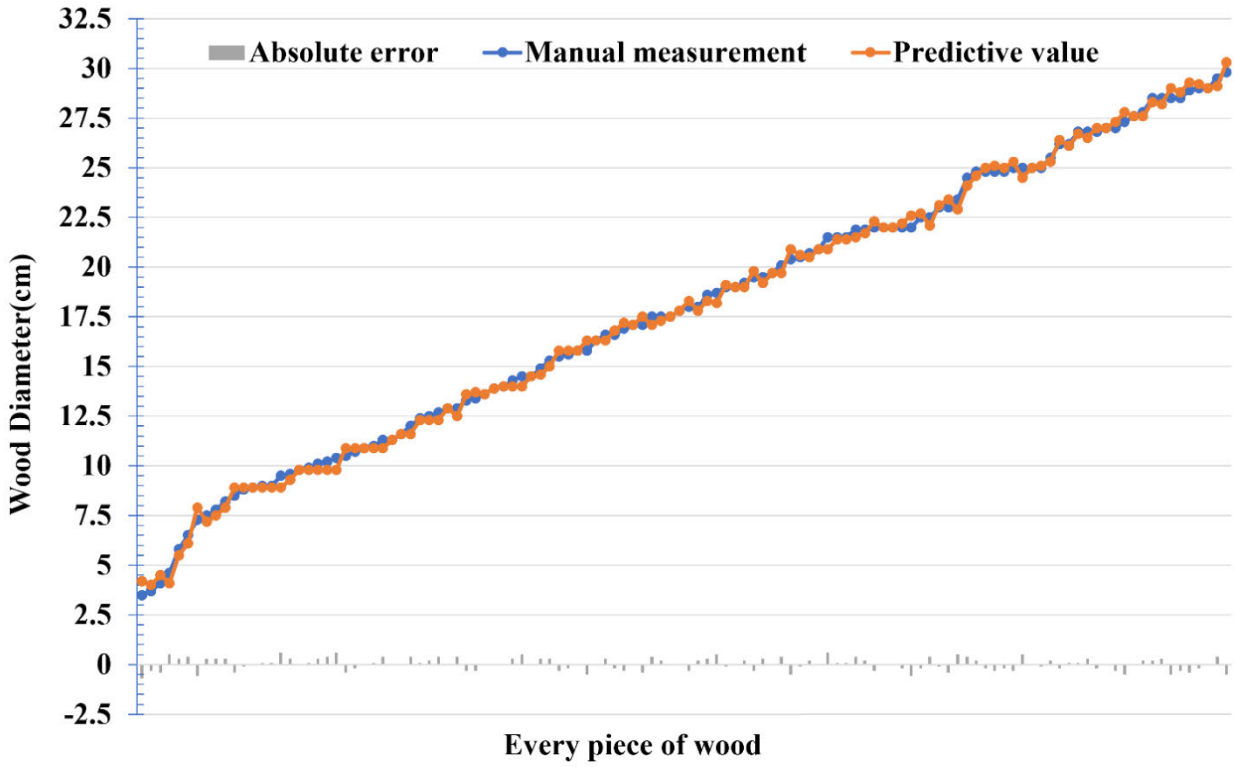

FIGURE 16. Comparison chart of visual measurement accuracy.

the credibility of the measurement method applied to the automatic sorting of wood diameter grades.

5. Leakage rate: The number of wood roots with unmeasured diameters was counted and their ratio to the total was calculated for evaluating the robustness of the measurement method.

6. Average single-root measurement time: To validate the efficiency of this method, the average time required for visual measurement of each piece of wood was calculated.

7. Average Single Sorting Completion Time: In order to verify the working efficiency of the automatic sorting system for wood diameter grade, the time required to complete the sorting of all 120 pieces of wood was counted, and the average value was taken at the end.

TABLE 4. Evaluation results of each index.

Evaluation index	5mm error ratio	Diameter class false detection rate	Missing rate	Average single measurement time
error	12.7%	8.3%	1.7%	0.5S/ Single

TABLE 5. Absolute error of measurement for each group of grades.

Groups	A	B	C	D	E	F
Error/cm	0.42	0.2	0.24	0.21	0.24	0.18
Groups	G	H	I	J	K	L
Error/cm	0.22	0.26	0.27	0.24	0.19	0.31

2) EXPERIMENTAL RESULTS AND ANALYSIS

Using the method of this study, 120 wood diameters were measured, 118 woods were detected and 2 were missed. A line graph corresponding to the measured and manually measured values of 118 woods was made as shown in Figure 11, with the y-axis representing the diameter of the wood in cm,and the x-axis representing each wood. The graph shows that the measured value is not much different from the manual measurement value, in order to visualize the measurement accuracy of the method of this study, a bar graph is added to the graph to represent the absolute error of the measurement value of each wood, and it can be seen that most of the error value is less than 5 mm, and the measurement accuracy is high.

In order to further explore the reliability and efficiency of the method of this study, the average absolute error of diameter measurement of each diameter level, 5 mm error ratio, diameter level misdetection rate, leakage detection rate, average single measurement time and average single sorting completion time were taken as the performance evaluation indexes of the sorting system, and the results are shown in Tables 4 and 5.

After analyzing the results, due to the experiment set up by the 12 diameter level is more fine, the existence of the error led to the wood was divided into adjacent diameter level, resulting in 8.3% of the diameter level misdetection, the actual application can be divided into different diameter levels to avoid the diameter level misdetection. Leakage rate of 1.7%, 5 mm misdetection rate of 12.7%, resulting in 5 mm error is mainly due to two reasons, one is due to the wood was soaked in sewage caused by the color of the section and the skin color is similar to the skin, the model will be included in the division of the skin to make the final measurement result is large, the second is the timber section is too tilted, resulting in the section used to measure the diameter of the two points are not at the same level and the difference is large, which makes the measurement result is small. The result is small. The average single-root measurement time is 0.3 S/root, which meets the demand for production speed.

The average absolute errors of wood measurements for the 12 size classes are shown in Table 5, with the average absolute error for the smallest size class A (3.5-8 cm) being large compared to the other size classes, the average absolute errors for the B-K size classes being less than 0.3 cm, and the average absolute error for the largest size class L (28-29.8 cm) being second only to that of the A size class. There are many factors that cause measurement errors, such as the model segmentation error of the wood cross-section, ellipse fitting error and binocular vision ranging error, the smallest diameter level of wood measurement is more sensitive to the model segmentation and binocular vision ranging error while the largest diameter level of wood measurement is more sensitive to the ellipse fitting error, which can be further optimized in the algorithm. Overall, the measurement error of this research method is within 0.5 cm, and the leakage rate is only 1.7%, which has the advantages of higher accuracy and robustness to meet the production needs of wood diameter grade sorting.

V. CONCLUSION

In this paper, a wood diameter grade online automatic sorting system is designed, which mainly contains two parts: machine vision detection system and PLC control system, and through the information transmission and coordinated operation of each module, it realizes the automated online sorting of wood diameter grade, and an average of 10 pieces of wood can be sorted per minute. In addition, a wood diameter measurement algorithm based on deep learning combined with binocular camera is proposed, and the measurement error can be controlled within 5 millimeters, and the average single-root measurement time is 0.3 seconds, which ensures the efficiency and accuracy of the wood sorting system. It can be seen that the system can meet the production needs of automation in the timber industry, which not only improves the efficiency of timber sorting, but also reduces the workload of the sorters and the cost of timber sorting, and helps the development of the timber production and manufacturing industry.

REFERENCES

[1] Y. Chen, "A re-analysis of the concept of carbon neutrality," *China Population, Resour. Environ.*, vol. 32, no. 4, pp. 1–12, Feb. 2022, doi: 10.12062/cpre.20220203.

[2] S. Y. Yi, "Research on existing problems and countermeasures of forestry ecological protection in my country," *Agricult. Develop. Equip.*, vol. 1, pp. 91–93, 2023, doi: 10.3969/j.issn.1673-9205.2023.01.037.

[3] Q. Cai, S. Z. Chen, and R. Zhao, "Interaction relationship between central forestry investment and forestry economic growth in China," *Scientia Silvae Sinicae*, vol. 51, no. 9, pp. 126–133, Sep. 2015.

[4] Z. Y. Huang and Y. Liu, "Application of grating projection-based measurement technology in cutting rough edge planks," *Forestry Machinery Woodworking Equip.*, vol. 39, no. 5, pp. 18–19, May,2011, doi: 10.3969/j.issn.2095-2953.2011.05.006.

[5] M. Maltamo, "Estimation of timber volume and stem density based on scanning laser altimetry and expected tree size distribution functions," *Remote Sens. Environ.*, vol. 90, no. 3, pp. 319–330, Apr. 2004, doi: 10.1016/j.rse.2004.01.006.

[6] Z. Pásztory, B. Heinzmann, and M.-C. Barbu, "Manual and automatic volume measuring methods for industrial timber," *IOP Conf. Ser., Earth Environ. Sci.*, vol. 159, Jun. 2018, Art. no. 012019, doi: 10.1088/1755-1315/159/1/012019.

[7] A. Baskar, *Digital Image Processing*. Upper Saddle River, NJ, USA: Prentice-Hall, 1964, pp. 1–20.

[8] F. Budiman, R. Mardiyanto, and R. Rachmat, "A handy and accurate device to measure smallest diameter of log to reduce measurement errors," in *Proc. Int. Seminar Intell. Technol. Appl. (ISITIA)*, Jul. 2016, pp. 423–428, doi: 10.1109/ISITIA.2016.7828697.

[9] S. Yella and M. Dougherty, "Automatically detecting the number of logs on a timber truck," *J. Intell. Syst.*, vol. 22, no. 4, pp. 417–435, Dec. 2013, doi: 10.1515/jisys-2013-0026.

[10] L. Yan, X. Gao, T. Y. Wang, L. Pang, L. Su, and Y. Kang, "A measurement method of log diameter for rotary cutting machines based on image processing," *J. Intell. Syst.*, vol. 55, no. 5, pp. 125–133, 2019, doi: 10.11707/j.1001-7488.20190514.

[11] G. H. Chen, Q. Zhang, M. Chen, and H. Yin, "Rapid detection algorithms for log diameter classes based on binocular vision," *J. Beijing Jiaotong Univ.*, vol. 42, no. 2, pp. 22–30, Jan. 2018, doi: 10.1109/icsai.2017.8248480.

[12] A. Kruglov and E. Shishko, "Log pile measurement through 3D modeling," in *Proc. 40th Int. Conf. Telecommun. Signal Process. (TSP)*, Jul. 2017, pp. 263–266, doi: 10.1109/TSP.2017.8075983.

[13] A. V. Kruglov, "The algorithm of the roundwood volume measurement via photogrammetry," in *Proc. Int. Conf. Digit. Image Comput., Techn. Appl. (DICTA)*, Nov. 2016, pp. 1–5, doi: 10.1109/dicta.2016.7797088.

[14] A. Kruglov, "The image analysis algorithm for the log pile photogrammetry measurement," *WSEAS Trans. Signal Process.*, vol. 13, pp. 134–145, Apr. 2017.

[15] N. Samdangdech and S. Phiphobmongkol, "Log-end cut-area detection in images taken from rear end of eucalyptus timber trucks," in *Proc. 15th Int. Joint Conf. Comput. Sci. Softw. Eng. (JCSSE)*, Jul. 2018, pp. 1–6, doi: 10.1109/JCSSE.2018.8457388.

[16] H. Tang, K. Wang, J. Gu, X. Li, and W. Jian, "Application of SSD framework model in detection of logs end," *J. Phys., Conf.*, vol. 1486, no. 7, Apr. 2020, Art. no. 072051, doi: 10.1088/1742-6596/1486/7/072051.

[17] P. P. Yu, Y. H. Lin, Y. F. Lai, S. Y. Cheng, and P. J. Lin, "Dense log end face detection method using the hybrid of BiFPN and YOLOv5s," *J. Forestry Eng.*, vol. 8, no. 1, pp. 126–134, Oct. 2022, doi: 10.13360/j.issn.2096-1359.202204006.

[18] Y. H. Lin, H. L. Zhao, Z. C. Yang, and M. Lin, "An equal length log, volume inspection system using deep-learning and Hough transformation," *J. Forestry Eng.*, vol. 6, no. 1, pp. 136–142, Aug. 2020, doi: 10.13360/j.issn.2096-1359.202003022.

[19] P. Yang, J. Zheng, Z. Feng, Z. Ding, S. Li, Q. Huang, and L. Kong, "Research on detection and segmentation methods of dense-stacked logs using mask R-CNN," *J. Forestry Eng.*, vol. 7, no. 2, pp. 135–142, Oct. 2021, doi: 10.13360/j.issn.2096-1359.202106001.

[20] N. Ma, X. Zhang, H. T. Zheng, and J. Sun, "Shufflenet v2: Practical guidelines for efficient CNN architecture design," *Proc. Eur. Conf. Comput. Vis. (ECCV).*, 2018, pp. 116–131, doi: 10.1007/978-3-030-01264-9_8.

[21] C. R. Liu, "Application of improved residual network in sensing image classification," *Sci. Technol. Eng.*, vol. 21, no. 31, pp. 13421–13429, Nov. 2021, doi: 10.3969/j.issn.1671-1815.2021.31.029.

[22] K. Han, Y. Wang, Q. Tian, J. Guo, C. Xu, and C. Xu, "GhostNet: More features from cheap operations," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2020, pp. 1577–1586, doi: 10.1109/cvpr42600.2020.00165.
[23] J. Huang, "Potato seed bud eye detection algorithm based on lightweight convolutional neural network," *Trans. Chin. Soc. Agricult. Eng.*, vol. 39, no. 9, pp. 172–182, Mar. 2023, doi: 10.11975/j.issn.1002-6819.202303035.
[24] X. B. Liu, "Object detection method based on CloU improved bounding box loss function," *Chin. J. Liquid Crystals Displays*, vol. 38, no. 5, pp. 656–665, May 2023.
[25] R. J. Ma, "Recognition method for maturity of Pitaya based on improved YOLOv5," *J. Shenyang Agricult. Univ.*, vol. 54, no. 2, pp. 196–206, Feb. 2023, doi: 10.3969/j.issn.1000-1700.2023.02.008.
[26] Q. Gao, Y. Pan, L. Zhu, and J. Yan, "Improved YOLOv5 remote sensing target detection method based on SIOU function," *Changjiang Inf. Commun.*, vol. 35, no. 11, pp. 5–8, Nov. 2022, doi: 10.3969/j.issn.1673-1131.2022.11.002.
[27] H. Zhang, L. An, Q. Zhang, Y. Guo, X. Song, and Q. Gao, "SGBM algorithm and BM algorithm analysis and research," *Geomatics Spatial Inf. Technol.*, vol. 35, no. 11, pp. 5–8, Oct,2016, doi: 10.3969/j.issn.1672-5867.2016.10.063.
[28] K. He, G. Gkioxari, P. Dollár, and R. Girshick, "Mask R-CNN," in *Proc. IEEE Int. Conf. Comput. Vis. (ICCV)*, Oct. 2017, pp. 2980–2988.
[29] W. He, C. Li, X. Nie, X. Wei, Y. Li, Y. Li, and S. Luo, "Recognition and detection of aero-engine blade damage based on improved cascade mask R-CNN," *Appl. Opt.*, vol. 60, no. 17, pp. 5124–5133, Jun. 2021.
[30] D. Bolya, C. Zhou, F. Xiao, and Y. J. Lee, "YOLACT: Real-time instance segmentation," in *Proc. IEEE/CVF Int. Conf. Comput. Vis. (ICCV)*, Oct. 2019, pp. 9156–9165, doi: 10.1109/ICCV.2019.00925.
[31] X. Wang, T. Kong, C. Shen, Y. Jiang, and L. Li, "SOLO: Segmenting objects by locations," in *Computer Vision—ECCV*. Glasgow, U.K.: Springer, 2020, pp. 649–665, doi: 10.1007/978-3-030-58523-5_38.

JIAXIN LIU was born in Jiujiang, Jiangxi, in 1996. He received the B.S. degree in measurement and control technology and instrumentation from the Jincheng College, Nanjing University of Aeronautics and Astronautics. He is currently pursuing the M.S. degree in mechanical engineering with the Fujian University of Technology. His research interests include machine vision and image processing technology.

LINGHUA KONG received the bachelor's degree in physics from Nankai University, in 1983, the master's degree in physics from the Institute of High Energy Physics, Chinese Academy of Sciences, in 1988, and the Ph.D. degree in mechanical engineering from the Department of Mechanical Engineering, McGill University, Canada, in 2004. He received a Postdoctoral Fellowship with the Georgia Institute of Technology, in 2005. He is currently a Professor with the School of Mechanical and Automotive Engineering, Fujian University of Technology. He designed and developed a variety of new products and equipment, and obtained 15 patents: published 20 influential articles included in SCI/EI as the first author. His main research interests include multispectral and plasma fields.

JISHI ZHENG received the Ph.D. degree in engineering from Central South University, in 2015. From January 2019 to July 2019, he was a Visiting Scholar with the Robotics Laboratory, Department of Computer and Electronic Engineering, University of Essex, U.K. He is currently a person in charge of the Internet of Things and the Director of the Department of Traffic Information and Control. He has presided over and participated in more than ten provincial and municipal scientific research projects. His main research interests include the application of artificial intelligence in the industry and the research on the flight control algorithm of drones. He is also a part-time Executive Director of the Fujian Aeronautical Society.

ZHIGANG DING received the master's degree in bioengineering from Jilin University, in 2007. He has been a Senior Engineer with the Fujian University of Technology, since 2008. He is currently the Head of the Vehicle Engineering Laboratory. He has participated in more than ten national 863 programs and provincial and municipal scientific research projects. He is mainly engaged in the research and development of new energy vehicle power system assembly, product visual inspection, and intelligent equipment.

LI FENG received the B.S. degree in electronic information engineering from Pingxiang College, in 2020. She is currently pursuing the M.S. degree in control engineering with the Donghua University of Science and Technology. Her research interests include radon solution preparation system control and design.

YANLU LV was born in Xiamen, Fujian, in 1997. He received the bachelor's degree in mechanical manufacturing and automation from the School of Physics and Electromechanics, Longyan University. He is currently pursuing the master's degree in mechanical engineering with the School of Mechanical and Automotive Engineering, Fujian University of Technology. His research interests include machine vision and image processing.