

Article

3D Positioning Method for Pineapple Eyes Based on Multiangle Image Stereo-Matching

Anwen Liu ¹, Yang Xiang ^{1,*} , Yajun Li ^{1,2}, Zhengfang Hu ¹, Xiufeng Dai ¹, Xiangming Lei ¹ and Zhenhui Tang ¹¹ College of Mechanical and Electrical Engineering, Hunan Agriculture University, Changsha 410128, China² Intelligent Equipment Research Center, Beijing Academy of Agriculture and Forestry Sciences, Beijing 100097, China

* Correspondence: xy@hunau.edu.cn; Tel.: +86-13319585960

Abstract: Currently, pineapple processing is a primarily manual task, with high labor costs and low operational efficiency. The ability to precisely detect and locate pineapple eyes is critical to achieving automated pineapple eye removal. In this paper, machine vision and automatic control technology are used to build a pineapple eye recognition and positioning test platform, using the YOLOv5 target detection algorithm to quickly identify pineapple eye images. A 3D localization algorithm based on multiangle image matching is used to obtain the 3D position information of pineapple eyes, and the CNC precision motion system is used to pierce the probe into each pineapple eye to verify the effect of the recognition and positioning algorithm. The recognition experimental results demonstrate that the mAP reached 98%, and the average time required to detect one pineapple eye image was 0.015 s. According to the probe test results, the average deviation between the actual center of the pineapple eye and the penetration position of the probe was 1.01 mm, the maximum was 2.17 mm, and the root mean square value was 1.09 mm, which meets the positioning accuracy requirements in actual pineapple eye-removal operations.

Keywords: pineapple eye; three-dimensional; YOLOv5; stereo-matching

Citation: Liu, A.; Xiang, Y.; Li, Y.; Hu, Z.; Dai, X.; Lei, X.; Tang, Z. 3D Positioning Method for Pineapple Eyes Based on Multiangle Image Stereo-Matching. *Agriculture* **2022**, *12*, 2039. <https://doi.org/10.3390/agriculture12122039>

Academic Editor: Yanbo Huang

Received: 13 October 2022

Accepted: 26 November 2022

Published: 28 November 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2022 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

Pineapple is a fruit with a high added economic value. In 2018, China's yearly pineapple production was approximately 1.64 million tons [1]. Approximately 30% of pineapples are utilized for production and processing [2]. The processing of pineapple is complicated, especially because even after the pineapple has been skinned, there are still many eyes on its surface that need to be removed. Currently, the main way to remove pineapple eyes is to do so manually with special tools, which is labor intensive and has high labor costs and low production efficiency. The key to automatically removing pineapple eyes is to rapidly and accurately identify and locate pineapple eyes.

Machine vision technology is frequently utilized in fruit recognition and quality inspection because of its noncontact nature, high speed, and high precision [3]. In traditional machine vision technology, targets are primarily recognized based on characteristics such as color, shape, and texture. Li et al. [4] proposed a field recognition system for pineapple based on monocular vision through threshold segmentation, morphological processing, and other operations to recognize pineapples and obtain pineapple center point information. Lin et al. [5] presented a segmentation method using texture and color features, and Leung-Malik textures and HSV color features were fused to realize the detection and recognition of citrus fruit. Lv et al. [6] proposed a method to deepen the fruit region and improve the edge definition in images by using a histogram equalization algorithm. Then, the R-B color difference image based on histogram equalization was obtained, and green apple recognition was realized. Kurtulmus et al. [7] used circular Gabor texture analysis for green citrus object recognition. When the fruit surface is uneven in color, shadowed, or obscured

due to environmental factors such as light, the recognition quality of traditional machine vision technology is significantly reduced [8].

By applying machine learning technology to fruit image analysis, a better application effect and higher efficiency can be obtained [9]. Li Han et al. [10] used a naive Bayes classifier to classify fruit and nonfruit regions, and the interference caused by the color similarity of green tomatoes and green foliage backgrounds was eliminated to improve the fruit recognition accuracy. Wang et al. [11] proposed a litchi recognition algorithm based on K-means clustering, which can better resist the influence of illumination changes and can maintain high accuracy for recognition under occlusion and different lighting conditions. Zhao et al. [12] extracted the Haar-like features of grayscale images and used the AdaBoost classifier for classification and recognition. In the actual environment, the detection accuracy rate of ripe tomatoes reached 96%, and the classifier structure was simple.

In recent years, object detection based on deep learning has shown great advantages in the field of fruit image recognition [13,14]. The convolutional neural network, with its fast detection speed and excellent ability to extract target features, not only reduces the workload but also improves the recognition speed and accuracy [15]. Zhang Xing et al. [16] proposed a study on pineapple picking and recognition under a complex field environment based on the improved YOLOv3. The multiscale fusion training network was used to detect single-category pineapple, and a detection and recognition rate of approximately 95% was achieved using this method. Tian et al. [17] proposed an improved YOLOv3 model to identify apples at different growth stages in orchards. The model was used with the DenseNet method to process low-resolution feature layers; this method effectively enhances feature propagation, promotes feature reuse, improves network performance, and has good recognition performance under apple overlap and occlusion conditions. Yu et al. [18] proposed a mask R-CNN algorithm to identify 100 wild strawberry images. The results demonstrated that the average recognition accuracy was 95.78% and the recall rate was 95.41%. Zhang et al. [19] proposed a real-time detection method for grape clusters based on the YOLOv5s deep learning algorithm. By training and adjusting the parameters of the YOLOv5s model on the data set, the fast and accurate detection of grape clusters was finally realized. The test results showed that the precision, recall, and mAP of the grape cluster detection network were 99.40%, 99.40%, and 99.40%, respectively.

Studies related to fruit positioning, which are widely used, have mainly focused on the three-dimensional positioning of fruit for robot picking, and methods include binocular stereo vision, structured light stereo vision, and monocular stereo vision. In binocular stereo vision, not only can the image information of different angles of the target be obtained, but the three-dimensional position information of the target through stereo matching can also be obtained [20]. Therefore, this is a widely used approach in fruit and vegetable recognition [21], positioning [22], and acquisition of phenotypic parameters [23]. Luo et al. [24] proposed a method for solving and positioning enclosure based on binocular stereo vision. When the depth distance was within 1000 mm, the positioning error was less than 5 mm. However, the calibration process of the binocular camera is complex, and the calculational burden of the algorithm was relatively large [25]. Stereovision, which is based on structured light, is a combination of structured light technology and binocular stereo vision technology. Through structured light matching, the corresponding pixels of the left and right cameras are subjected to stereo matching, parallax calculation, and recovery of the three-dimensional data of the scene. Zhang et al. [26] used a machine vision system based on a near-infrared array structure and three-dimensional reconstruction technology to realize the recognition and positioning of apple stems and calyxes. However, structured light stereo vision is easily affected by illumination [27]. Monocular stereo vision positioning can be divided into monocular camera positioning of one, two, or more images. The positioning of a single image mainly relies on the mapping relationship between the known spatial information of the characteristic light points, lines, or other image features to obtain the position coordinate information [28]. Generally, images from different perspectives are obtained using the positioning method by changing the position

of the camera, and the matching relationship of image feature points is used to obtain the relative positional relationship between the cameras during multiple shots to realize the positioning of the target. Zhao et al. [29] used a monocular color camera to build a vision system to locate the picking point of litchi clusters and realize the three-dimensional positioning of litchi clusters.

To date, there have been no research reports on pineapple eye machine-vision recognition or positioning. Based on the analysis of the existing research in the field of fruit recognition and positioning, deep learning technology based on convolutional neural networks is proposed in this paper to carry out research on pineapple eye recognition. On this basis, combined with the entire circumference-image-acquisition-of-pineapple method, the three-dimensional localization of pineapple eyes is realized by applying the stereo-matching method of monocular and multiangle images.

2. Materials and Methods

2.1. Structure and Working Principle of the Test Platform

The structure of the pineapple eye recognition and positioning test platform is shown in Figure 1. The notebook is an HP-Shadow Elf equipped with an Intel i7-10750H CPU@2.60 GHz processor, 16 GB RAM, and an NVIDIA GeForce GTX1650Ti graphics card. The 64-bit Windows 10 operating system is installed, and the software development environment is Visual Studio2017 + OpenCV4.0.0. The color camera is an Imaging Source DFK41BU02 with a resolution of 1280(H) × 960(V), a frame rate of 15 fps, and an 8.5 mm Computar lens. A CR-9600-R ring light source is installed directly under the camera lens. The Mitsubishi FX3U-32MT PLC controller is used as the control core, and the PLC is connected to the notebook through the serial communication port. The motion platform is composed of a clamping cylinder, servo motor, linear slide, probe cylinder, and probe. The peeled pineapple is clamped using the clamping cylinder and rotated at a precise angle by the servo motor to acquire the entire circumference of the pineapple image. In this paper, a probe is used to evaluate the accuracy of the identification and positioning algorithm. The probe is installed on the probe cylinder and can be inserted into the pineapple through the telescopic movement of the probe cylinder. The probe cylinder, which can accurately move, is installed and positioned in the direction parallel to the pineapple axis.

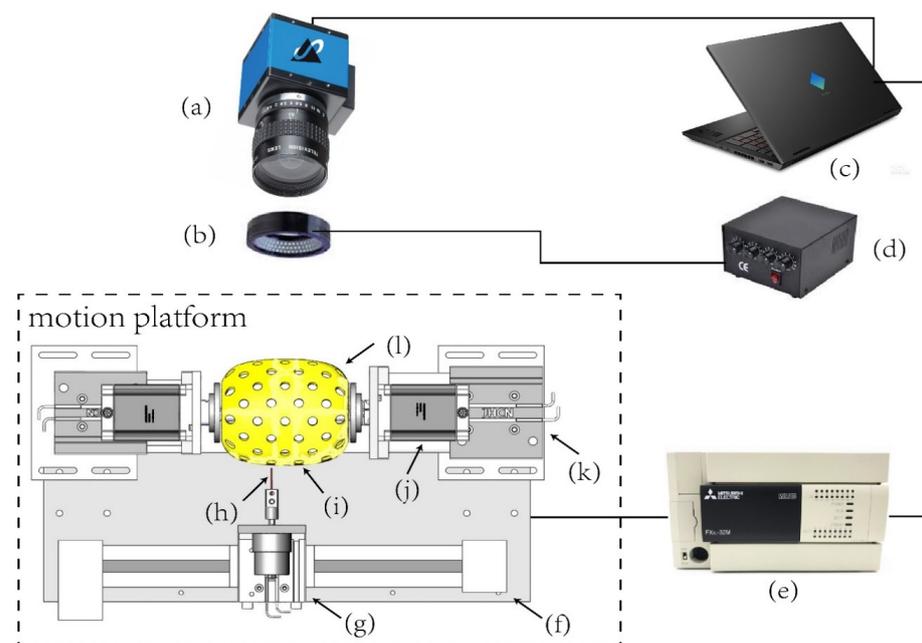


Figure 1. Structure of the test platform. (a) color camera, (b) ring light source, (c) notebook, (d) light source controller, (e) PLC controller, (f) linear slide, (g) probe cylinder, (h) probe, (i) pineapple eye, (j) servo motor, (k) clamping cylinder, and (l) pineapple.

2.2. Image Acquisition of Pineapple Eyes

Goodfarmer Philippine pineapples, which were manually peeled and placed on the test platform for image acquisition, were used for the experiments. Before image acquisition, the dot calibration plate was used to reduce the lens distortion and perspective distortion caused by the tilt of the camera [30]. To obtain the images of all pineapple eyes and provide a sufficient number of images for multiangle image stereo matching, images of pineapples were collected in 60° intervals, and 6 images were collected for each pineapple. Figure 2 shows images of the same pineapple collected from different angles. From this figure, there are obvious differences in the shape and size of pineapple eyes.

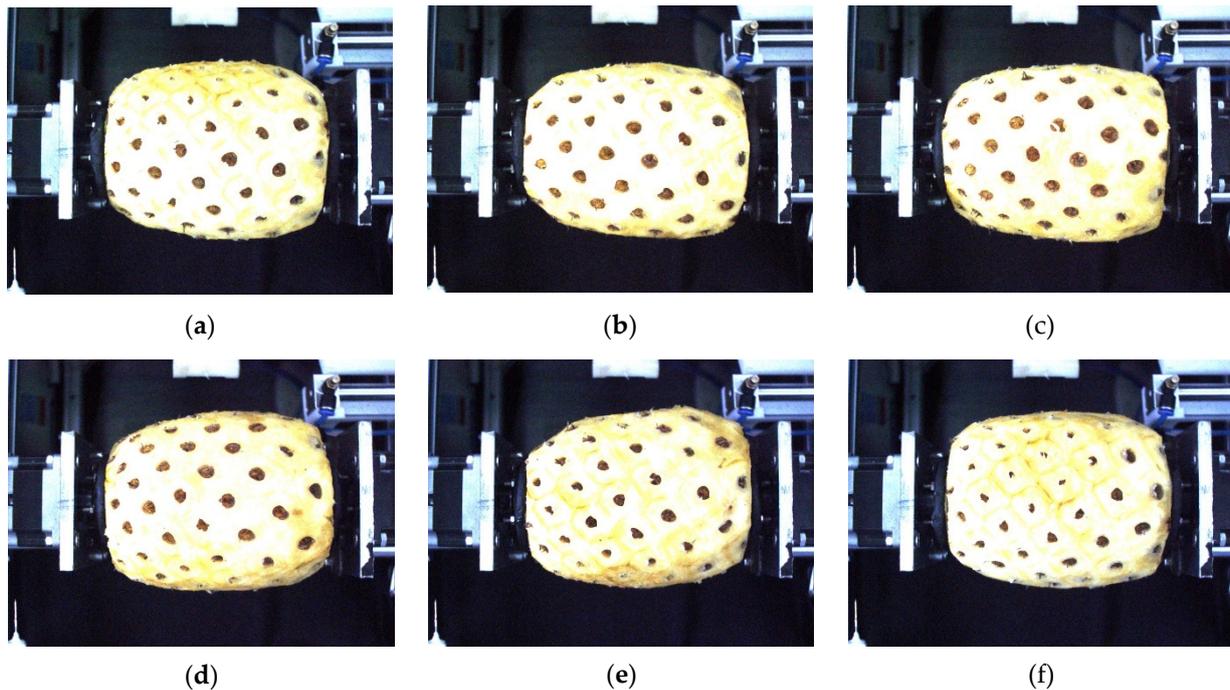


Figure 2. Images of the same pineapple at different angles. (a) 0 degrees; (b) 60 degrees; (c) 120 degrees; (d) 180 degrees; (e) 240 degrees; (f) 300 degrees.

2.3. Pineapple Eye Recognition Algorithm Based on YOLOv5

In this paper, YOLOv5 is selected as the target detection network for pineapple eye recognition. Among the commonly used object detection networks, strong detection performance is achieved with the YOLOv5 network [31], which uses mosaic data enhancement, adaptive anchor frame calculation, and adaptive image scaling at the input end. In the backbone network, the features of the target adopted through Focus and CSPNet (cross-stage partial network) can be quickly extracted. In the neck network, FPN (feature pyramid network) and PANet are used for multiscale fusion of the extracted features. GIoU (generalized intersection over union) loss is used as the loss function of the target detection frame in the output end. NMS (nonmaximum suppression) is introduced to filter out the overlapping candidate frames and obtain the best prediction output. These improvements ensure the detection accuracy and speed of small targets and have the advantages of a shallow structure, small weight file, and relatively low requirements for the configuration of the mounted equipment.

There are 4 versions of YOLOv5 [32]: YOLOv5s, YOLOv5m, YOLOv5l, and YOLOv5x. The width and depth of the YOLOv5s model are the initial values. This model is small and fast and is suitable for the detection of small and simple datasets. The YOLOv5m and YOLOv5x models have the deepest depths and are suitable for detection on large and complex datasets. As the depth of the network increases, the detection accuracy is improved, while the detection speed is reduced. In YOLOv5, the learning ability of

the neural network improves, the amount of calculation is reduced, and high detection accuracy is maintained. To maximize the detection speed while maintaining sufficient detection accuracy, YOLOv5l is used in this paper as the pineapple eye detection model. The structure of YOLOv5l is shown in Figure 3.

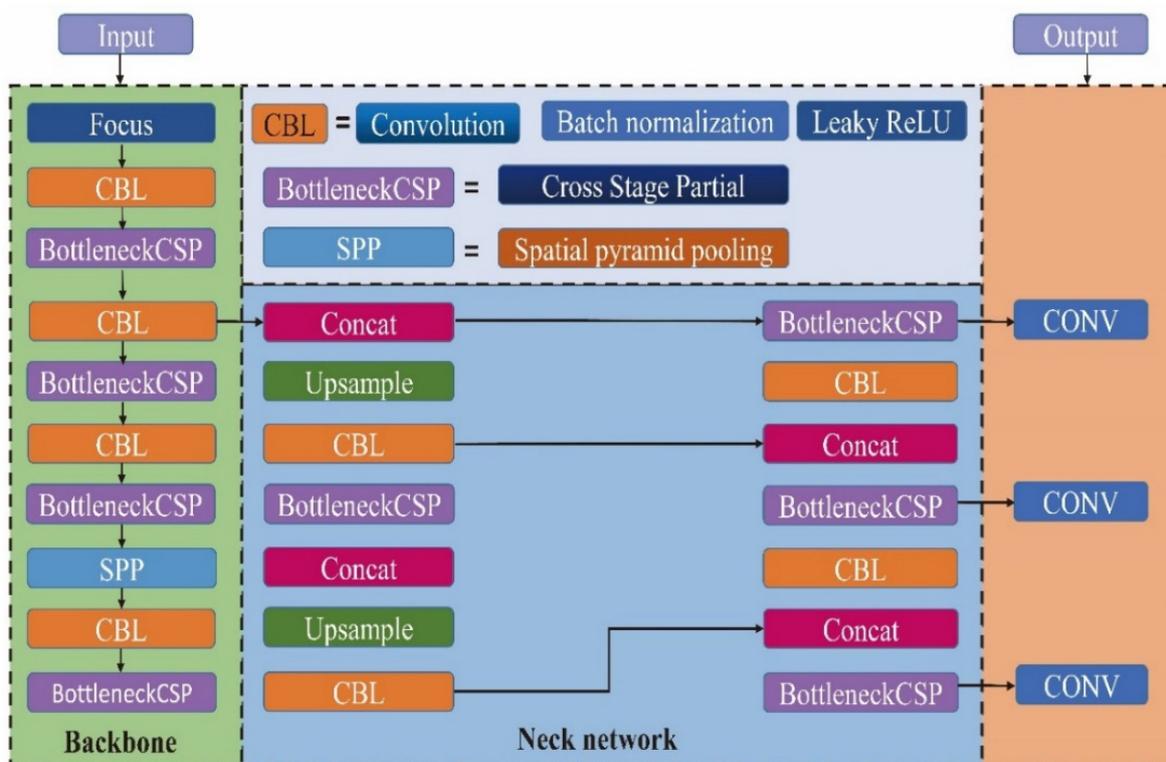


Figure 3. YOLOv5l model structure.

To construct the experimental dataset, 240 pineapple images were obtained from 40 pineapples. Then, the image was processed with data enhancements, such as rotation and horizontal and vertical mirror images, to improve the robustness of the recognition mode, and 600 pineapple images were finally obtained, with a total of approximately 18,000 pineapple eyes. The pineapple eye images were manually labeled one by one by labeling software. Pineapple eyes in the image were marked with a rectangular box and then named P. The labeling information was stored in the PASCALVOC (Pattern Analysis, Statical Modeling and Computational Learning, Visual Object Classes) format [33], in which the coordinates, labels, and serial numbers of each box are contained. The pineapple eye image, labeled data, and other files were saved according to the PASCALVOC dataset directory structure to build the pineapple eye dataset.

The 600 pineapple eye images enhanced by the dataset were divided into a training set, validation set, and test set at an 8:1:1 ratio. Because the size of the pineapple eye target is small, to improve the detection accuracy, the input size is 640×640 pixels, 32 images were taken as a batch, and the weight parameters were updated once for each batch of images trained.

YOLOv5 incorporates the current mainstream detection approach FPN (feature pyramid network) [34] and inherits the grid generation idea of the YOLO algorithm. The 640×640 feature plot is divided into grid areas of equal size $S \times S$ cells (usually 80×80 , 40×40 , or 20×20). After maximum suppression, the output end of the network outputs the prediction information of all grid information. The prediction information of each grid includes the classification probability and confidence of the target as well as the center coordinates and length and width of the box surrounding the detection target. The classification probability represents the classification information of the predicted target in the grid region, and the confidence represents the probability of the detection target in the

grid region. The central coordinates and length-and-width information of the box represent the specific size and position of the target predicted by the grid.

2.4. Three-Dimensional Positioning Algorithm for Pineapple Eyes

In this paper, images of pineapples are collected using 30° intervals; obviously, the same pineapple eye appears on multiple consecutive images. By analyzing these images and matching the same pineapple eyes in different images, the parallax information of the pineapple eye can be obtained. The depth information of the pineapple eye can be obtained through triangulation. In this paper, two images with an angle difference of 90° are used as a group for stereo matching analysis to obtain the three-dimensional position information of all pineapple eyes. Considering the high similarity of pineapple images from different angles, the traditional stereo-vision-matching algorithm is not expected to perform well. In addition, a large amount of calculation is required in this algorithm, which also has low efficiency. Therefore, this algorithm is not suitable for the needs of actual production. Figure 4a,b show the comparison of the γ degree and $\gamma + 90$ -degree pineapple eye images. Here, one pineapple eye appears in both images.

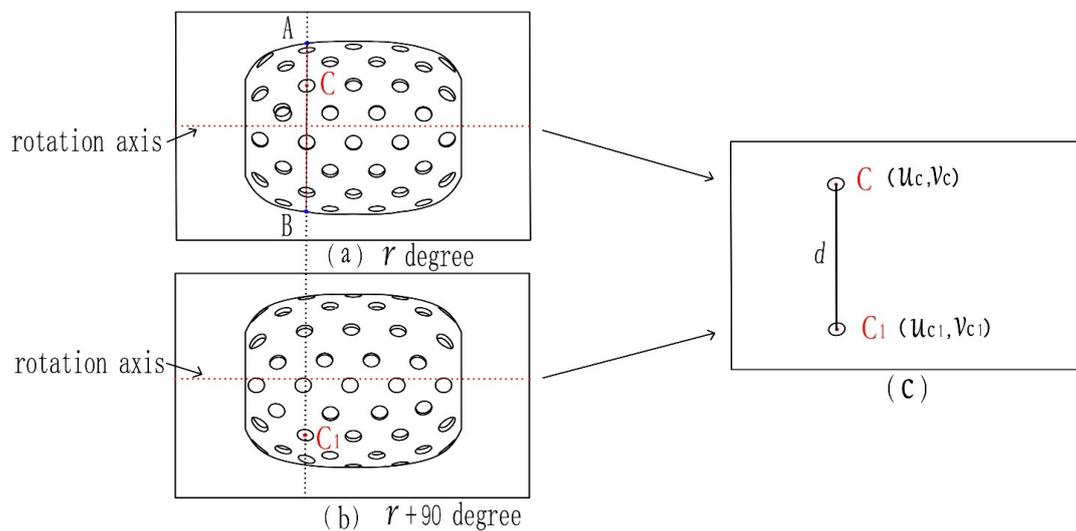


Figure 4. Epipolar constrained stereo matching. (a) γ degree; (b) $\gamma + 90$ degree; (c) calculate d of distance schematic diagram.

The central coordinates (u_c, v_c) and (u_{c1}, v_{c1}) are used to describe the position of the pineapple eye in the two images. Therefore, the position of the pineapple eye in the two images must satisfy the following two constraints: (1) the center of the pineapple eye is located on the same vertical line in the two images, that is, $v_{c1} = v_c$. (2) The row coordinates of the center of the pineapple eye on the two images can be predicted by the displacement of the center of the pineapple eye after the pineapple is rotated by 90°, namely, $u_{c1} = u_c + d$.

In order to obtain the value d in Figure 4c, Figure 5 is used to describe the solution process in detail, f is the focal length, and S is the distance between point O and point p , the optical center of the camera; R is the radius on the contour of the pineapple cross-section through C of the pineapple eye. We can obtain Formula (1).

$$\begin{cases} \eta = \arctan\left(\frac{l_0 d_x}{f}\right) \\ R = S \sin \eta \\ \frac{l_1 d_x}{R \sin r} = \frac{f}{S - R \cos r} \\ d = d_1 + d_2 = R \sin r + R \cos r \end{cases} \quad (1)$$

where d_x represents the physical size of a pixel on the u -axis, which is 0.00465 mm in this paper, η is the angle between the Op and the Ap , and r is the OG and the OC_1 .

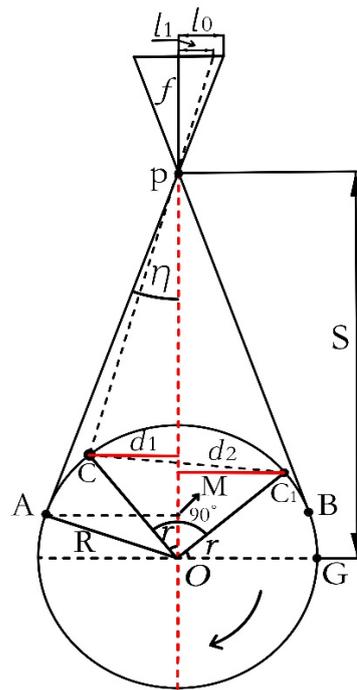


Figure 5. Schematic diagram of the 90° rotation distance of the pineapple eye center point.

Since the contour of the pineapple cross section through C of the pineapple eye is not an ideal circle, and due to system errors such as installation and imaging, u_{c1} and v_{c1} cannot fully meet the above constraints; therefore, a certain tolerance Δ is added when finding a matching pineapple eye in the $\gamma + 90$ degree image. In other words, we search for the target pineapple eye within the rectangular box $(u_{c1} - \Delta, v_{c1} - \Delta, u_{c1} + \Delta, v_{c1} + \Delta)$. To ensure that there is only one pineapple eye in the rectangular box, Δ is set to a third of the minimum distance between the two pineapple eyes in the image. Obviously, according to the above constraints, the pineapple eye below the rotation axis in Figure 4a is not found in Figure 4b, so there is no need to perform a matching operation.

In this paper, a 3D localization algorithm for pineapple eyes based on monocular multiangle image matching is proposed. After obtaining the image coordinates of the same pineapple eye in two images with a difference of 90°, the depth of the pineapple eye is calculated by triangulation. The information is then used to calculate the three-dimensional position information of the pineapple eye. The camera coordinate system $O_c-X_cY_cZ_c$ is established with the camera optical center as the origin, as shown in Figure 6. The center point C of any pineapple eye is selected as the measurement object. (u_c, v_c) represents the pixel value of pineapple eye center C under the imaging plane, O_1 is the intersection of the imaging plane of the pineapple eye center point C and the camera optical axis, and the pixel value is (u_0, v_0) .

The circle in Figure 7 is the cross-sectional profile of the pineapple through point C. ψ is the angle between the line segment OC and the optical axis of the camera, which satisfies the formula $\psi = \arctan\left(\frac{h_1}{h_2}\right)$. The pineapple is rotated clockwise in the direction indicated by the arrow in the figure. p refers to the optical center of the camera. The distance between point C and point p of the camera optical center is W, the distance between point O and point p of the camera optical center is S, and l_1 is the number of pixels in the axial direction of the pineapple eye imaging plane. When the pineapple rotates clockwise by 90°, which is equivalent to a 90° counterclockwise rotation of the camera, as shown in the

dotted line in Figure 8, l_2 is the number of pixels in the axial direction of the pineapple eye imaging plane after rotation. The following formula can be obtained from Figure 7:

$$\begin{cases} \alpha = \arctan\left(\frac{l_1 \times d_x}{f}\right) \\ \beta = \arctan\left(\frac{l_2 \times d_x}{f}\right) \end{cases} \quad (2)$$

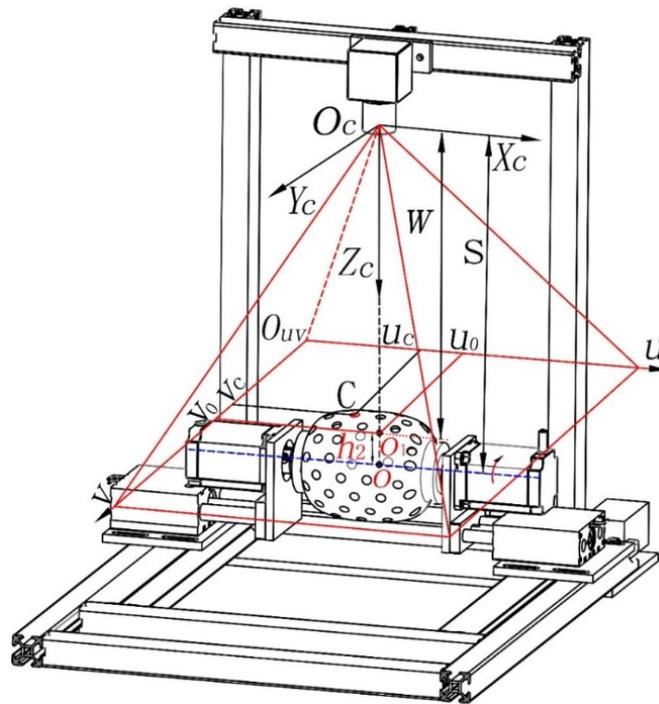


Figure 6. Camera coordinate system for the pineapple eye.

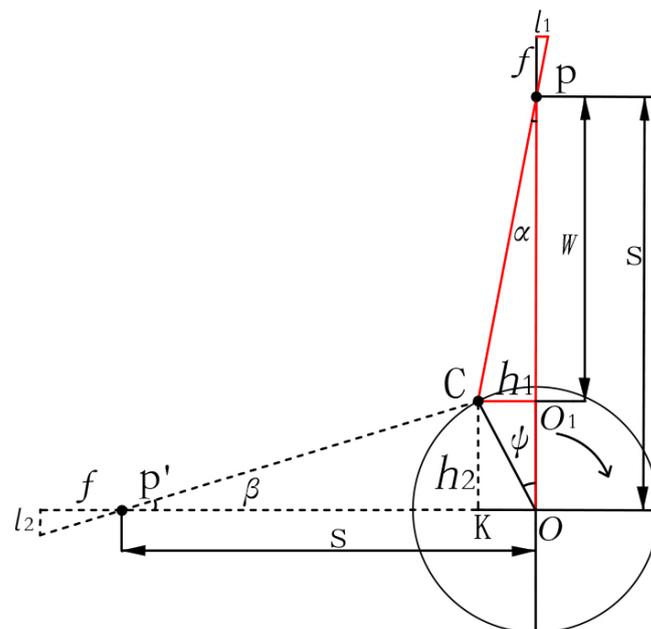


Figure 7. Schematic diagram of the pineapple eye depth information calculation.

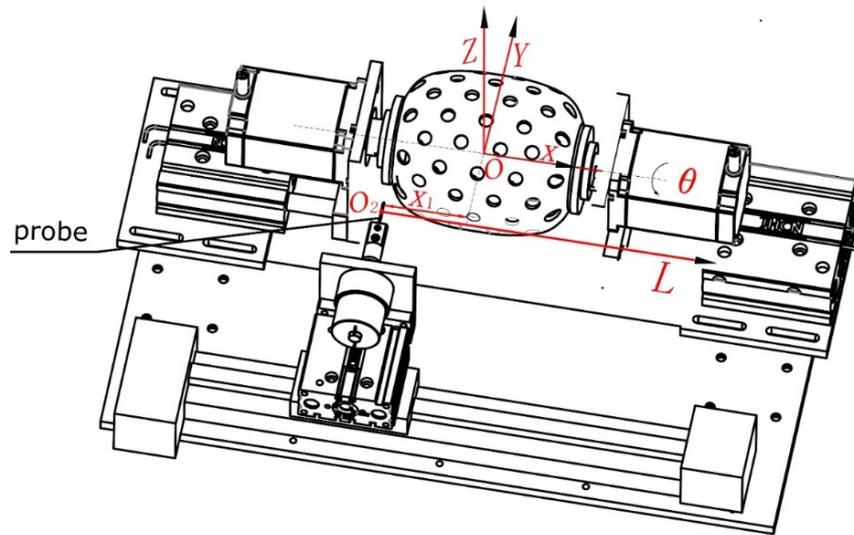


Figure 8. Three-dimensional positioning schematic diagram.

In this formula, d_x represents the physical size of a pixel on the u -axis, which is 0.00465 mm in this paper. α and β can be solved by using Formula (2), and h_1 and h_2 in Figure 7 can be simultaneously solved according to the following equations.

$$\begin{cases} h_1 = \frac{S(1-\tan(\beta))\tan(\alpha)}{1-\tan(\alpha)\tan(\beta)} \\ h_2 = \frac{S(1-\tan(\alpha))\tan(\beta)}{1-\tan(\alpha)\tan(\beta)} \end{cases} \quad (3)$$

In Formula (3), h_1 is the distance between point C and point O_1 , mm. h_2 is the distance between point C and point K, mm. Figure 7 shows that the center point C of the pineapple eye is imaged at time t , and the object distance of the imaging plane is $W = S - h_2$. Then, the number of pixels of pineapple eye point C on the imaging plane and in the camera coordinate system are determined using the following equation:

$$\begin{cases} \frac{d_x(u_c - u_0)}{f} = \frac{X_c}{W} \\ \frac{d_x(v_c - v_0)}{f} = \frac{Y_c}{W} \end{cases} \quad (4)$$

In other words, at time t , the center point C of the pineapple eye fulfills the matrix in the camera coordinate system, with the camera optical center serving as the origin:

$$\begin{bmatrix} X_c \\ Y_c \\ Z_c \end{bmatrix} = W \begin{bmatrix} u_c & -u_0 & 0 \\ v_c & -v_0 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \frac{d_x}{f} \\ \frac{d_x}{f} \\ 1 \end{bmatrix} \quad (5)$$

To facilitate subsequent experiments and the operation of removing pineapple eyes in practical engineering applications, the pineapple three-dimensional space coordinate with O as the center is established $O_{-}XYZ$. The geometric vector approach is used to translate the camera coordinates into the 3D space coordinates of the pineapple eye, as shown in Figure 8.

$$\begin{cases} X = -X_c \\ Y = Y_c \times \cos(-\alpha) + (S - Z_c) \times \sin(-\alpha) \\ Z = S - ((S - Z_c) \times \cos(\alpha) + Y_c \times \sin(\alpha)) \end{cases} \quad (6)$$

Furthermore, the three-dimensional coordinates of the pineapple eye (X, Y, Z) are converted to the probe, which can be used for eye removal after changing to the eye-removal tool. Position L and pineapple rotation angle θ are represented by the space coordinates

(L, θ) , and O_2 is the starting point of the probe. As shown in Figure 9, the corresponding conversion formula is as follows:

$$\begin{cases} L = X_1 - X \\ \theta = \frac{Y}{S-Z} \times 180/PI \end{cases} \quad (7)$$

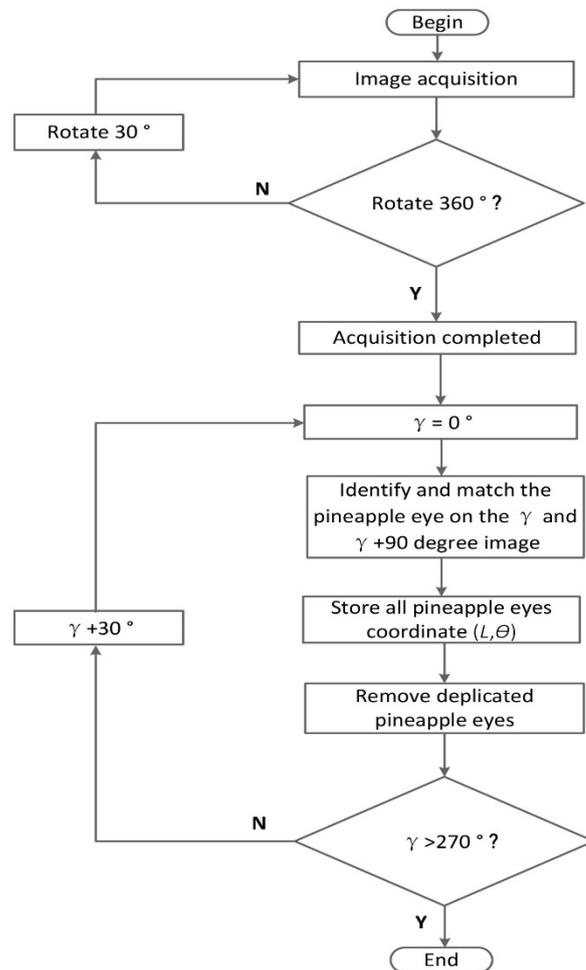


Figure 9. Flow diagram of the 3D positioning method for pineapple eyes.

In Equation (7), X_1 is the distance from the optical center of the horizontal axis camera to the starting point of the probe. Because the pineapple is rotating during the image acquisition process, all the calculated coordinates of the pineapple eyes are the result of the calculation of the current pineapple angle conditions. To obtain the coordinates of all pineapple eyes for the whole pineapple in the same coordinate space, we should reverse rotate the coordinates of all pineapple eyes to the 0° position. Therefore, Formula (7) should be modified to the following:

$$\begin{cases} L = X_1 - X \\ \theta = \frac{Y}{S-Z} \times \frac{180}{PI} - \gamma \end{cases} \quad (8)$$

The position information of all pineapple eyes can be obtained after image stereo matching and pineapple eye position computation. To ensure that the position information of each pineapple eye is calculated, the image acquisition angle interval is set to 30 degrees, which leads to the same pineapple eye being calculated in multiple sets of images. This results in more calculated pineapple eyes than the actual number of pineapple eyes. To avoid the same pineapple eye being repeatedly calculated, a successful match of a pineapple

eye in the image is marked. When using the image and the next picture, the marked pineapple eye does not participate in the matching calculation.

2.5. Flow Diagram of 3D Positioning Algorithm

The flow diagram of the 3D positioning algorithm for pineapple eyes based on multi-angle image stereo matching in the study is shown in Figure 9. It mainly includes all pineapple eye image acquisition to identify and match the pineapple eye on the γ and $\gamma + 90$ -degree image. When matching images on the γ degree and $\gamma + 90$ degree, all the pineapple eye coordinates (L, θ) are stored in a list. When matching the next set of images ($\gamma + 30$ degree and $\gamma + 120$ degree), some pineapple eyes which are duplicated with the previous set of images will inevitably be obtained. Because the pineapple eye coordinate (L, θ) is a global coordinate, the coordinates (L, θ) are approximate. By comparing the newly obtained pineapple eye coordinates with the pineapple eye coordinates stored in the list, it is easy to find and eliminate duplicate pineapple eyes. In this paper, the Euclidean distance judgment is used as the judgment basis; when the distance between the two pineapple eyes is less than 1 mm, the two pineapple eyes are considered to be duplicate pineapple eyes.

2.6. Probe Positioning Test

In this paper, a probe test method is proposed for evaluating the positioning accuracy of the positioning system. The probe mounted on the linear slide, as illustrated in Figure 10, may be accurately moved and positioned in the direction of the pineapple axis. At the same time, the servo drive motor rotates the pineapple at a precise angle. Therefore, according to the coordinates (L, θ) of any pineapple eye, the probe can be moved to the position of the pineapple eye and inserted into the pineapple eye through the extension action of the probe cylinder. The deviation er (error) between the actual center of the pineapple eye and the probe penetration position can be calculated to evaluate the positioning accuracy of the pineapple eye:

$$er = \sqrt{(W_2/2 - W_1 - 0.99)^2 + (H_2/2 - H_1 - 0.99)^2} \tag{9}$$

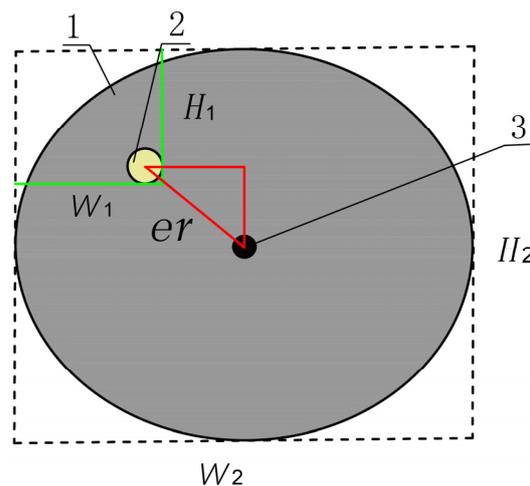


Figure 10. Measurement principle of the probe position error. 1. pineapple eye, 2. probe, and 3. pineapple eye center point.

In Equation (9), er is the error, and W_1 is the distance between the left edge of the pineapple eye and the right edge of the probe, in mm. W_2 is the maximum length of the pineapple eye in the horizontal direction, in mm. H_1 is the distance between the upper

edge of the pineapple eye and the lower edge of the probe, in mm. H_2 is the maximum length of the pineapple eye in the vertical direction, in mm. The probe radius is 0.99 mm.

Using five Goodfarmer Philippine pineapples, the diameter of the pineapple eye was 9–12 mm (manual measurement) after manual peeling. The positioning test is carried out on the built-in test platform. When the probe reaches each pineapple eye position, a Vernier caliper is used to successively measure the distances W_1 , W_2 , H_1 , and H_2 , as shown in Figure 11.

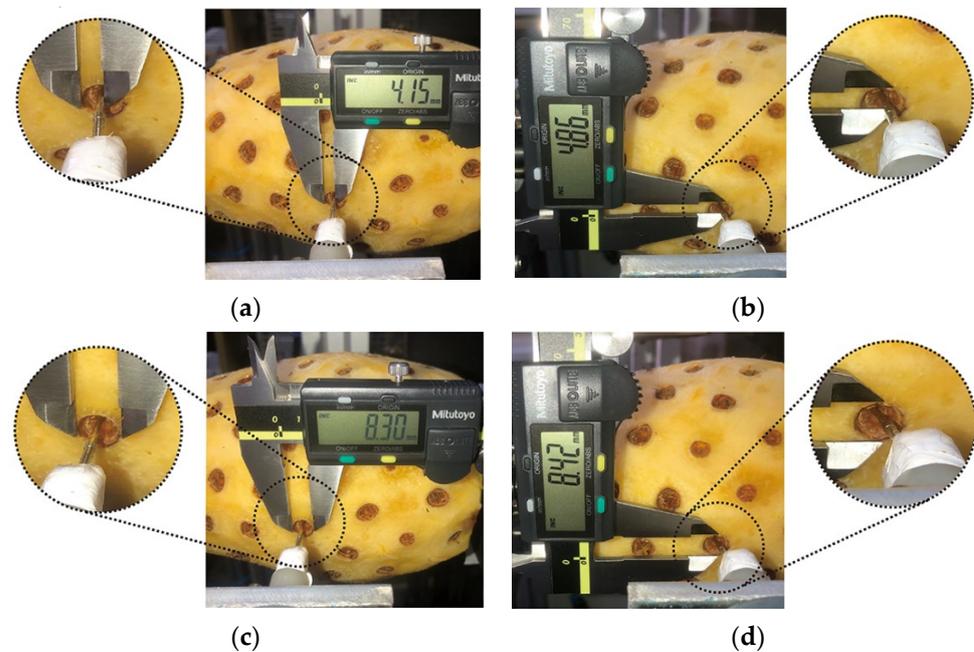


Figure 11. Measuring the pineapple eye error with a Vernier caliper. (a) W_1 measurement; (b) H_1 measurement; (c) W_2 measurement; (d) H_2 measurement.

3. Results and Discussion

3.1. YOLOv5 Model Performance Evaluation

To evaluate the detection effect of the pineapple eye recognition model, the model recognition accuracy and detection efficiency are mainly measured from four parameters: recall (R), precision (P), average accuracy (AP), and detection time of a single pineapple eye.

$$\begin{cases} P = \frac{TP}{TP+FP} \\ R = \frac{TP}{TP+FN} \\ AP = \int_0^1 PdR \end{cases} \quad (10)$$

The AP value in Formula (10) is the area between the P – R curve and the coordinate axis, TP represents the number of positive samples (pineapple eyes) correctly predicted as positive samples, TN denotes the number of negative samples correctly predicted as negative samples, FP indicates the number of negative samples predicted as positive samples, and FN suggests the number of positive samples predicted as negative samples.

The curve of network model training is shown in Figure 12. Figure 12a shows the loss function curve of training, with a minimum value of 0.01689. Figure 12b shows the accuracy P (precision) curve, and the maximum accuracy is 97.8%. Figure 12c shows the recall rate R (recall) curve, and the maximum recall rate is 97.5%. Figure 12d shows the mean average precision curve when the IOU threshold is set to 0.5.

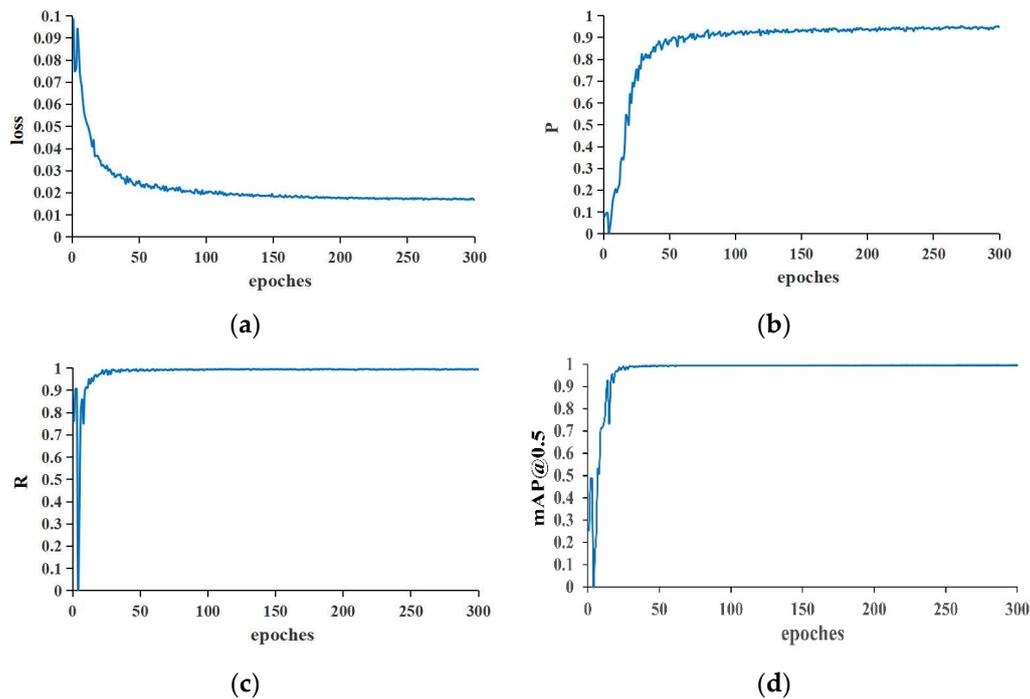


Figure 12. Model training results. (a) Value of loss varies with the number of iterations; (b) P vary with the number of iterations; (c) R vary with the number of iterations; (d) mAP@0.5 vary with the number of iterations.

The P – R curve is a graph that depicts the relationship between precision and recall. The abscissa represents R , while the ordinate represents P . The region contained in the P – R curve and the coordinate axis is AP . The larger the area between the curve and the coordinate axis is, the better the model recognition effect. Figure 13 shows the P – R curve with a threshold of 0.5 generated in the training process. Since there is only one recognition target in this paper, the AP is equal to the mAP (mean Average Precision). The mAP is 99.2%.

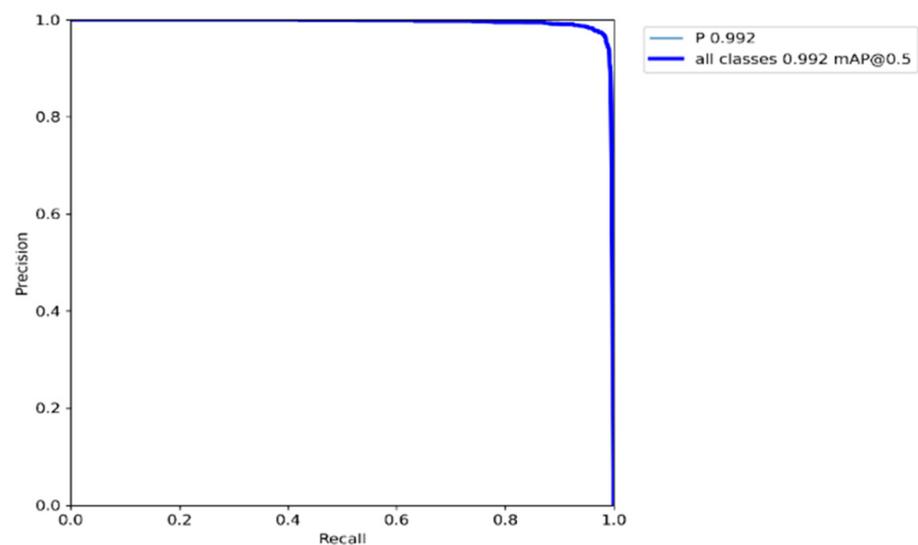


Figure 13. P – R curve.

To further verify the YOLOv5l model performance for pineapple eyes, the YOLOv5l network was compared with YOLOv5s, YOLOv5m, and YOLOv5x on 60 images in the test set; the target distribution of the test set was actually 1806 pineapple eyes. Then, the test

set images were input into the above models, respectively. The target recognition results of the pineapple eyes in the image samples of the test set by the model are shown in Table 1. The YOLOv5 (l, s, m, and x) values of mAP at a confidence of 0.5 were 98%, 97.6%, 97.8%, and 98%, respectively, showing the effectiveness of the proposed model. Additionally, the average times required to detect one pineapple eye image were 0.015 s, 0.012 s, 0.019 s, and 0.024 s, respectively. Figure 14 shows the YOLOv5l detection effect diagram with a confidence level greater than 0.5.

Table 1. Identification results for the pineapple eyes in test set.

Models	Precision (%)	Recall (%)	mAP (%)	Average Time(s)
YOLOv5l	98.0	96.6	98.0	0.015
YOLOv5s	98.3	96.2	97.6	0.012
YOLOv5m	97.9	96.3	97.8	0.019
YOLOv5x	98.1	96.5	98.0	0.024



Figure 14. YOLOv5l detection effect diagram.

Average time is the time to detect one pineapple eye image.

In order to further analyze the accuracy of the YOLOv5l model in pineapple eye image detection, the training results of YOLOv5l and the target detection model Mask R-CNN were compared with a threshold of 0.5, as shown in Table 2. As can be seen from Table 2, the mAP and detection speed of YOLOv5l are significantly higher than Mask R-CNN.

Table 2. Comparison models of YOLOv5l and Mask R-CNN.

Models	mAP (%)	Average Time (s)
YOLOv5l	99.2	0.015
Mask R-CNN	97.5	0.021

3.2. Result of Probe Positioning Test

The probe positioning test result, as shown in Figure 15, reveals that of the five Good-farmer Philippine pineapples after manual peeling (460 pineapple eyes in total, 444 pineapple eyes were successfully recognized), the deviation between the actual center of the pineapple eye and the probe puncture position was 1.01 mm, and the maximum was 2.17 mm, with a root mean square value of 1.09 mm.

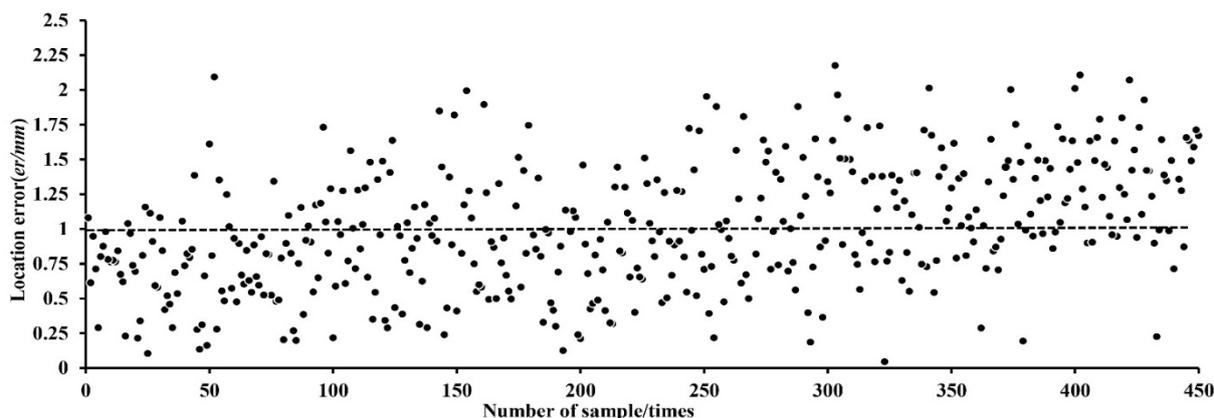


Figure 15. Probe positioning test.

3.3. Discussion

The YOLOv5 model has high detection accuracy on the self-built pineapple eye dataset. In the sample images of the whole test set, the accuracy, recall, and *AP* of the model are higher than 96%, indicating that the YOLOv5 recognition algorithm is feasible. The reason why a few pineapple eyes could not be successfully identified is that the pineapple eyes on both sides of the image are prone to distortion. This situation increases the recognition difficulty, resulting in some pineapple eye recognition errors. Therefore, further research on the optimization methods of models and parameters is needed to improve detection accuracy.

The localization experiment demonstrates that collecting images of the entire pineapple circumference at even intervals and employing multiangle image matching with high positioning precision may effectively accomplish three-dimensional localization of the pineapple eye. Simultaneously, pineapple eye coordinates have been converted into a form that can be directly applied by the actuator, which provides a good foundation for the further development of pineapple eye-removal equipment for practical operations.

4. Conclusions

A pineapple eye recognition algorithm was presented based on deep learning. YOLOv5 was used as the target detection network for pineapple eye recognition. The 600 pineapple eye images enhanced by the dataset are divided into a training set, validation set, and test set with an 8:1:1 ratio. The values in the final model validation of precision, recall, and mAP (mean average precision) were 97.8%, 97.5%, and 99.2%, respectively. The YOLOv5l network was compared with YOLOv5s, YOLOv5m, and YOLOv5x on 60 images in the test set. The YOLOv5 (l, s, m, and x) values of mAP were 98%, 97.6%, 97.8%, and 98%, showing the effectiveness of the proposed model. Additionally, the average times required to detect one pineapple eye image were 0.015 s, 0.012 s, 0.019 s, and 0.024 s. The detection results of YOLOv5l and Mask R-CNN were further compared, and the results showed that YOLOv5l was significantly higher than that of Mask R-CNN in both the mAP and detection speed.

A pineapple eye location algorithm based on monocular multiangle image stereo matching was proposed. Two images with different angles of 90° were selected as a group for stereo-matching analysis to obtain the three-dimensional position information of all pineapple eyes, establish a camera three-dimensional coordinate system with the camera optical center as the origin, and obtain the three-dimensional space coordinates (X, Y, Z) of the all pineapple eye through the geometric vector method. To facilitate subsequent experiments and the operation of removing pineapple eyes in practical engineering applications, in this paper, the three-dimensional space coordinate (X, Y, Z) of the pineapple eye was transformed into the space coordinate (L, θ) with the probe (or eye-removal tool) position L and the rotation angle θ of the pineapple as the reference. The probe test results showed that the average deviation between the actual center of the pineapple eye and the puncture position of the probe was

1.01 mm, the maximum was 2.17 mm, the root mean square value was 1.09 mm, and the positioning accuracy met the needs of the automated eye-removal operations.

The pineapple eye recognition and positioning algorithm proposed in this paper provides an important theoretical basis for the development of automatic pineapple-eye-removal equipment. The practical application performance of the algorithm needs to be verified and improved in the actual eye-removal operation. At the same time, only one variety of pineapple was tested, and the peeling operation was performed manually. The applicability of the algorithm to different varieties of pineapples and machine-peeled pineapples also needs to be further verified.

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

Funding: This research was funded by the Natural Science Foundation of Hunan Province of China, grant number 2021JJ30363.

Institutional Review Board Statement: Not applicable.

Data Availability Statement: All data are presented in this article in the form of figures and tables.

Acknowledgments: We gratefully acknowledge Mingliang Wu, Ying Xiong and anonymous referees for thoughtful review of this research as well as the assistance of Yanfei Li with statistical analyses.

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

References

- Jin, Y. Pineapple Market and Industry Investigation and Analysis Report. *Agric. Prod. Mark.* **2021**, *8*, 46–47.
- Gong, Y. Research on Strategies for Optimization and Upgrading of Pineapple Industry in Zhanjiang. Master's Thesis, Guangdong Ocean University, Zhanjiang, China, 2020.
- Jia, W.; Zhang, Z.; Shao, W.; Hou, S.; Ji, Z.; Liu, G.; Yin, X. FoveaMask: A Fast and Accurate Deep Learning Model for Green Fruit Instance Segmentation. *Comput. Electron. Agric.* **2021**, *191*, 106488. [[CrossRef](#)]
- Li, B.; Ning, W.; Wang, M.; Li, L. In-Field Pineapple Recognition Based on Monocular Vision. *Trans. Chin. Soc. Agric. Eng.* **2010**, *26*, 345–349. [[CrossRef](#)]
- Lin, G.; Zou, X. Citrus Segmentation for Automatic Harvester Combined with AdaBoost Classifier and Leung-Malik Filter Bank. *IFAC-Pap.* **2018**, *51*, 379–383. [[CrossRef](#)]
- Lv, J.; Wang, F.; Xu, L.; Ma, Z.; Yang, B. A Segmentation Method of Bagged Green Apple Image. *Sci. Hort.* **2019**, *246*, 411–417. [[CrossRef](#)]
- Kurtulmus, F.; Lee, W.S.; Vardar, A. Green Citrus Detection Using “Eigenfruit”, Color and Circular Gabor Texture Features under Natural Outdoor Conditions. *Comput. Electron. Agric.* **2011**, *78*, 140–149. [[CrossRef](#)]
- Wang, D.; He, D. Fusion of Mask RCNN and Attention Mechanism for Instance Segmentation of Apples under Complex Background. *Comput. Electron. Agric.* **2022**, *196*, 106864. [[CrossRef](#)]
- Kasinathan, T.; Singaraju, D.; Uyyala, S.R. Insect Classification and Detection in Field Crops Using Modern Machine Learning Techniques. *Inf. Process. Agric.* **2021**, *8*, 12. [[CrossRef](#)]
- Li, H.; Zhang, M.; Gao, Y. Green ripe tomato detection method based on machine vision in greenhouse. *Trans. Chin. Soc. Agric. Eng.* **2017**, *33*, 328–334+388.
- Wang, C.; Zou, X.; Tang, Y.; Luo, L.; Feng, W. Localisation of Litchi in an Unstructured Environment Using Binocular Stereo Vision. *Biosyst. Eng.* **2016**, *145*, 39–51. [[CrossRef](#)]
- Zhao, Y.; Gong, L.; Zhou, B.; Huang, Y.; Liu, C. Detecting Tomatoes in Greenhouse Scenes by Combining AdaBoost Classifier and Colour Analysis. *Biosyst. Eng.* **2016**, *148*, 127–137. [[CrossRef](#)]
- Altaheri, H.; Alsulaiman, M.; Muhammad, G. Date Fruit Classification for Robotic Harvesting in a Natural Environment Using Deep Learning. *IEEE Access* **2019**, *7*, 117115–117133. [[CrossRef](#)]
- Koirala, A.; Walsh, K.B.; Wang, Z.; McCarthy, C. Deep Learning—Method Overview and Review of Use for Fruit Detection and Yield Estimation. *Comput. Electron. Agric.* **2019**, *162*, 219–234. [[CrossRef](#)]
- Lv, J.; Xu, H.; Han, Y.; Lu, W.; Xu, L.; Rong, H.; Yang, B.; Zou, L.; Ma, Z. A Visual Identification Method for the Apple Growth Forms in the Orchard. *Comput. Electron. Agric.* **2022**, *197*, 106954. [[CrossRef](#)]

16. Zhang, X.; Gao, Q.; Pan, D. Picking recognition research of pineapple in complex field environment based on improved YOLOv3. *J. Chin. Agric. Mech.* **2021**, *42*, 201–206. [[CrossRef](#)]
17. Tian, Y.; Yang, G.; Wang, Z.; Wang, H.; Li, E.; Liang, Z. Apple Detection during Different Growth Stages in Orchards Using the Improved YOLO-V3 Model. *Comput. Electron. Agric.* **2019**, *157*, 417–426. [[CrossRef](#)]
18. Yu, Y.; Zhang, K.; Yang, L.; Zhang, D. Fruit Detection for Strawberry Harvesting Robot in Non-Structural Environment Based on Mask-RCNN. *Comput. Electron. Agric.* **2019**, *163*, 104846. [[CrossRef](#)]
19. Zhang, C.; Ding, H.; Shi, Q.; Wang, Y. Grape Cluster Real-Time Detection in Complex Natural Scenes Based on YOLOv5s Deep Learning Network. *Agriculture* **2022**, *12*, 1242. [[CrossRef](#)]
20. Ji, W.; Meng, X.; Qian, Z.; Xu, B.; Zhao, D. Branch Localization Method Based on the Skeleton Feature Extraction and Stereo Matching for Apple Harvesting Robot. *Int. J. Adv. Robot. Syst.* **2017**, *14*, 1729881417705276. [[CrossRef](#)]
21. Rong, X.; Jiang, H.; Ying, Y. Recognition of Clustered Tomatoes Based on Binocular Stereo Vision. *Comput. Electron. Agric.* **2014**, *106*, 75–90. [[CrossRef](#)]
22. Wang, C.; Tang, Y.; Zou, X.; Luo, L.; Chen, X. Recognition and Matching of Clustered Mature Litchi Fruits Using Binocular Charge-Coupled Device (CCD) Color Cameras. *Sensors* **2017**, *17*, 2564. [[CrossRef](#)]
23. Ge, L.; Yang, Z.; Sun, Z.; Zhang, G.; Zhang, M.; Zhang, K.; Zhang, C.; Tan, Y.; Li, W. A Method for Broccoli Seedling Recognition in Natural Environment Based on Binocular Stereo Vision and Gaussian Mixture Model. *Sensors* **2019**, *19*, 1132. [[CrossRef](#)]
24. Luo, W.; Schwing, A.G.; Urtasun, R. Efficient Deep Learning for Stereo Matching. In Proceedings of the IEEE Conference on Computer Vision & Pattern Recognition, Las Vegas, NV, USA, 27–30 June 2016; pp. 5695–5703.
25. Guo, H.; Liu, Y.F.; Wang, Y.; Shen, X. Calibration of Binocular Vision Measurement of Large Gear Workpiece Welding. *J. Donghua Univ. Sci.* **2013**, *4*, 455–459.
26. Zhang, B.; Huang, W.; Wang, C.; Gong, L.; Zhao, C.; Liu, C.; Huang, D. Computer Vision Recognition of Stem and Calyx in Apples Using Near-Infrared Linear-Array Structured Light and 3D Reconstruction. *Biosyst. Eng.* **2015**, *139*, 25–34. [[CrossRef](#)]
27. Hongsheng, S.; Zhenwei, W.; Hong, C. Three-Dimensional Reconstruction of Complex Spatial Surface Based on Line Structured Light. In Proceedings of the IECON 2021—47th Annual Conference of the IEEE Industrial Electronics Society, Toronto, ON, Canada, 13–16 October 2021; pp. 1–5.
28. Chen, C.; Tian, Y.; Lin, L.; Chen, S.; Li, H.; Wang, Y.; Su, K. Obtaining World Coordinate Information of UAV in GNSS Denied Environments. *Sensors* **2020**, *20*, 2241. [[CrossRef](#)]
29. Zhao, D.-A.; Lv, J.; Ji, W.; Zhang, Y.; Chen, Y. Design and Control of an Apple Harvesting Robot. *Biosyst. Eng.* **2011**, *110*, 112–122. [[CrossRef](#)]
30. Tsai, R. A Versatile Camera Calibration Technique for High-Accuracy 3D Machine Vision Metrology Using off-the-Shelf TV Cameras and Lenses. *IEEE J. Robot. Autom.* **1987**, *3*, 323–344. [[CrossRef](#)]
31. Olenskyj, A.G.; Sams, B.S.; Fei, Z.; Singh, V.; Raja, P.V.; Bornhorst, G.M.; Earles, J.M. End-to-End Deep Learning for Directly Estimating Grape Yield from Ground-Based Imagery. *Comput. Electron. Agric.* **2022**, *198*, 107081. [[CrossRef](#)]
32. Wu, T.-H.; Wang, T.-W.; Liu, Y.-Q. Real-Time Vehicle and Distance Detection Based on Improved Yolo v5 Network. In Proceedings of the 2021 3rd World Symposium on Artificial Intelligence (WSAI), Guangzhou, China, 18–20 June 2021; pp. 24–28.
33. Zhou, X.; Wei, G.; Fu, W.L.; Du, F. Application of Deep Learning in Object Detection. In Proceedings of the 2017 IEEE/ACIS 16th International Conference on Computer and Information Science (ICIS), Wuhan, China, 24–26 May 2017.
34. Xu, X.; Zhang, X.; Zhang, T. Lite-YOLOv5: A Lightweight Deep Learning Detector for On-Board Ship Detection in Large-Scene Sentinel-1 SAR Images. *Remote Sens.* **2022**, *14*, 1018. [[CrossRef](#)]