fractal and fractional

MDPI

Article

RGB-D Camera and Fractal-Geometry-Based Maximum Diameter Estimation Method of Apples for Robot Intelligent Selective Graded Harvesting

Bin Yan [1,2,3,*] and Xiameng Li [4]

[1] College of Automation and Information Engineering, Xi'an University of Technology, Xi'an 710048, China
[2] College of Mechanical and Electronic Engineering, Northwest A&F University, Xianyang 712100, China
[3] Shaanxi Key Laboratory of Complex System Control and Intelligent Information Processing, Xi'an 710048, China
[4] Faculty of Liberal Arts, Northwest University, Xi'an 710127, China
* Correspondence: yanbin@nwafu.edu.cn

Abstract: Realizing the integration of intelligent fruit picking and grading for apple harvesting robots is an inevitable requirement for the future development of smart agriculture and precision agriculture. Therefore, an apple maximum diameter estimation model based on RGB-D camera fusion depth information was proposed in the study. Firstly, the maximum diameter parameters of Red Fuji apples were collected, and the results were statistically analyzed. Then, based on the Intel RealSense D435 RGB-D depth camera and LabelImg software, the depth information of apples and the two-dimensional size information of fruit images were obtained. Furthermore, the relationship between fruit depth information, two-dimensional size information of fruit images, and the maximum diameter of apples was explored. Based on Origin software, multiple regression analysis and nonlinear surface fitting were used to analyze the correlation between fruit depth, diagonal length of fruit bounding rectangle, and maximum diameter. A model for estimating the maximum diameter of apples was constructed. Finally, the constructed maximum diameter estimation model was experimentally validated and evaluated for imitation apples in the laboratory and fruits on the Red Fuji fruit trees in modern apple orchards. The experimental results showed that the average maximum relative error of the constructed model in the laboratory imitation apple validation set was ±4.1%, the correlation coefficient (R^2) of the estimated model was 0.98613, and the root mean square error (RMSE) was 3.21 mm. The average maximum diameter estimation relative error on the modern orchard Red Fuji apple validation set was ±3.77%, the correlation coefficient (R^2) of the estimation model was 0.84, and the root mean square error (RMSE) was 3.95 mm. The proposed model can provide theoretical basis and technical support for the selective apple-picking operation of intelligent robots based on apple size grading.

Keywords: fractal geometry; harvesting robot; apple; nonlinear surface fitting; multivariate regression analysis

Citation: Yan, B.; Li, X. RGB-D Camera and Fractal-Geometry-Based Maximum Diameter Estimation Method of Apples for Robot Intelligent Selective Graded Harvesting. *Fractal Fract.* **2024**, *8*, 649. https://doi.org/10.3390/fractalfract8110649

Academic Editors: Jinxi Zhang, Xuefeng Zhang, Driss Boutat and Dayan Liu

Received: 26 September 2024
Revised: 1 November 2024
Accepted: 5 November 2024
Published: 7 November 2024

1. Introduction

Nowadays, apple picking still mainly relies on manpower [1–5], which not only requires high labor and time requirements but also poses significant safety risks when picking apples from high places. With the rapid development of machine vision, robotics technology, and artificial intelligence technology [6–15], apple-picking robots are gradually becoming a new direction to replace traditional manual picking in order to reduce labor costs [16–18]. Therefore, exploring and optimizing the key technologies of apple-picking robots in depth to achieve efficient and automated picking operations is of great significance for ensuring efficient apple harvesting and promoting the intelligent development of agricultural production. Among them, intelligent perception of fruit information on apple

trees [19–23] is one of the core technologies of harvesting robots. On the other hand, the current apple-picking operation and the fruit-grading operation based on apple size after harvesting are based on two disjointed processes. In order to improve the efficiency of apple harvesting grading, the picking robot can intelligently detect the maximum diameter of apples on the fruit tree online and then selectively and accurately pick the fruits based on the diameter size, in order to achieve the integration of "intelligent apple picking and grading". This is an inevitable requirement for the future development trend of smart agriculture and precision agriculture, with important strategic significance and practical value.

The financial benefits of robot intelligent selective graded harvesting for apple fruits contains the following: (1) Improving picking and grading efficiency: The apple-picking robot adopts intelligent operation, unmanned driving, autonomous pathfinding, and autonomous path planning, which can operate continuously and greatly improve picking and grading efficiency. On the other hand, robots can accurately identify, pick, and grade apples based on their size characteristics, avoiding waste and errors in manual picking and grading. (2) Reducing labor costs: Traditional apple-picking and grading methods require a large amount of human resources, while robot picking and grading technology can replace some manual labor and reduce the labor intensity of farmers. On the other hand, robot harvesting and grading can reduce losses caused by human factors and ensure the quality of fruits. In addition, robot picking avoids the damage that may be caused to apples due to improper operation during manual picking and grading, improving the quality and commodity rate of fruits. (3) The adoption of advanced robot picking and grading technology can enhance the modernization level and market competitiveness of orchards, bringing more business opportunities and development space for orchards. These helps orchards achieve sustainable development, bring long-term economic returns, and increase market share in apple sales.

Up until now, scholars have conducted research on apple size estimation [24–29], but most of the research is focused on assembly line operations such as fruit quality detection, grading, and packaging under structured and environmentally controlled conditions. The growth environment of fruits on apple trees is unstructured, and the distance and angle between the camera and the fruits on the tree vary. Therefore, the above algorithm cannot directly estimate the size of fruits on apple trees.

On the other hand, Zeze Fan et al. [30] obtained the mapping relationship between the width and height values of the apple detection box and the actual physical size of the fruit based on size measurement, conversion models, and parameters of the Sony ICLE-6000 camera. Based on this mapping relationship, they used artificial intelligence algorithms to obtain the fruit detection box and obtain the diameter prediction information of the apple. However, the experimental verification of the apple diameter estimation model is based on setting the angle between the camera and the apple target to 0° and keeping the distance between the camera and the apple fixed at 3 m. In reality, the RGB-D depth camera installed on the apple-picking robot obtains a wider field of view, and the apples are not only distributed in the central area of the camera's field of view. Therefore, the angle between the fruit and the camera will not be a fixed value of 0°, and due to the presence of the apple crown, the depth distance between each apple and the camera will vary and is not fixed. Therefore, it is not possible to directly apply this model to picking robots for real-time estimation of apple diameter on fruit trees in orchard environments. Shenglian Lu et al. [31] used two apple image capture modes and combined the YOLOv4 convolutional neural network model [32] to detect the position of fruits in two-dimensional images. They estimated the length and diameter of apples on fruit trees using artificial apple reference points. However, due to the fact that this method requires the use of a reference object (an artificial apple model with a size of 90 mm × 80 mm) placed in the camera's field of view to estimate the actual length and diameter of the fruit based on the size of the imitation apple, it is not suitable for robots that need to perform dynamic picking operations in apple orchards.

The research content of this article is related to intelligent agricultural robots; therefore, it is relevant to the research on multi-agent and non-linear systems in this context. Reference [33] proposed a performance-based neural network consensus control method for time-varying non-linear MAS with strict feedback to ensure that the synchronization error quickly converges to a tight error set when all system states are subject to full state constraints. A self-adaptive PI event-triggered control method was proposed in [34] for a class of multiple-input multiple-output (MIMO) nonlinear systems with uncertain input delay. Through the proposed method, MIMO nonlinear systems can dynamically handle input delays of different durations while still maintaining excellent tracking performance. The estimation of the maximum diameter of fruits for apple-picking robots requires precise measurement of the maximum diameter of fruits in an unstructured orchard environment, that is, under constrained conditions (various natural environments in the orchard), with a small margin of error.

Up until now, there have been no reports on the research of real-time estimation of apple diameter size on fruit trees in actual orchard operation scenarios providing technical support for robot "intelligent apple picking and grading fusion". Therefore, this study utilized the Intel RealSense D435 RGB-D depth camera and LabelImg software, based on multiple regression analysis and nonlinear surface fitting methods, to explore the relationship between fruit depth information, two-dimensional size information of fruit images, and the maximum diameter of apples. Furthermore, the correlation between fruit depth, diagonal length of fruit bounding rectangle, and maximum diameter was analyzed. Finally, a model for estimating the maximum diameter of apples based on RGB-D camera fused depth information was proposed. The proposed model can provide theoretical basis and technical support for the selective apple-picking operation of intelligent robots based on apple size grading.

2. Acquisition and Analysis of Maximum Diameter Data for Apples

2.1. Measurement Method for Maximum Diameter of Fruit

According to the differences in fruit tree planting layout, apple orchards are divided into two categories: traditional and modern standardized (or standard). The traditional apple orchard has significant characteristics, with staggered branches between fruit trees, as well as narrow and closed space between rows, which restricts the smooth operation of mechanical work in the orchard. For modern standardized apple orchards (as shown in Figure 1), the dwarf rootstock dense planting cultivation mode dominates. This mode, with its low crown, convenient management, low labor demand, superior ventilation and light transmission performance, uniform fruit coloring, superior quality, and ease of intelligent agricultural machinery operation, has become an effective way to achieve orchard standardization, scale, and intelligent management. This model is widely used in advanced countries for apple production worldwide and is also a core development direction in the modernization process of the apple industry. Among them, compared with other cultivation modes such as "V-shaped", "Y-shaped", and "wall-shaped", the spindle-shaped cultivation mode, as the mainstream type of dwarf rootstock dense planting mode, is highly favored due to its multiple fruiting branches and strong fruit-bearing capacity, and it has become the mainstream choice in modern standard orchard construction. Therefore, the spindle-shaped apple trees in the standardized orchard at the Apple Experimental Station of Northwest A&F University in Baishui County of Shaanxi Province were used as the research object in the study.

Figure 1. Spindle-shaped apple trees in a modern orchard.

The LF01 digital vernier caliper (measurement accuracy: 0.01 mm) produced by Quzhou Gangtuo Tools Co., Ltd. (Quzhou, China), was used to measure and record the true maximum diameter of the fruit on the spindle-shaped apple tree in the modern orchard located at the Baishui Apple Test Station of Northwest A&F University. The measurement process is shown in Figure 2. Among them, multiple measurements of the maximum equatorial circle diameter were taken for each fruit, and the measured maximum cross-sectional diameter data were recorded as the maximum cross-sectional diameter value of the apple.

Figure 2. Measuring the maximum diameter of the apple.

2.2. Measurement Results and Statistical Analysis of Maximum Diameter of Apples

Based on the maximum diameter value of Red Fuji apples measured using a vernier caliper in the previous section, we drew a distribution chart of the data, as shown in Figure 3. As shown in the figure below, the distribution range of the maximum diameter of apples was 50–100 mm, with the concentrated distribution range of the maximum diameter of fruits being 70–100 mm.

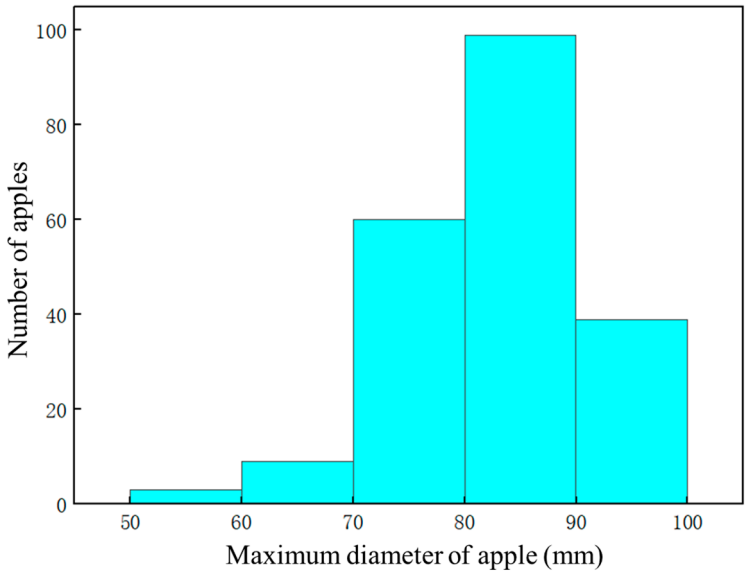


Figure 3. Maximum diameter distribution of fruit on apple trees in a modern orchard.

3. Calculation of Apple Depth Information and Two-Dimensional Image Size

3.1. Apple Depth Information Acquisition Based on RGB-D Camera

The Intel Realsense D435 camera (Intel Corporation, Santa Clara, CA, USA) imaging belongs to structured light-based 3D vision, and its depth measurement is based on the principle of optical triangulation, similar to laser detection. The difference is that the projection light (structured light) used in structured light 3D vision technology is an encoded light source, and the near-infrared signal emitter projects light with certain structural characteristics onto the surface of the object to be tested. The near-infrared images are received by the left and right near-infrared cameras, which can achieve real-time acquisition of depth images of the entire visual range scene and thus obtain information such as brightness, distance mapping, and 3D coordinates of the target object within the visual range. The module composition of the D435 camera and the assembly of the camera on the apple-picking robot are shown in Figure 4, and the schematic diagram of the structured light 3D vision principle is shown in Figure 5.

Figure 4. Module composition of the Intel Realsense D435 camera and its assembly on the apple-picking robot.

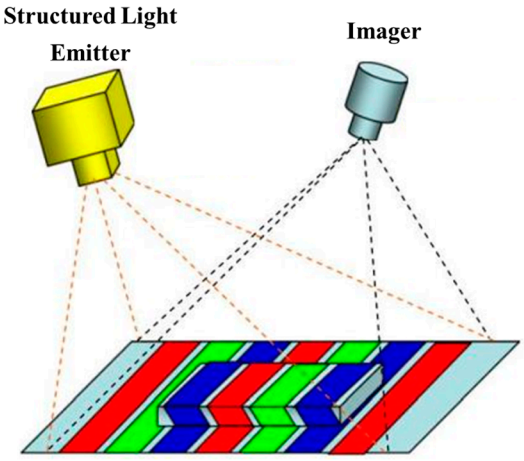


Figure 5. Schematic diagram of structured light 3D vision principle.

In the modern standard apple orchard of the Baishui Apple Experimental Station at Northwest A&F University, an apple two-dimensional image and depth data acquisition experimental platform was built. The main equipment includes a laptop (Lenovo Legion Y7000P, Intel (R) Core (TM) i7-9750H CPU, 2.6 GHz, 16 GB internal storage), Intel Realsense D435 depth camera, digital vernier caliper, etc. For different apple trees, fruit images and depth data were collected at different distances from the apple trees. The experimental scenario is shown in Figure 6.

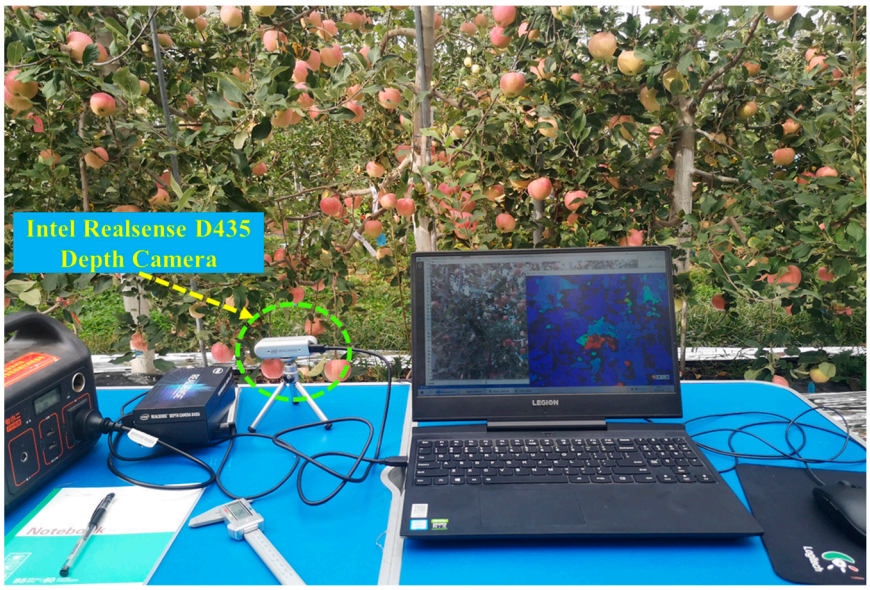


Figure 6. Acquisition of two-dimensional image and depth information of apples.

Before capturing the image and recording data, we measured the light intensity of the environment at that time utilizing the digital luminometer (Model: TES-1332A), and the light intensities of high light and low light environments were 1.119×10^5 lux and 1.85×10^4 lux, respectively.

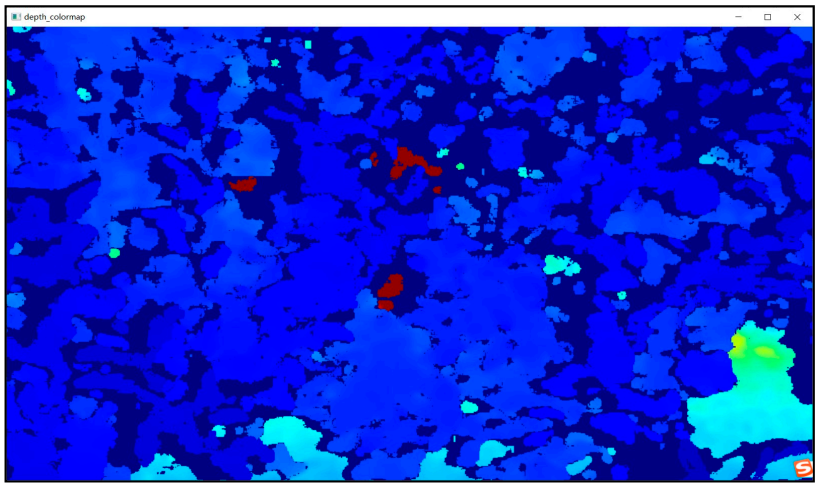
The acquisition of an apple's depth information is mainly based on the Intel Realsense D435 depth camera. Due to the inconsistency in size between the near-infrared images captured by the depth camera and the RGB color images, the first step is to match and align the RGB images captured by D435 with the near-infrared images. This was implemented using the Python (version 3.8) programming language, and the main steps are as follows:

(1) Configure the pipeline for streaming transmission to obtain the depth scale of the depth sensor.

(2) Create alignment objects and set the flow type for aligning depth frames, then perform the alignment operation between depth frames and color frames.
(3) Obtain the internal parameters of the aligned frames and color frames, and after verifying that the two frames are valid, obtain the aligned and matched color image frames and depth image frames.

After matching and aligning the RGB images obtained by the depth camera with the near-infrared images, RGB and depth pseudo-color images of the fruit were collected at different positions from the apple tree. At the same time, the depth information of all pixels in the image was automatically extracted and stored in an Excel (version 2016) spreadsheet. The collected RGB color images of apples and their corresponding depth pseudo-color images are shown in Figure 7a and Figure 7b, respectively (The symbol in the lower right corner of Figure 7 is the symbol of the input method software when taking a screenshot of the computer).

(**a**)



(**b**)

Figure 7. Apple RGB image (**a**) and corresponding apple depth pseudo-color image (**b**) collected.

The pseudocode that computes the depth information is shown as follows:

```
//Initialize depth camera
initialize_depth_camera()
//Configure camera parameters such as resolution, frame rate, etc.
set_camera_configuration(
stream_type: DEPTH,
resolution_width: 640,
resolution_height: 480,
frame_rate: 30
)
//Activate camera capture
start_camera_capture()
//Define a function to obtain depth information of a certain location
function get_depth_at_position(x, y):
//Waiting for the camera to capture the next frame
frame = wait_for_next_frame()
//Check if the frame has been successfully obtained
if frame is null:
raise_error("Failed to retrieve frame")
//Extract depth data from frames
depth_data = extract_depth_data(frame)
//Check if the depth data is valid
if depth_data is invalid:
raise_error("Invalid depth data")
//Get the depth value of the specified location
depth_value = depth_data[y][x]
//Convert depth values from one unit to another as needed
depth_value_in_meters = depth_value/1000
return depth_value_in_meters
//Call a function to obtain depth information for a specific location
try:
x_position = 320
y_position = 240
depth_at_specific_position = get_depth_at_position(x_position, y_position)
except error as e:
print("Error retrieving depth information: " + e)
```

LabelImg (Version 1.8.1) is a visual image calibration tool software. In this study, the tool was used to read the collected apple images, and the cursor was manually moved to the centroid position of the target apple. LabelImg can automatically display the two-dimensional coordinate values of the pixel point in real time on the interface, as shown in Figure 8. The non-English term in Figure 8 are belonged to the storage path of the image in the computer.

Figure 8. Obtaining the two-dimensional centroid position of an apple based on LabelImg software.

After reading the center coordinates of the two-dimensional image of the target apple, the depth values at the corresponding coordinate points were obtained from the Excel file that stored the depth data of all pixel points in the image. The depth value distribution of all experimental apples was statistically analyzed as shown in Figure 9. It can be seen from the figure that the depth value distribution range of the obtained apples was 40–240 cm, with the main depth value distribution in the range of 80–200 cm.

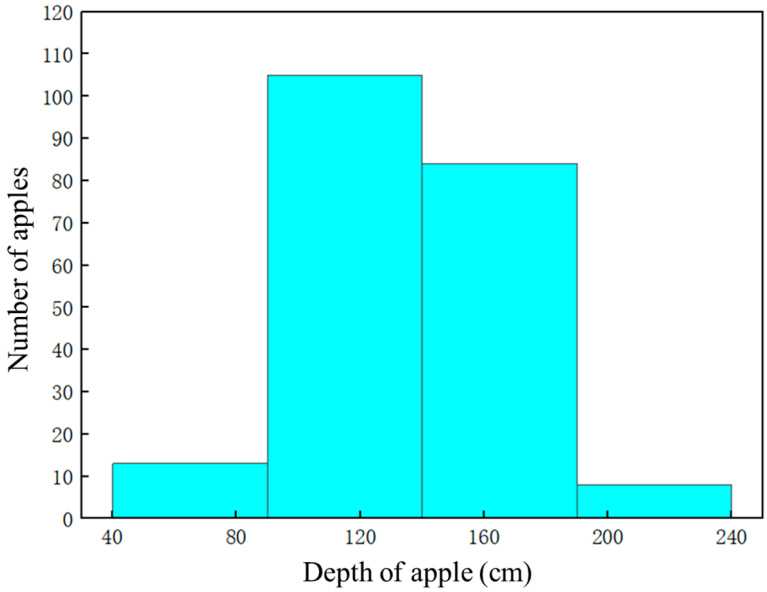


Figure 9. Depth distribution of apples.

3.2. Calculation of Two-Dimensional Size of Apples in Images

The LabelImg tool was utilized to read the collected 2D RGB image of the apple, and the software's built-in manual annotation function was used to draw the minimum bounding rectangle of the target apple. The software can automatically display the length and width dimensions of the rectangle (as shown in Figure 10). The non-English term in Figure 10 are belonged to the storage path of the image in the computer.

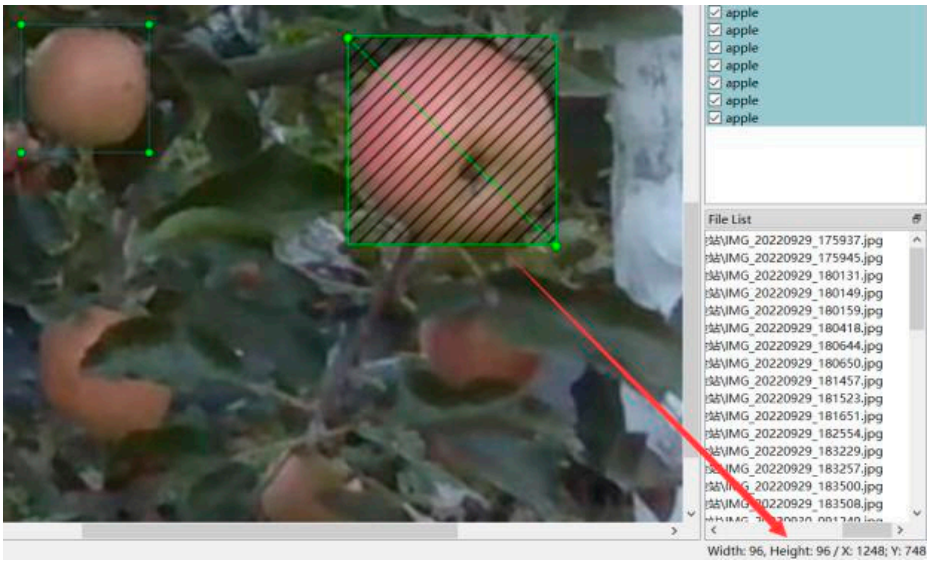


Figure 10. Obtaining the size of the circumscribed rectangle of an apple based on LabelImg software.

Due to the close positive correlation between the diagonal length of the bounding rectangle of an apple image and its length and width dimensions, the diagonal length of

the bounding rectangle of an apple image (in pixels) was used to characterize the two-dimensional size of the fruit image. The diagonal length distribution of the bounding rectangle of apples collected in the orchard experiment is shown in Figure 11. It can be seen from the figure that the distribution range of the length of the bounding rectangle of apples was 40–160 pixels (px).

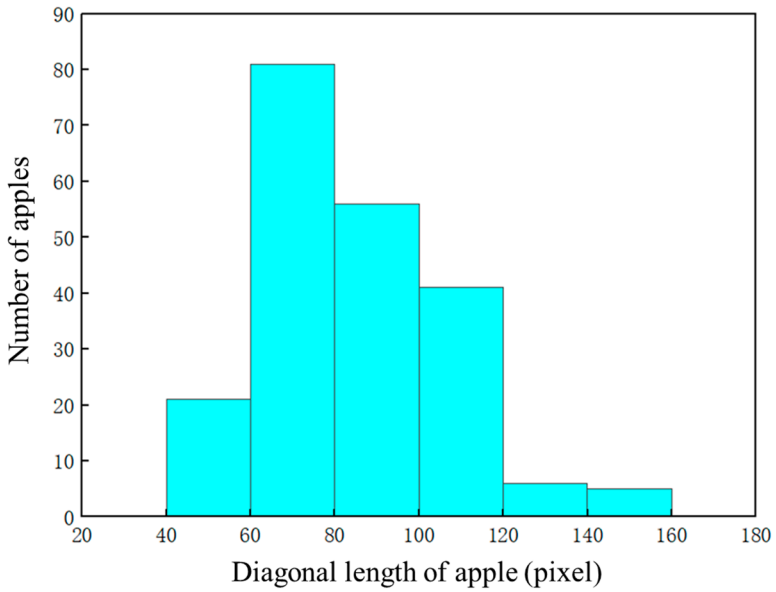


Figure 11. Diagonal length distribution of apple circumscribed rectangle.

4. Construction of the Maximum Diameter Estimation Model for Apples

4.1. Multivariate Regression Analysis of the Estimation Model

Based on the Intel Realsense D435 RGB-D depth camera, this study explored the relationship between fruit depth information, two-dimensional size information, and maximum diameter on apple trees. Furthermore, a fruit maximum diameter estimation model for selective apple picking by picking robots was proposed. Through multiple regression analysis and nonlinear surface fitting methods, the correlation between fruit depth, two-dimensional size, and maximum diameter was analyzed, and a multiple regression model containing the above parameter variables was established.

Binary quadratic polynomials can be used to fit three-dimensional surface models, and the normalized mathematical expression is as follows:

$$z = c_{20}x^2 + c_{10}x + c_{02}y^2 + c_{01}y + c_{11}xy + c_{00} \quad (1)$$

In the formula, x, y, and z represent data obtained through actual measurement, and c_{20}, c_{10}, c_{02}, c_{01}, c_{11}, and c_{00} are the undetermined coefficients of the three-dimensional surface model expression. The expression for the sum of squared errors of this surface model is as follows:

$$Q(c_{20},\ c_{10},\ c_{02},\ c_{01},\ c_{11},\ c_{00}) = \sum_{i=1}^{n}\left[z_i - \left(c_{20}{x_i}^2 + c_{10}x_i + c_{02}{y_i}^2 + c_{01}y_i + c_{11}x_iy_i + c_{00}\right)\right]^2 \quad (2)$$

In the equation, (x_i, y_i, z_i) represents the measured and sampled data, where $i = 1, 2, \ldots, n$.

The principle for solving the undetermined coefficient terms in the above binary quadratic polynomial is to minimize the sum of squared errors of the three-dimensional surface model. According to the knowledge of calculus, this problem is transformed into solving the extremum of a six variable function $Q(c_{20}, c_{10}, c_{02}, c_{01}, c_{11}, c_{00})$, and solving the

following equations (Equation (3)) simultaneously can obtain the values of the coefficients of each undetermined term of the fitted three-dimensional surface model.

$$\begin{array}{ccc} \frac{\partial Q}{\partial c_{20}} = 0 & \frac{\partial Q}{\partial c_{02}} = 0 & \frac{\partial Q}{\partial c_{11}} = 0 \\ \frac{\partial Q}{\partial c_{10}} = 0 & \frac{\partial Q}{\partial c_{01}} = 0 & \frac{\partial Q}{\partial c_{00}} = 0 \end{array} \tag{3}$$

Based on the fitting theory of the three-dimensional surface model mentioned above, non-linear surface fitting was used to analyze the correlation between the dependent variable (maximum diameter of the apple) and the independent variables (fruit depth information, two-dimensional size information). The dependent variables of the maximum diameter estimation model for apples were fitted to the general form of a quadratic polynomial model, and the equation was expressed as follows:

$$z = ax^2 + bx + cy^2 + dy + exy + f \tag{4}$$

In the formula, z is the estimated maximum diameter of the apple (dependent variable, unit: mm), y is the depth information of the fruit (unit: cm), x is the minimum diagonal length of the bounding rectangle of the apple (unit: pixel), f is the intercept, b and d are linear coefficients, a and c are quadratic coefficients, and e is the interaction coefficient.

4.2. Nonlinear Surface Fitting

The three-dimensional spatial distribution map of the maximum diameter, diagonal length of the bounding rectangle, and depth parameters of Red Fuji apples collected from the standard orchard at the Baishui Apple Experimental Station of Northwest A&F University for constructing the maximum diameter estimation model is shown in Figure 12.

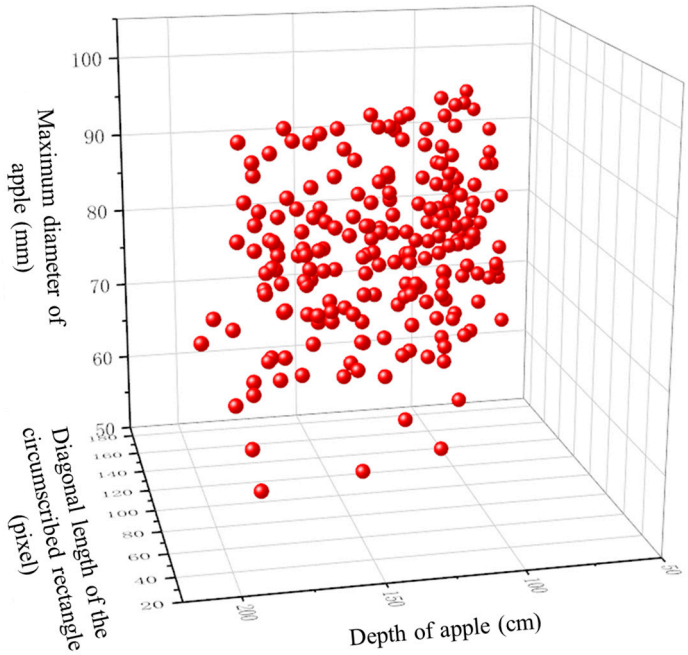


Figure 12. Three-dimensional distribution of the parameters of the maximum diameter, diagonal length of the circumscribed rectangle, and the depth of apples.

Origin is a scientific drawing and data analysis software developed by OriginLab corporation that runs on the Microsoft Windows operating system. Origin supports the drawing of various 2D/3D graphics, including powerful data analysis functions such as statistics, signal processing, curve fitting, and peak analysis. This study conducted multiple regression nonlinear surface fitting analysis on the measured maximum diameter, fruit depth information, and minimum external rectangle diagonal length parameter data of

apples using Origin (version 2021) software. The nonlinear surface fitting function setting interface in Origin software is shown in Figure 13. The non-English term and symbol * in the software interface in Figure 13 are prompts for parameter settings, function selection, and other operation buttons for the non-linear surface fitting function.

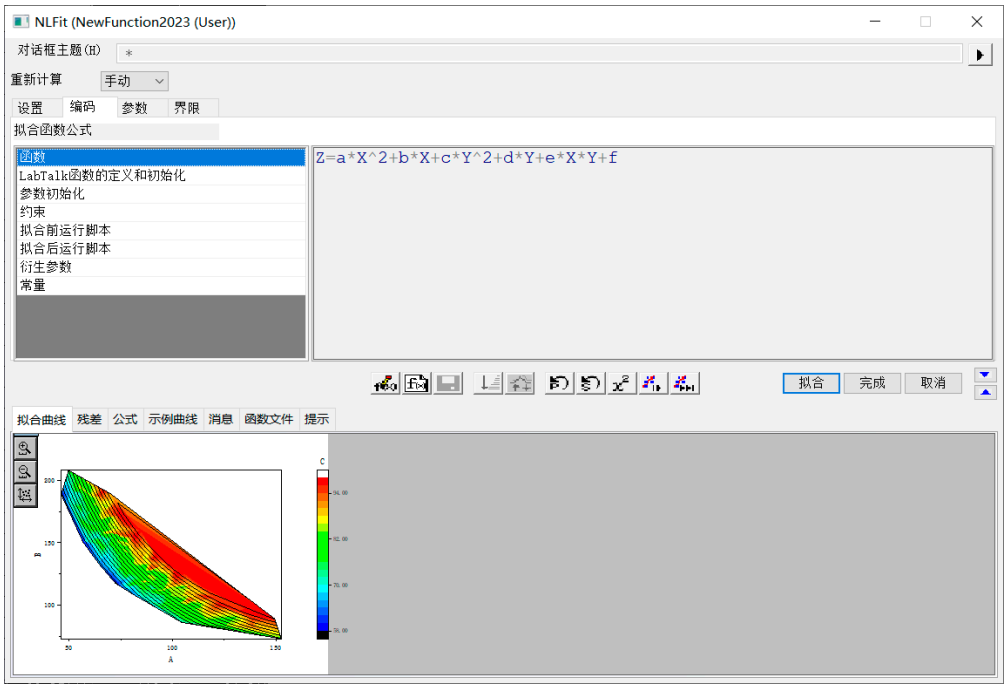


Figure 13. Origin software-based model nonlinear surface fitting setting interface.

The least squares method is a widely used nonlinear surface fitting algorithm in practical applications that determines the optimal fitting surface by minimizing the sum of squared distances between data points and the fitting surface. In this study, the fitting of the estimation model was performed using the Levenberg–Marquardt algorithm (LMA) and nonlinear least squares method for nonlinear 3D surface fitting. The multiple regression based on the Origin2021 execution estimation model resulted in a final fitting state of 'successful'. The coefficient of determination (R^2) reflects the proportion of the total variation of the dependent variable that can be explained by the independent variable through the regression relationship. The higher the value, the better the model. The key fitting parameter results of the constructed maximum transverse diameter estimation model for apples are shown in Table 1. It can be seen that R^2 was 0.80889, indicating that the model had good fitting correlation, and the correlation between variables was high. The model fitting result was good.

Table 1. Fitting results of the estimate model.

Estimating Model Fitting Parameters/Indicators	Parameter Values
Equation	$z = ax^2 + bx + cy^2 + dy + exy + f$
Reduced Chi-Sqr	13.49046
R^2 (COD)	0.80889
Adjusted R^2	0.80421
Freedom	204
Fitting state	Success (100)

The evaluation analysis of the coefficient terms of the multiple regression equation for the maximum diameter estimation model of fruits obtained by nonlinear surface fitting is

shown in Table 2. As shown in the table below, the correlation between the coefficients of the estimated model was above 95%, indicating a good fitting effect of the model.

Table 2. Evaluation and analysis of coefficient terms of the multiple regression equation.

Estimate Model Coefficients	Parameter Values	Standard Deviation	t Value	Probability > |t|	Correlation
a	−0.00372	0.00184	−2.02229	0.04445	0.99975
b	1.22207	0.67377	1.8138	0.07118	0.99998
c	−0.00112	9.95579×10^{-4}	−1.12579	0.26158	0.99986
d	0.62816	0.49219	1.27627	0.20331	0.99999
e	0.00241	0.00273	0.88122	0.37924	0.99993
f	−82.58015	61.00251	−1.35372	0.17732	0.99998

The results of the multiple regression analysis of variance for the estimation model are shown in Table 3 below:

Table 3. Variance analysis of multiple regression of the estimate model.

Analysis Indicators	DF	Sum of Squares	Mean Square	F Value	Probability > F
Regression	5	11,648.65712	2329.73142	172.69474	2.82079×10^{-71}
Residual	204	2752.05373	13.49046	-	-
Uncorrected overall	210	1,448,389.83	-	-	-
Corrected overall	209	14,400.71085	-	-	-

The expression of the maximum diameter estimation model for apples based on multiple regression and nonlinear surface fitting is as follows:

$$z = -0.00372x^2 + 1.22207x - 0.00112y^2 + 0.62816x + 0.00241xy - 82.58015 \quad (5)$$

Among them, z is the estimated maximum diameter of the apple, y is the depth of the fruit, and x is the minimum diagonal length of the bounding rectangle of the apple.

Based on the actual maximum diameter, depth, and diagonal length of the bounding rectangle of modern orchard apples, the nonlinear surface fitting results of the apple maximum diameter estimation model is shown in the 3D figure below (Figure 14). Among them, the red small spheres represent each apple data point, and it can be clearly seen that the fitting surface covered the red small sphere data points well, indicating a good fitting effect.

On the other hand, the Intel Realsense D435 depth camera was utilized to collect apple image data in the study. The D435 camera is equipped with the OV2740 RGB sensor, which has a pixel size (the actual size represented by a single pixel in the length and width directions) of 1.4 μm × 1.4 μm. Due to the small differences in the size of individual pixels among different depth cameras, the maximum diameter estimation model for fruits needs to be applied to other depth cameras by multiplying the diagonal length of the minimum bounding rectangle of the obtained apple by a scaling factor k and then substituting it into the expression of the estimation model. Based on the pixel size of the D435 camera, the mathematical expression for the scaling factor k is shown in Equation (6), where k_0 is the pixel size of the camera used, measured in micrometers (μm).

$$k = \frac{k_0}{1.4} \quad (6)$$

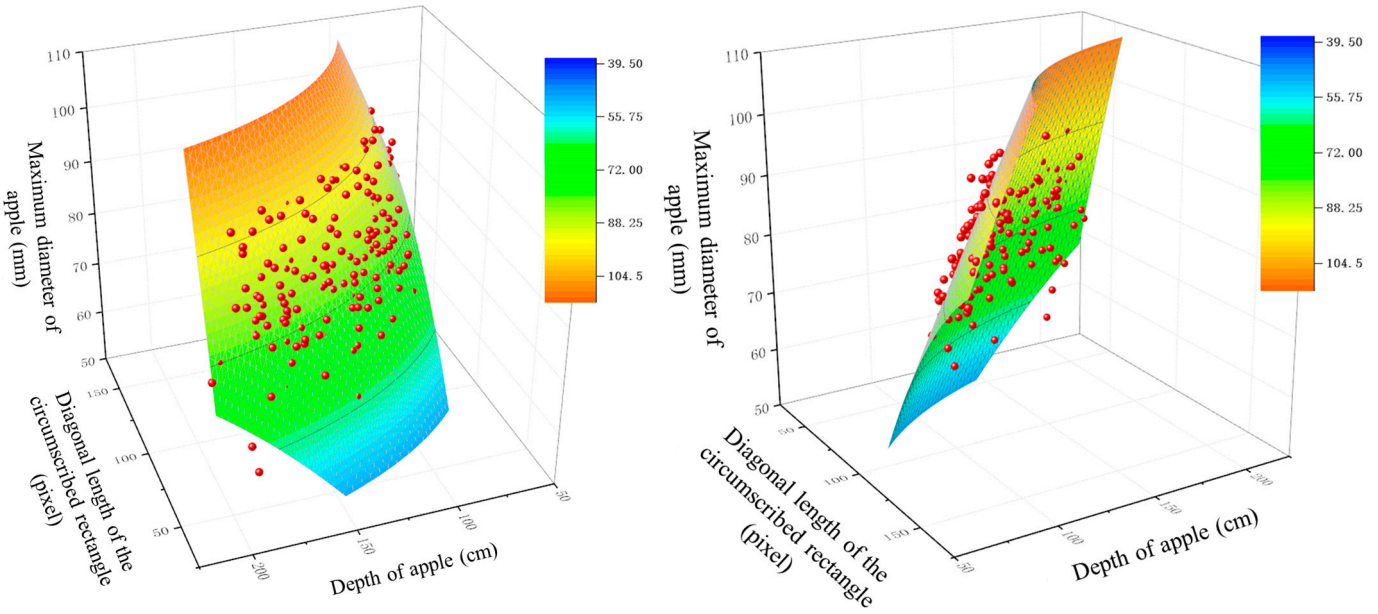


Figure 14. The 3D schematic diagram of the apple maximum diameter estimate model.

Correspondingly, the expression of the maximum diameter estimation model for apples is shown in Equation (7) below:

$$z = -0.00372(kx)^2 + 1.22207kx - 0.00112y^2 + 0.62816kx + 0.00241kxy - 82.58015 \quad (7)$$

5. Estimation Experiment and Result Analysis of Maximum Diameter of Apples

5.1. Estimation and Verification of the Maximum Diameter of Imitation Apples

Obtaining the maximum diameter parameter of imitation apples: Measurement of the maximum diameter parameter of imitation apples was conducted in the laboratory at the fully mechanized apple research base of the Ministry of Agriculture and Rural Affairs. The LF01 digital vernier caliper (measurement accuracy: 0.01 mm) produced by Quzhou Gangtuo Tools Co., Ltd., was utilized to measure the maximum equatorial circle diameter of each imitation fruit multiple times. We recorded the maximum cross-sectional diameter data obtained as the maximum cross-sectional diameter value of the imitation apple. The measurement operation is shown in Figure 15.

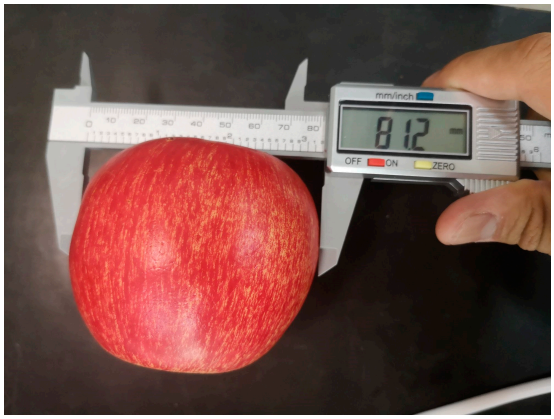


Figure 15. Measuring the maximum diameter of the imitation apple.

The measurement results and data distribution statistics of the maximum diameter of imitation apples are shown in Figure 16. It can be seen from the figure that the distribution range of the maximum diameter of imitation apples was 30–105 mm.

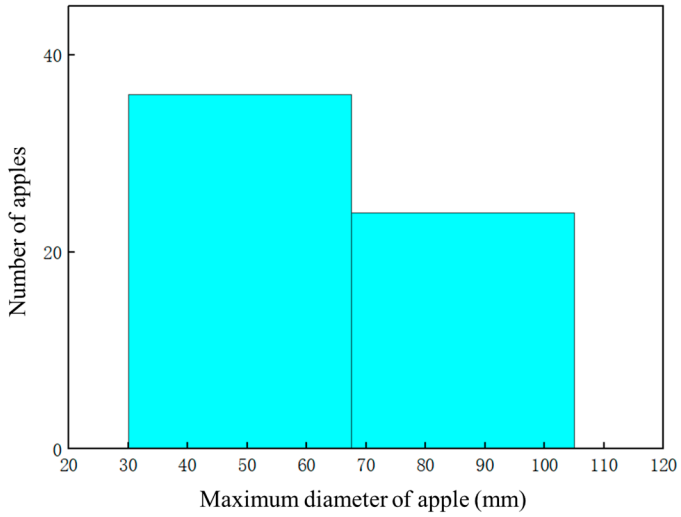


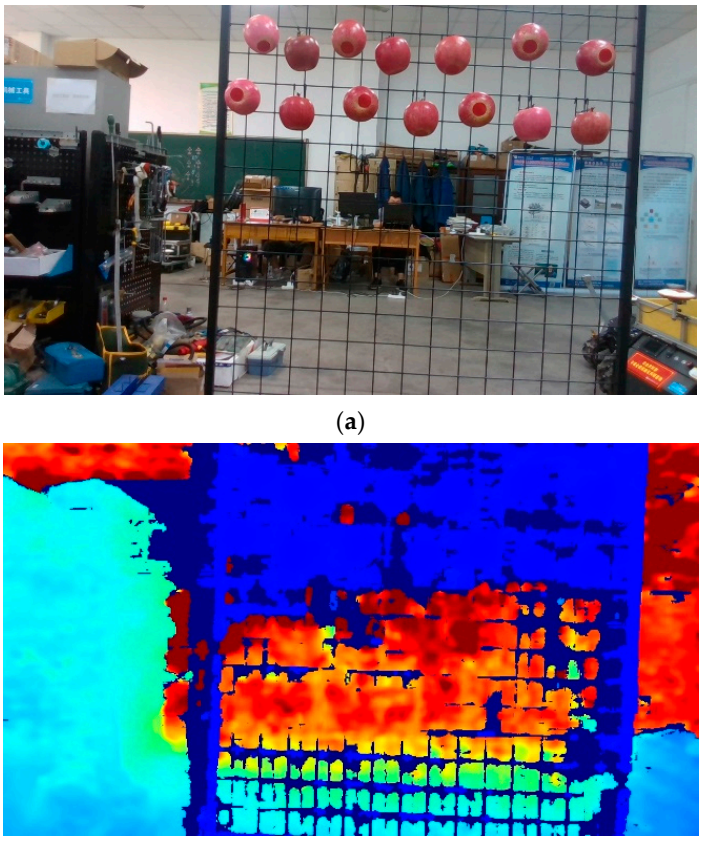
Figure 16. The maximum diameter distribution of imitation apples.

Imitation apples of different sizes in random positions were hung on the apple fruit hanging device in order to simulate the growth posture of real apples on fruit trees more realistically. We placed the fruit hanging device at different distances and positions within the field of view of the RealSense D435 camera to collect RGB and depth images of apples in order to obtain information on apples of different sizes at different distances from the camera, making the collected data more comprehensive and better verifying the maximum diameter estimation model obtained. The scenario diagram of the imitation apple validation test for the maximum diameter estimation model of apples conducted in the laboratory is shown in Figure 17.

Figure 17. Validation test situations for maximum diameter estimation of imitation apples in the laboratory.

Imitation apples of different sizes in the validation set were placed at different depths, positions, and directions facing the camera plane. Based on the D435 camera to capture RGB images of apples and combined with the image data captured by the near-infrared camera, the corresponding depth information of apples was obtained and stored. An

example of using the D435 camera to capture RGB images of apples and pseudo-color images containing depth information is shown in Figure 18.



(**a**)

(**b**)

Figure 18. Obtaining the RGB image (**a**) and corresponding pseudo-color image (**b**) containing depth information of apples based on the D435 camera.

The imitation apple depth information collected through the above operation was statistically analyzed, and the distribution of depth values is shown in Figure 19. From the figure, it can be seen that the depth distribution range of the imitation apples in the validation set was 50–300 cm.

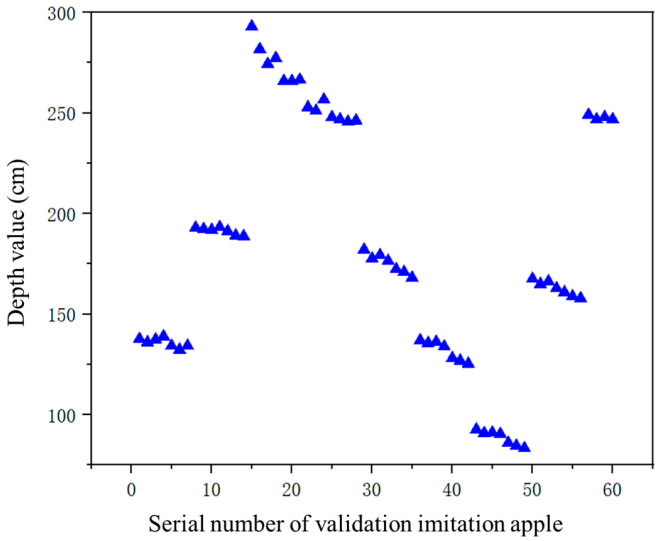


Figure 19. Depth value distribution of the validation of imitation apples.

Based on LabelImg software, we manually selected and obtained the length and width information of the bounding rectangle of each corresponding apple in the two-dimensional image (as shown in Figure 20), and then we calculated the diagonal length of the rectangle based on the Pythagorean theorem. The non-English term in Figure 20 are belonged to the storage path of the image in the computer.

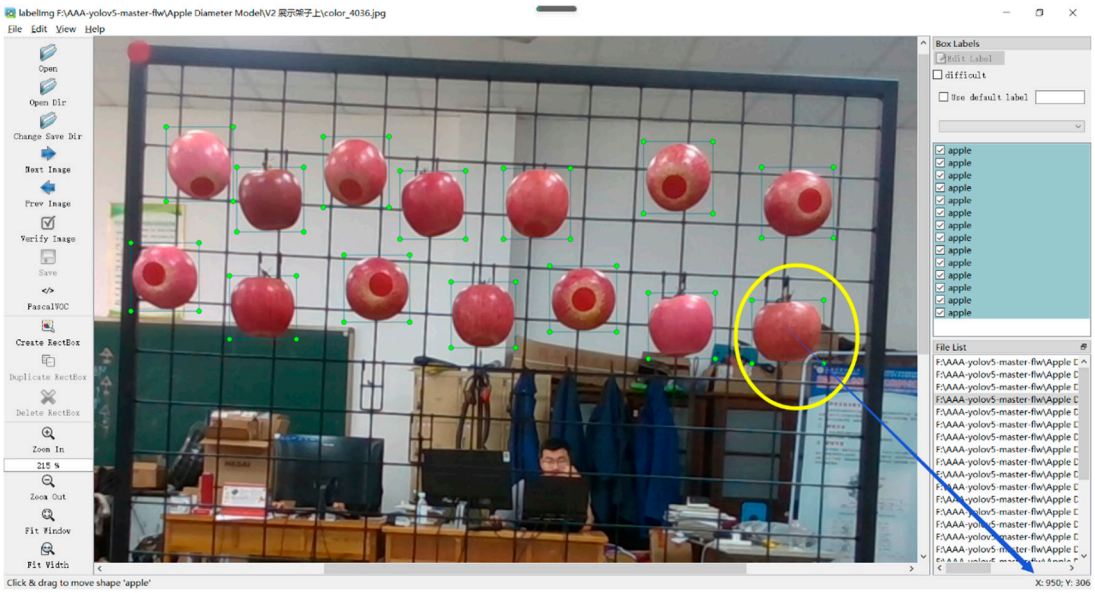


Figure 20. Obtaining the size of the external rectangle of an imitation apple based on LabelImg software.

The diagonal length distribution of the external rectangle in the imitation apple image is shown in Figure 21. From the figure, it can be seen that the diagonal length distribution range of the external rectangular box was 0–200 pixels.

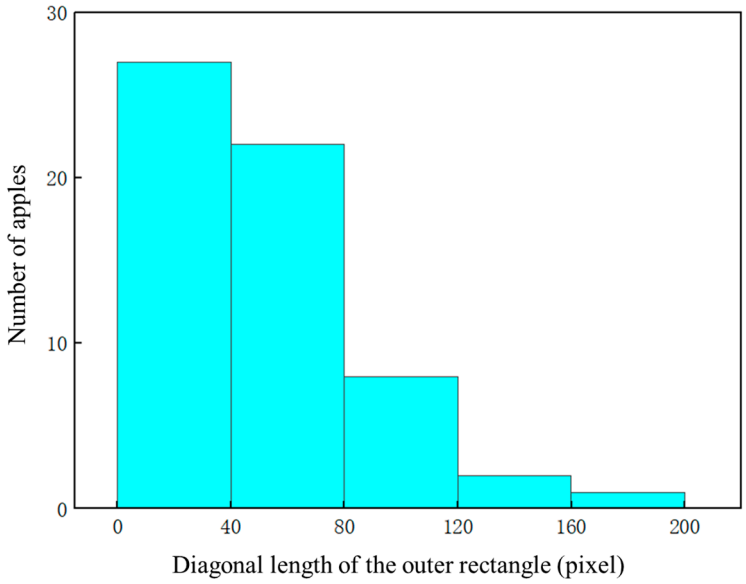


Figure 21. Diagonal length distribution of the outer rectangle of imitation apple images.

Verification results and analysis of the estimation model for imitation apples: Based on the constructed apple maximum diameter estimation model that integrates fruit depth information, the accuracy of the model was evaluated using experimental verification methods. The maximum diameter estimation results of 60 validation sets of imitation apples were compared with their actual measured values, as shown in Figure 22. It can be seen that there was a high degree of agreement between the two, which proves that the obtained maximum diameter estimation model is relatively accurate.

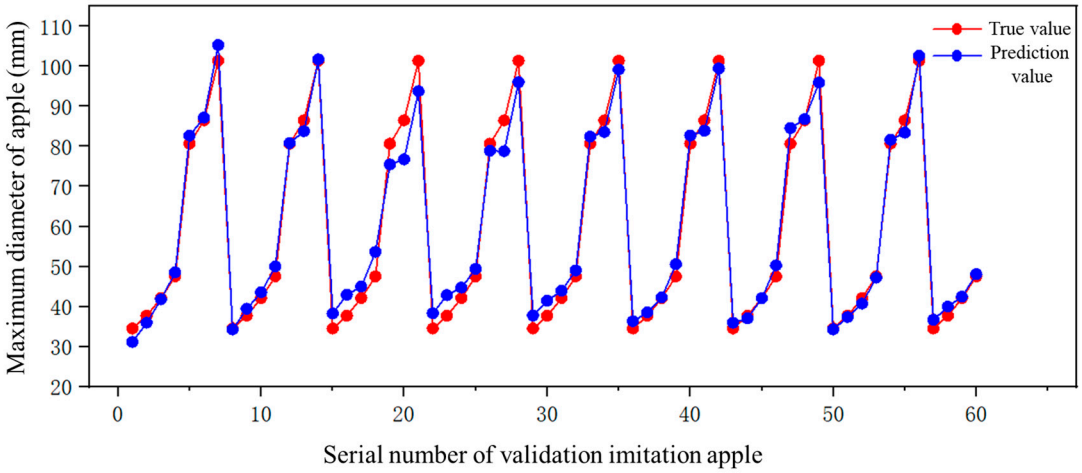


Figure 22. Validation results of the maximum outer diameter estimation model of imitation apples.

Root mean squared error (RMSE) is an evaluation metric used to measure the deviation between predicted and actual data values. The smaller the RMSE value, the more accurate the predicted results are. To measure the difference between the predicted value and the actual measurement value of the maximum diameter of the validation set apple using the fruit maximum diameter estimation model, the maximum diameter estimation relative error (*REEMax_Diameter*), average estimation relative error (*MREEMax_Diameter*), and root mean square error (RMSE) indicators are defined, which include variables such as the maximum diameter estimation value (*EMax_Diameter*), the maximum diameter actual value (*TMax_Diameter*), and the number of apples (Num_{apple}). The calculation formulas are defined as follows:

$$REE_{Max\_Diameter} = \frac{E_{Max\_Diameter} - T_{Max\_Diameter}}{T_{Max\_Diameter}} \times 100\% \quad (8)$$

$$MREE_{Max\_Diameter} = \pm\frac{\sum|E_{Max\_Diameter} - T_{Max\_Diameter}|}{\sum T_{Max\_Diameter}} \times 100\% \quad (9)$$

$$RMSE = \sqrt{\frac{\sum(E_{Max\_Diameter} - T_{Max\_Diameter})^2}{Num_{apple}}} \quad (10)$$

The relative error distribution of the maximum diameter estimation for all validation sets of imitation apples is shown in Figure 23, with an average relative error of ±4.1% for the maximum diameter estimation.

The linear correlation fitting results between the maximum diameter estimation model of the imitation apple validation set and the actual measured values is shown in Figure 24, where R^2 was 0.98613 and root mean square error (RMSE) was 3.21 mm.

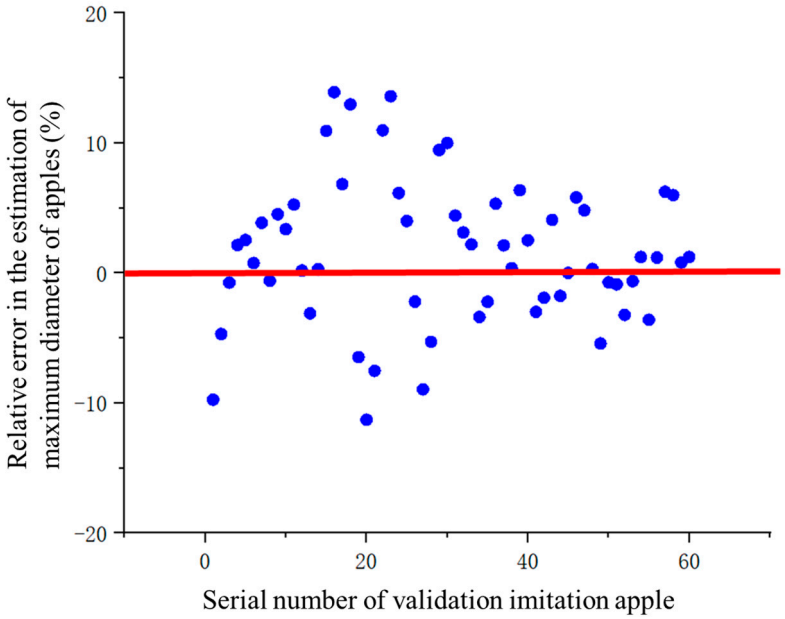


Figure 23. Distribution of the relative error in the estimation of the maximum diameter of apples in the validation set.

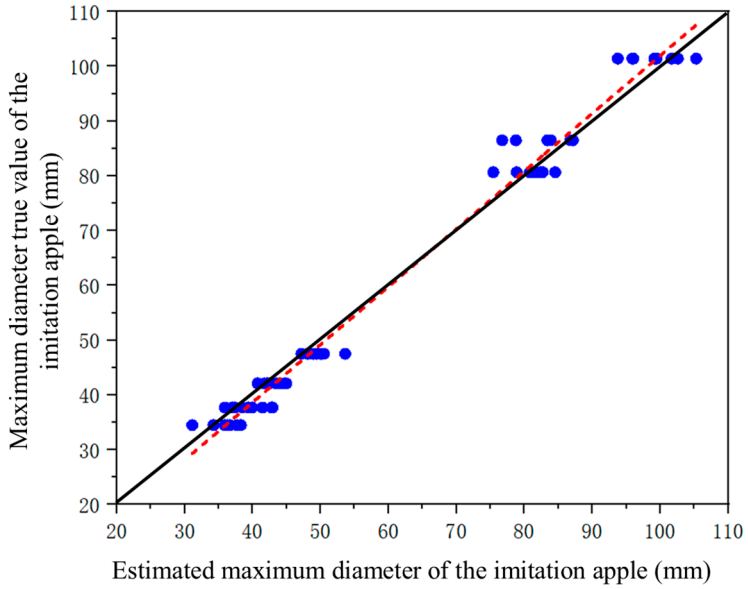


Figure 24. Linear correlation fitting results between the estimated maximum diameter and the true value of the imitation apple in the validation set.

5.2. Estimation and Verification of the Maximum Diameter of Fruits on Modern Apple Orchard Trees

Image acquisition was conducted on fruit trees located in the modern apple orchard at the Baishui Apple Experimental Station of Northwest A&F University. A total of 110 randomly selected fruits of Red Fuji apple trees under spindle-shaped cultivation mode were selected as the orchard apple validation set for the maximum diameter estimation model of apples.

The LF01 digital vernier caliper (measurement accuracy: 0.01 mm) was utilized to accurately measure and record the true maximum transverse diameter parameter value of the fruit on the apple tree. Based on the D435 depth camera installed on the apple-picking robot, two-dimensional RGB images of fruits were obtained at different distances from apple trees for different apple trees. Combined with the data collected by the near-infrared

camera, the corresponding depth information of the apples was obtained and stored. The experimental scene is shown in Figure 25.

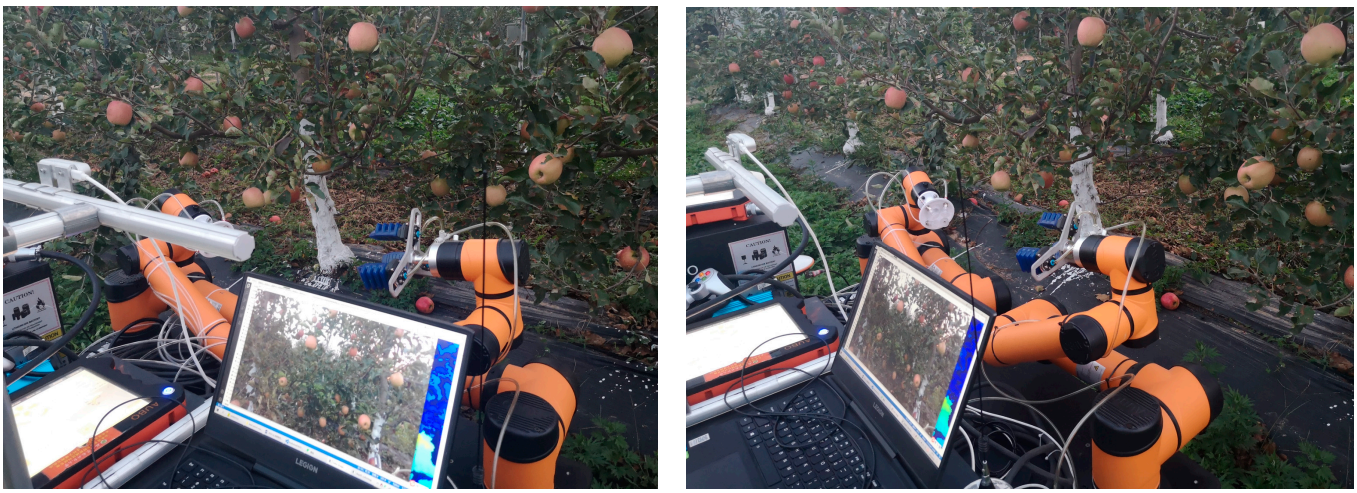


Figure 25. Verification situation of the test for estimating the maximum diameter of fruit on apple trees.

The actual measurement results of the maximum diameter of the validation set fruits on apple trees is shown in Figure 26. According to the analysis of the figure, the true maximum diameter distribution range of apples was 50–110 mm, with a concentrated distribution range of 70–110 mm.

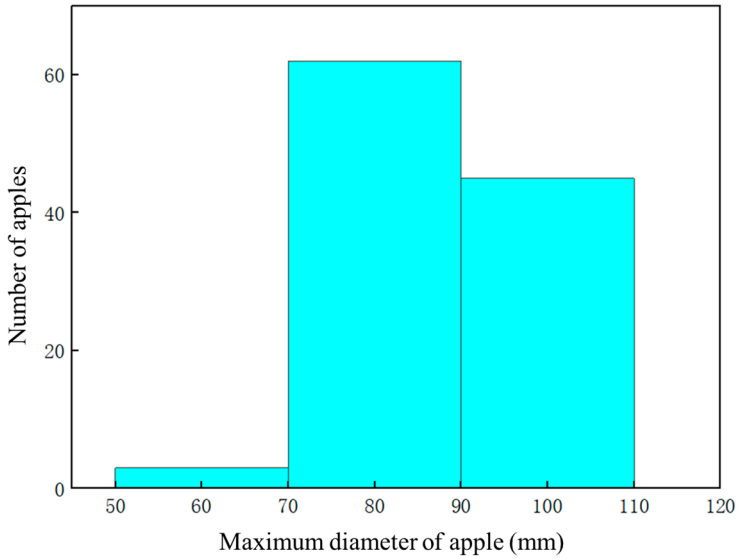


Figure 26. The maximum diameter distribution of fruit of the validation set on apple trees.

Based on LabelImg software and the collected RGB images, the length and width information of the bounding rectangle box of each corresponding validation set fruit two-dimensional image on apple trees was obtained. Then, based on the Pythagorean theorem, the diagonal length of the apple bounding rectangle box was calculated. The length calculation result is shown in Figure 27. As shown in the figure, the diagonal length of the fruit was mainly distributed between 40 and 140 pixels.

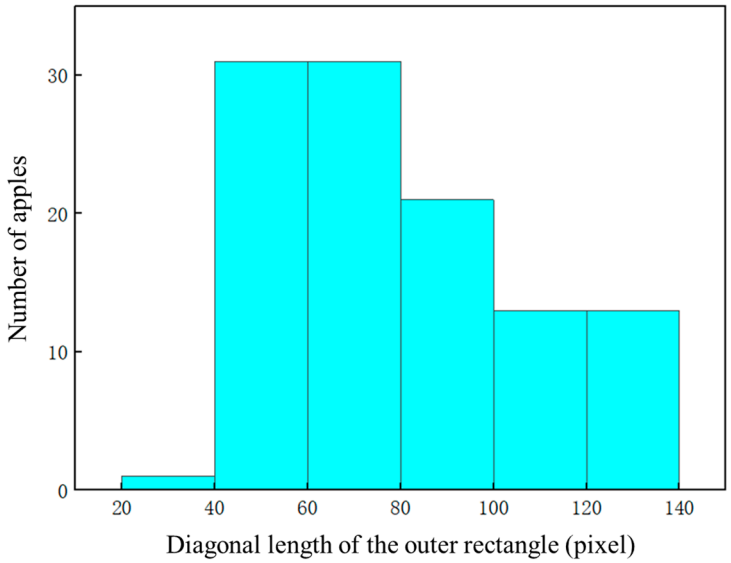


Figure 27. Distribution of the length of fruit outer rectangle diagonal on apple trees.

After obtaining the center position coordinates of the two-dimensional images of the validation set of apples on the fruit trees, the depth values corresponding to the centroid position of the fruit were extracted from the image data containing depth information collected by the D435 camera. The distribution statistics of the depth values of the validation set of apples in the orchard are shown in Figure 28. It can be seen from the following figure that the distribution range of the obtained apple depth values was 80–240 cm.

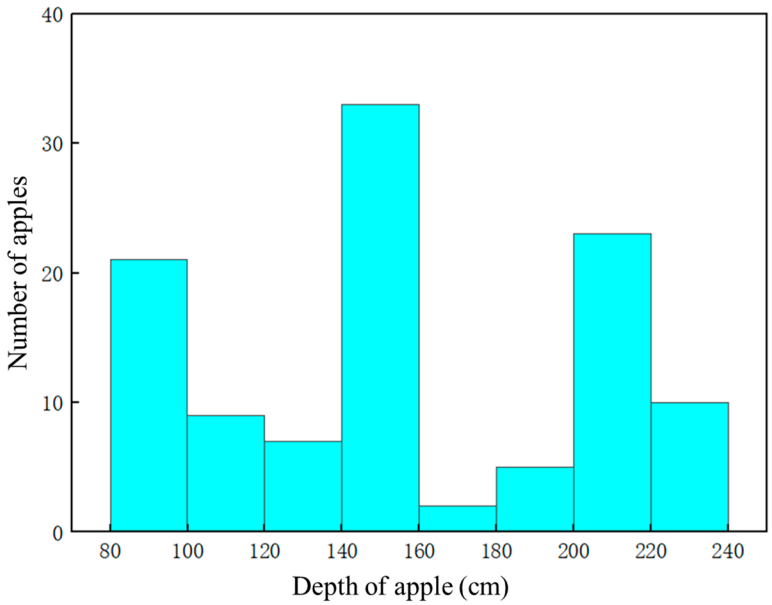


Figure 28. Distribution of fruit depth information on apple trees.

Based on the constructed model for estimating the maximum diameter of apples, its accuracy was evaluated through experimental verification. The comparison between the estimated maximum diameter of 110 validation set fruits on Red Fuji apple trees using the constructed model and their actual measurement values is shown in Figure 29.

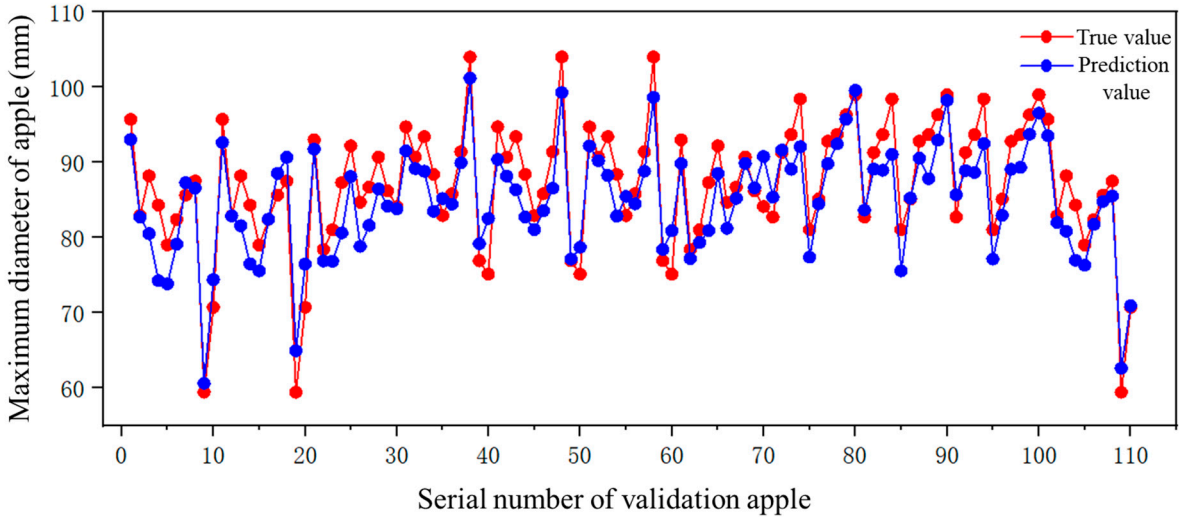


Figure 29. Validation results of the maximum diameter estimation model of apples.

Further analysis and evaluation were conducted on the validation results of the maximum diameter estimation model for fruit estimation on apple trees. The relative error of the maximum diameter estimation for all obtained validation set fruits on fruit trees was calculated, and its distribution statistics are shown in Figure 30. It can be seen that the relative estimation error of the model for the vast majority of apples was within ±10%, with an average maximum diameter estimation relative error of ±3.77%.

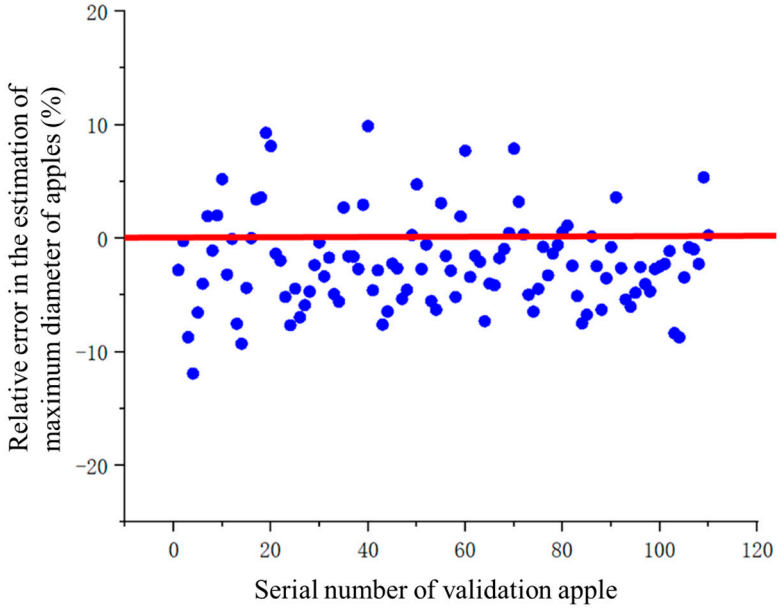


Figure 30. Error distribution of the estimated value of the maximum diameter of validation set apples.

The linear correlation between the estimated value and the true value of the maximum transverse diameter estimation model for the validation set fruits on apple trees is shown in Figure 31. Among them, R^2 was 0.84, root mean square error (RMSE) was 3.95 mm, indicating a good fitting degree, and overall indicating that the constructed model for estimating the maximum transverse diameter of apples based on RGB-D camera fusion depth information is relatively accurate, satisfying the requirements of selective grading and picking of fruits on apple trees by harvesting robots.

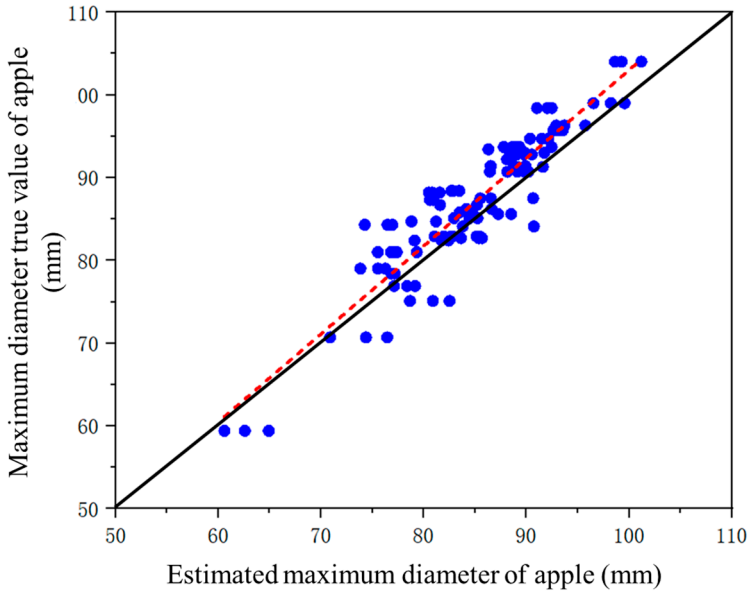


Figure 31. The linear correlation fitting results between the estimated value and true value of the maximum diameter of validation set apples in the modern orchard.

The process of estimating the maximum diameter of fruits on apple trees mainly includes the following steps: Obtaining apple images based on a depth camera, detecting the length and width of the bounding rectangle of the apple, calculating the diagonal length of the rectangle, obtaining the depth value of the center position of the rectangle based on the depth camera, inputting the diagonal length and depth value into the fruit maximum diameter calculation model, and obtaining the maximum diameter of the fruit. The flowchart, which shows the overall methodological approach, is shown as follows in Figure 32:

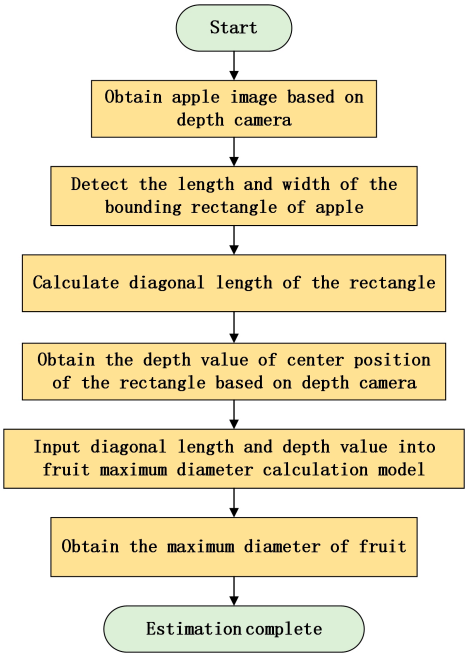


Figure 32. The flowchart of the overall methodological approach.

Based on the proposed maximum lateral diameter estimation model for apples, it is embedded into the YOLOv5 apple target recognition and localization algorithm to achieve real-time estimation of the maximum lateral diameter while detecting and locating apple

targets. The real-time estimation effect of the picking robot on the maximum transverse diameter of the apple target is shown in the following Figure 33 (the maximum transverse diameter estimation value of the corresponding fruit is displayed above the detection box).

Figure 33. Real-time estimation of the maximum diameter of apple targets.

6. Conclusions

In response to the inability of existing algorithms for estimating the external dimensions of apples to provide theoretical and algorithmic support for robots to selectively pick and grade apples based on their size, a method for estimating the diameter of apples on fruit trees for the actual working scenario of picking robots in orchards was proposed in this study, in order to support in the integration of "intelligent picking and grading" of apple targets on fruit trees.

The maximum transverse diameter parameter of Red Fuji apples was collected from the fruits of modern apple orchards, and the results were statistically analyzed. The depth information of apples on fruit trees and the diagonal length parameter of the bounding rectangle of the fruit were obtained based on the D435 depth camera and LabelImg software. Based on Origin software, multiple regression analysis and nonlinear surface fitting methods were used to explore the relationship between fruit depth, diagonal length of the bounding rectangle, and maximum transverse diameter. A model for estimating the maximum transverse diameter of apples based on multiple regression and nonlinear surface fitting was constructed. Finally, maximum diameter estimation experiments were conducted to verify and evaluate the constructed model for imitation apples and fruits on the Red Fuji fruit tree in modern apple orchards at the Apple Full Mechanization Research Base and the Baishui Apple Test Station of Northwest A&F University. The experimental results show that the average maximum relative error of the constructed model in the laboratory imitation apple validation set was $\pm$4.1%, the correlation coefficient (R^2) of the estimated model was 0.98613, and the root mean square error (RMSE) was 3.21 mm. The average maximum diameter estimation relative error on the modern orchard Red Fuji fruit validation set in the apple testing station was $\pm$3.77%, the correlation coefficient (R^2) of the estimation model was 0.84, and the root mean square error (RMSE) was 3.95 mm. The proposed model can provide a theoretical algorithm basis and technical support for the selective apple-picking work of intelligent robots based on apple size grading.

7. Discussion and Future Work

According to the actual test results of the proposed fruit maximum diameter estimation model in apple orchards, it can be seen that the accuracy of the model can basically meet the requirements of fruit selective grading and picking for apple-picking robots based on the maximum diameter of the fruit. However, the estimation error of fruit diameter in

the experimental results shows that there is still significant room for improvement in the accuracy of the model. Therefore, further work needs to collect more apple data of different fruit diameter sizes, expand the dataset used for estimating model modeling, and obtain a more accurate estimation model. On the other hand, the apple variety targeted in this study is the widely planted Red Fuji apple, so the effectiveness of the proposed fruit diameter estimation model in estimating the maximum diameter of other apple varieties is unknown. The excellent universality of the estimation model is necessary for the future development of smart agriculture. Therefore, in later research work, it is necessary to include data from multiple commonly planted apple varieties in the original dataset for modeling a universal estimation model of the maximum diameter of various apple fruits.

In addition, the night picking operation of apple-picking robots is also a development direction for smart agriculture and precision agriculture. Therefore, in later research, it is necessary to collect depth image data of apples at night and then conduct estimation experiments on the maximum diameter of apple fruits at night to verify and optimize the maximum diameter estimation model of apples, ultimately achieving the day/night selective picking operation of apple-picking robots.

Author Contributions: All authors contributed to this manuscript. B.Y. contributed to the development of the algorithm, obtaining data, programming, and writing. B.Y. also performed the experiments, analyzed the results, and contributed to funding acquisition. X.L. contributed to the original draft preparation. All authors have read and agreed to the published version of the manuscript.

Funding: This work was financially supported by the National Natural Science Foundation of China (NSFC) (Grant No. 62406244), the National Natural Science Foundation of China (NSFC) (Grant No. 62473311), and the Doctoral Research Project of Xi'an University of Technology (Grant No. 103-451123011).

Data Availability Statement: The original contributions presented in the study are included in the article, and further inquiries can be directed to the corresponding author.

Acknowledgments: We sincerely thank the editors and reviewers for their detailed comments and efforts toward improving our study.

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

References

1. Yan, B.; Fan, P.; Wang, M.; Shi, S.; Lei, X.; Yang, F. Real-time apple picking pattern recognition for picking robot based on improved YOLOv5m. *Trans. CSAM* **2022**, *53*, 28–38+59. [CrossRef]
2. Yan, B.; Fan, P.; Lei, X.; Liu, Z.; Yang, F. A Real-Time Apple Targets Detection Method for Picking Robot Based on Improved YOLOv5. *Remote Sens.* **2021**, *13*, 1619. [CrossRef]
3. Gao, F.; Fu, L.; Zhang, X.; Majeed, Y.; Li, R.; Karkee, M.; Zhang, Q. Multi-class fruit-on-plant detection for apple in SNAP system using Faster R-CNN. *Comput. Electron. Agric.* **2020**, *176*, 105634. [CrossRef]
4. Kang, H.; Zhou, H.; Chen, C. Visual Perception and Modeling for Autonomous Apple Harvesting. *IEEE Access* **2020**, *8*, 62151–62163. [CrossRef]
5. Yang, F.; Lei, X.; Liu, Z.; Fan, P.; Yan, B. Fast Recognition Method for Multiple Apple Targets in Dense Scenes Based on CenterNet. *Trans. CSAM* **2022**, *53*, 265–273. [CrossRef]
6. Ma, H.; Li, Y.; Zhang, X.; Li, Y.; Li, Z.; Zhang, R.; Zhao, Q.; Hao, R. Target Detection for Coloring and Ripening Potted Dwarf Apple Fruits Based on Improved YOLOv7-RSES. *Appl. Sci.* **2024**, *14*, 4523. [CrossRef]
7. Liu, J.; Zhao, G.; Liu, S.; Liu, Y.; Yang, H.; Sun, J.; Yan, Y.; Fan, G.; Wang, J.; Zhang, H. New Progress in Intelligent Picking: Online Detection of Apple Maturity and Fruit Diameter Based on Machine Vision. *Agronomy* **2024**, *14*, 721. [CrossRef]
8. Sekharamantry, P.K.; Melgani, F.; Malacarne, J.; Ricci, R.; de Almeida Silva, R.; Marcato Junior, J. A Seamless Deep Learning Approach for Apple Detection, Depth Estimation, and Tracking Using YOLO Models Enhanced by Multi-Head Attention Mechanism. *Computers* **2024**, *13*, 83. [CrossRef]
9. Gao, F.; Wu, Z.; Suo, R.; Zhou, Z.; Li, R.; Fu, L.; Zhang, Z. Apple detection and counting using real-time video based on deep learning and object tracking. *Trans. CSAE* **2021**, *37*, 217–224. [CrossRef]
10. Zhang, Z.; Jia, W.; Shao, W.; Hou, S.; Ze, J.; Zheng, Y. Green Apple Detection Based on Optimized FCOS in Orchards. *Spectrosc. Spectr. Anal.* **2022**, *42*, 647–653. [CrossRef]
11. Sun, J.; Qian, L.; Zhu, W.; Zhou, X.; Dai, C.; Wu, X. Apple detection in complex orchard environment based on improved RetinaNet. *Trans. CSAE* **2022**, *38*, 314–322. [CrossRef]

12. Wang, Z.; Wang, J.; Wang, X.; Shi, J.; Bai, X.; Zhao, Y. Lightweight Real-time Apple Detection Method Based on Improved YOLO v4. *Trans. CSAM* **2022**, *53*, 294–302.
13. Hu, G.; Zhou, J.; Chen, C.; Li, C.; Sun, L.; Chen, Y.; Zhang, S.; Chen, J. Fusion of the lightweight network and visual attention mechanism to detect apples in orchard environment. *Trans. CSAE* **2022**, *38*, 131–142. [CrossRef]
14. Ye, R.; Gao, Q.; Qian, Y.; Sun, J.; Li, T. Improved YOLOv8 and SAHI Model for the Collaborative Detection of Small Targets at the Micro Scale: A Case Study of Pest Detection in Tea. *Agronomy* **2024**, *14*, 1034. [CrossRef]
15. Yang, S.; Yao, J.; Teng, G. Corn Leaf Spot Disease Recognition Based on Improved YOLOv8. *Agriculture* **2024**, *14*, 666. [CrossRef]
16. Yan, B.; Li, X.; Yan, W. Deep Learning-Based Biomimetic Identification Method for Mask Wearing Standardization. *Biomimetics* **2024**, *9*, 563. [CrossRef]
17. Yan, B.; Quan, J.; Yan, W. Three-Dimensional Obstacle Avoidance Harvesting Path Planning Method for Apple-Harvesting Robot Based on Improved Ant Colony Algorithm. *Agriculture* **2024**, *14*, 1336. [CrossRef]
18. Yan, B.; Liu, Y.; Yan, W. A Novel Fusion Perception Algorithm of Tree Branch/Trunk and Apple for Harvesting Robot Based on Improved YOLOv8s. *Agronomy* **2024**, *14*, 1895. [CrossRef]
19. Kang, H.; Chen, C. Fruit detection, segmentation and 3D visualisation of environments in apple orchards. *Comput. Electron. Agric.* **2020**, *171*, 105302. [CrossRef]
20. Fu, L.; Majeed, Y.; Zhang, X.; Karkee, M.; Zhang, Q. Faster R-CNN-based apple detection in dense-foliage fruiting-wall trees using RGB and depth features for robotic harvesting. *Biosyst. Eng.* **2020**, *197*, 245–256. [CrossRef]
21. Long, Y.; Li, N.; Gao, Y.; He, M.; Song, H. Apple fruit detection under natural condition using improved FCOS network. *Trans. CSAE* **2021**, *37*, 307–313. [CrossRef]
22. Song, H.; Jiang, M.; Wang, Y.; Song, L. Efficient detection method for young apples based on the fusion of convolutional neural network and visual attention mechanism. *Trans. CSAE* **2021**, *37*, 297–303. [CrossRef]
23. Song, H.; Ma, B.; Shang, Y.; Wen, Y.; Zhang, S. Detection of Young Apple Fruits Based on YOLO v7-ECA Model. *Trans. CSAM* **2023**, *54*, 233–242. [CrossRef]
24. Sofu, M.; Er, O.; Kayacan, M.; Cetişli, B. Design of an automatic apple sorting system using machine vision. *Comput. Electron. Agric.* **2016**, *127*, 395–405. [CrossRef]
25. Vakilian, K.A.; Massah, J. An apple grading system according to european fruit quality standards using gabor filter and artificial neural networks. *Sci. Study Res. Chem. Chem. Eng.* **2016**, *17*, 75–81.
26. Li, L.; Peng, Y.; Li, Y. Design and experiment on grading system for online non-destructive detection of internal and external quality of apple. *Trans. CSAE* **2018**, *34*, 267–275. [CrossRef]
27. Zhang, Q.; Gu, B.; Ji, C.; Fang, H.; Guo, J.; Shen, W. Design and experiment of an online grading system for apple. *J. South China Agric. Univ.* **2017**, *38*, 117–124. [CrossRef]
28. Li, J.; Zhang, D.; Liu, B. Design of Conveyer and Turnover Mechanism of Apple Grader. *Trans. CSAM* **2009**, *40*, 158–161+157.
29. Hu, G.; Zhang, E.; Zhou, J.; Zhao, J.; Gao, Z.; Sugirbay, A.; Jin, H.; Zhang, S.; Chen, J. Infield Apple Detection and Grading Based on Multi-Feature Fusion. *Horticulturae* **2021**, *7*, 276. [CrossRef]
30. Fan, Z.; Liu, Q.; Chai, J.; Yang, X.; Li, H. Apple detection and grading based on color and fruit-diameter. *Comput. Eng. Sci.* **2020**, *42*, 1599–1607. [CrossRef]
31. Lu, S.; Chen, W.; Zhang, X.; Karkee, M. Canopy-attention-YOLOv4-based immature/mature apple fruit detection on dense-foliage tree architectures for early crop load estimation. *Comput. Electron. Agric.* **2022**, *193*, 106696. [CrossRef]
32. Yu, J.; Zhang, X.; Wu, T.; Pan, H.; Zhang, W. A Face Detection and Standardized Mask-Wearing Recognition Algorithm. *Sensors* **2023**, *23*, 4612. [CrossRef] [PubMed]
33. Long, S.; Huang, W.; Wang, J.; Liu, J.; Gu, Y.; Wang, Z. A Fixed-Time Consensus Control with Prescribed Performance for Multi-Agent Systems Under Full-State Constraints. *IEEE Trans. Autom. Sci. Eng.* **2024**, 1–10. [CrossRef]
34. Wang, J.; Wu, Y.; Chen, C.L.P.; Liu, Z.; Wu, W. Adaptive PI event-triggered control for MIMO nonlinear systems with input delay. *Inf. Sci.* **2024**, *677*, 120817. [CrossRef]