Estimating Xisha watermelon yield using sampling and global scanning based on drone remote sensing

Xiaofei Zhang[1], Yuqing Zheng[1], Ao Shen[1], Yi Xun[1,2], Qinghua Yang[3,4], Zhiheng Wang[1,2*]

(1. *College of Mechanical Engineering, Zhejiang University of Technology, Hangzhou 310032, China*;
2. *Taizhou Key Laboratory of Advanced Manufacturing Technology, Taizhou Institute, Zhejiang University of Technology, Taizhou 318014, Zhejiang, China*;
3. *College of Optical, Mechanical and Electrical Engineering, Zhejiang A & F University, Hangzhou 310032, China*;
4. *The Collaborative Innovation Center for Intelligent Production Equipment of Characteristic Forest Fruits in Hilly and Mountainous Areas of Zhejiang Province, Zhejiang A & F University, Hangzhou 310032, China*)

Abstract: Xisha watermelon, one variety of selenium-rich economic crops, has been widely cultivated in the semi-arid regions of northwest China. It is critical to estimate its yield for the decision making on the harvesting and market. This paper presents a yield estimation model for Xisha watermelon using drone remote sensing in digital agriculture. Firstly, the watermelon weight was estimated in three steps: object detection with YOLOv8n, contour fitting with the functions from the OpenCV library, and weight estimation of individual watermelon through a volume-to-weight model. Then, two yield estimation strategies were developed. 1) Sampling: the total yield of watermelons over the entire plot was calculated using the average yield within sampled units and the plot area. 2) Global scanning: an overall yield distribution of watermelons was obtained to scan the entire plot using orthoimagery. Finally, a series of field tests was carried out to verify the estimation in plantations. The results reveal that the average accuracy of the detection model was 0.986 using YOLOv8n. Once the number of watermelons exceeded 45, the relative error between the total estimated and the measured weight was less than 1.00%. The speed of sampling was 27.37 m^2/s for a 9000 m^2 field size of Xisha watermelon, approximately 50 times higher than that of global scanning. Compared with global scanning, the sampling-based estimation underestimated the count by 1.77% and the total weight by 5.10%, both of which fall within an acceptable range. Each estimation can be suitable for the specific scenarios of application. The sampling can be expected to provide the higher efficiency for the total field yield. While the global scanning can effectively represent the overall yield distribution of Xisha watermelons in the field. This study provides a new research approach and direction for fruit and vegetable yield estimation in precision agriculture based on UAV remote sensing technology.

Keywords: yield estimation, drone remote sensing, object detection, Xisha watermelon
DOI: 10.25165/j.ijabe.20261902.9946

Citation: Zhang X F, Zheng Y Q, Shen A, Xun Y, Yang Q H, Wang Z H. Estimating Xisha watermelon yield using sampling and global scanning based on drone remote sensing. Int J Agric & Biol Eng, 2026; 19(2): 272–281.

1 Introduction

Farm field data has been widely used to develop the agronomic operations in digital agriculture[1]. Intelligent equipment is necessary to acquire the farm field information and then process the high-throughput image data. Among them, drones (known as unmanned aerial vehicles, UAVs) can be expected to rapidly capture the high-throughput images in smart agriculture[2]. Image sensors can also be integrated with drones in various scenarios, such as weed detection[3], disease detection[4], crop growth monitoring[5], and yield estimation[6]. Among them, the crop yield can be estimated using crop area, density, distribution, total yield, and individual weight[7]. This is of great importance for decision-making on harvesting, transportation, storage, and sales in agricultural production[8].

A special variety of melon, Xisha watermelon (known as selenium sand melon), has been one of the most important economic crops in the semi-arid regions of northwest China[9]. The alkaline soil also allows the Xisha watermelon to have a large amount of mineral elements beneficial for human health, particularly in Zhongwei City in Ningxia Hui Autonomous Region[10]. Soil moisture can be maintained using a layer of gravel on the soil surface in arid mountainous areas, thus resulting in the sparse foliage in the maturation period of Xisha watermelons. As such, the cultivation of Xisha watermelon can greatly contribute to the emerging sustainable industry against land desertification. However, a significant challenge remains for the relatively limited varieties and short harvesting time from July to August under local climatic conditions. Furthermore, it is difficult to realize efficient harvesting, transportation, and sales in the large-scale planting area for only a short period. Fortunately, digital harvesting can be expected to predict the yield of Xisha watermelons in advance.

The linear relationship between fruit weight and size is an essential prerequisite to estimate the fruit weight using machine vision[11]. Huynh et al.[12] proposed a vision-based method that estimates sweet potato volume using a virtual slicing approach from single-view RGB images and establishes a volume-weight model to

Received date: 2025-05-30 **Accepted date:** 2026-01-11

Biographies: Xiaofei Zhang, PhD candidate, research interest: agricultural automation and robotics, Email: zxf843390242@163.com; **Yuqing Zheng**, MS candidate, research interest: drone remote sensing technology, Email: yuqin_zheng@foxmail.com; **Ao Shen**, MS candidate, research interest: drone remote sensing technology, Email: 545491180@qq.com; **Yi Xun**, PhD, Lecturer, research interest: robotics and automation, Email: xunyi@zjut.edu.cn; **Qinghua Yang**, PhD, Professor, research interest: agriculture robot, Email: robot@zjut.edu.cn.

***Corresponding author: Zhiheng Wang**, PhD, Associate Professor, research interest: robot motion control, embedded system development and application, UAV remote sensing. Zhejiang University of Technology, Hangzhou 310032, China. Tel: +86-13777860388, Email: wzh232@zjut.edu.cn.

achieve fast and accurate non-destructive mass estimation. Sabouri et al.[13] applied machine learning regression models to two-dimensional fruit size features extracted from RGB images to achieve non-destructive estimation of plum fresh weight. In the field of UAV-based yield estimation, Kalantar et al.[14,15] developed volume-weight estimation models for individual melons, demonstrating the feasibility of image-processing-based yield estimation. They also detailed the complete yield estimation workflow, with the estimated weights being approximately 3% lower than the measured values. Wittstruck et al.[16] used a pixel-based random forest approach combined with morphological filtering and ellipse fitting for pumpkin identification, followed by the establishment of an ellipsoid-based single-fruit weight estimation model, achieving a detection rate of 95% and enabling a two-dimensional yield distribution map. These studies collectively demonstrate that estimating the weight of specific fruits and vegetables through image-based measurements and computational algorithms is feasible and effective, thereby providing a solid foundation for conducting yield estimation of Xisha watermelon using UAV imagery.

In the field of watermelon yield estimation, Ekizi et al.[17,18] sequentially applied K-means clustering and a multiclass RBF-SVM to identify watermelons, improving the classification accuracy of watermelon, leaf, and soil features from 86.46% to 96.51%. Qiu et al.[19] proposed an improved U-Net-based segmentation model for watermelon fruits, achieving an accuracy of 99.03%. Jiang et al.[20] developed a YOLOv8s-based watermelon detection model with an accuracy of 99.20% and achieved over 96.61% counting accuracy using panoramic image stitching and overlapping segmentation. Yin et al.[21] proposed a UAV-based automatic watermelon counting method by combining YOLOv7-GCSF with DeepSORT, achieving 98.2% detection precision and 96.3% counting accuracy. In summary, existing studies on watermelon yield estimation mainly focus on object detection and counting, while a complete weight estimation model applicable to field-scale total yield assessment has not yet been established. However, total yield is a key indicator for economic benefit analysis and harvest planning in agricultural production. On the one hand, Xisha watermelons grow in semi-arid, gravel-mulched fields with uneven terrain, leading to variations in fruit altitude and tilt. These factors alter the apparent image size of watermelons, which in turn impacts the accuracy of the image-based volume-weight model. Therefore, it is necessary to evaluate the applicability of this model in real field conditions. In addition, most existing methods rely on full-field UAV scanning, which results in low counting efficiency under large-scale planting conditions and fails to meet the practical requirements for rapid yield estimation. To date, few studies have attempted to introduce a sampling-based strategy for watermelon yield estimation, nor have they addressed the weight estimation modeling problem caused by incomplete watermelon targets resulting from image cropping during the sampling process.

In this paper, a yield estimation model integrating sampling and global scanning is proposed to fully meet different field requirements, and the data acquisition methods of these two approaches are illustrated in Figure 1. The model is based on UAV digital orthophoto imagery and estimates individual fruit weight and field-scale total yield through object detection, contour fitting, and ellipsoid-based weight modeling. To overcome the low efficiency of traditional global scanning, a sampling method is introduced for rapid yield estimation. Meanwhile, to address the difficulty of weight estimation caused by incomplete fruit targets during sampling and image cropping, an ellipsoid point cloud estimation model is developed to enable reasonable weight estimation of partially cropped fruits. In addition, the global scanning mode is used to generate yield distribution maps of watermelon, and overlapping buffer zones are introduced to effectively resolve fruit attribution issues at image boundaries.

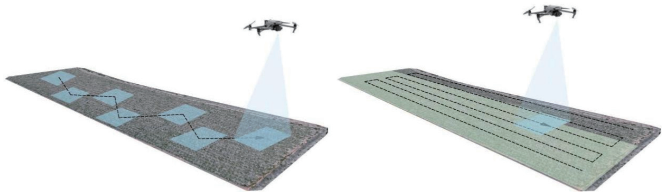
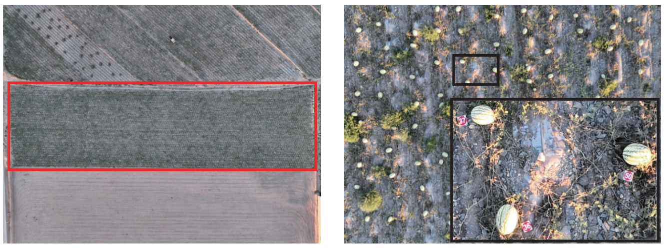


Figure 1 Sampling (left) and global scanning (right) for yield estimation

2 Materials and methods

2.1 Data collection

The experimental site was selected from the Xisha watermelon land in Zhongning County, Zhongwei City, Ningxia Hui Autonomous Region (36°56′59.4536″N; 105°16′57.6227″E). Xisha watermelon was cultivated using a gravel cover on the soil surface. Water resource utilization was optimized at planting points using pipe drip irrigation. Sparse foliage and vines of Xisha watermelons shared the minimal shading from leaves without overlapping fruits in the planting area[22]. The plot was approximately rectangular with the size of 50×190 m, as shown in Figure 2a. Watermelon plants were cultivated with a row spacing of 1.5 m and an intra-row spacing of 1.4 m. Furthermore, the thinning operation was performed to concentrate the nutrition for the Xisha watermelons in the period of fruit setting. Only one fruit remained per vine, leading to a relatively uniform distribution of watermelons in the field.



a. Overall view of selected field b. Marked Xisha watermelons

Figure 2 Experimental site

Aerial image data was collected from the watermelons at 20:04 on July 28, 2023 using a DJI Mavic 3M (SZ DJI Technology Co., Ltd., Shenzhen, China) drone equipped with a visible light camera from M3M model. Automatic cruise shooting mode was set in the drone with a focal length of 12 mm, flying at an altitude of 12 m and a speed of 2 m per second. The camera lens was also positioned vertically downward during image acquisition. A total of 796 images were collected with a resolution of 5280×3956 pixels, each containing approximately 130 watermelons.

Prior to data collection, 120 Xisha watermelons were randomly selected from the target plot and then marked with numbered wooden boards placed next to them. These boards faced upwards to be captured by the drone, as shown in Figure 2b. The lengths of the major and minor axes and the weights were measured for these 120 marked watermelons. Among them, the lengths of axis were measured using a vernier caliper with a precision of 1 mm. The weights were measured using an industrial electronic scale with a precision of 1 g.

2.2 Individual watermelon weight estimation

An axially symmetric shape is often assumed for watermelons[22]. To verify this assumption, two short axes perpendicular to the major axis and differing by 90° in orientation were measured. The results showed that the relative error between the two short axes was within 5%, supporting the assumption of an approximately symmetric cross-section. Considering the ellipsoidal shape of the watermelon, the following volume calculation formula was adopted for estimation:

$$V = \frac{4\pi a^2 b}{3} \tag{1}$$

where, V is the volume of the watermelon, m^3; a is the length of the short axis, m; and b is the length of the long axis, m.

Ground Sample Distance (GSD) represents the actual ground distance corresponding to a single pixel in the image. By using GSD, the pixel dimensions of the long and short axes of watermelons in the image can be converted into their estimated real-world sizes. The formula for calculating GSD is as follows:

$$\mathrm{GSD} = \frac{h \cdot \Delta p}{f} \tag{2}$$

where, h is the altitude at which the image was acquired; Δp is the sensor pixel size; and f is the focal length setting during image acquisition. The visible light camera in the M3M shared a sensor pixel size of 3.275 μm, and the focal length of the lens was 12 mm. The flight height of the drone relative to the planting surface of Xisha watermelon was 12 m, resulting in a GSD of 3.275 mm/pixel.

2.3 Object detection algorithm

An image dataset was randomly selected from 60 of 796 images in the Xisha watermelon fields via drone aerial photography. Then, the bounding boxes were manually labeled around the Xisha watermelons using MakeSense annotation tool. Among them, 30 images were selected as the training set, resulting in a total of about 4000 training targets. The remaining 30 images were used as the validation set to evaluate the performance of the model.

YOLOv8n algorithm was adopted as the base model for the object detection of the images. The training was conducted using a deep learning framework built on Windows 10, Python 3.8, CUDA 12.1, and PyTorch 2.1.0, on a computer equipped with an Intel(R) Core(TM) i7-7700HQ CPU, NVIDIA GTX 1050ti 4G GDDR5 GPU, and 8GB of RAM, with Visual Studio Code as the development environment. Accurate identification was realized with the Intersection over Union (IoU) confidence threshold of 0.7[15].

2.4 Contour extraction and fitting

The center position and boundary information of each detected box were obtained after object detection using YOLOv8n. Only one Xisha watermelon to be estimated was present within each detection box, according to the growth characteristics of Xisha watermelons after thinning operation. As shown in Figure 3, the contour extraction and fitting were then conducted on the individual watermelons in five steps:

1) Mask box construction. A mask box was constructed to detect the target. Twenty pixels were added into each side of the detection box with its center position as the origin. A mask box was then formed to expand the size. As such, the targets were fully captured from the mask box to more accurately extract the geometric features of Xisha watermelons. A mask was finally added into the image, according to the modified detection box.

2) Grayscale processing. The grayscale images were obtained to convert the content from the mask box. Some interferences were also removed from the irrelevant information. Among them, bitwise operations were also conducted with all pixels of 0 outside the mask box, and only the content within the detection box.

3) Binarization. A threshold was selected to binarize the grayscale image.

4) Contour extraction. The contours were extracted from the image using “cv2.findContours” function in OpenCV. The longest contour was retained in the image after extraction.

5) Ellipse fitting. The “cv2.fitEllipse” function was utilized to fit an ellipse, according to the scatter points of the longest contour. The ellipse was then drawn onto the initial image after coordinate mapping.

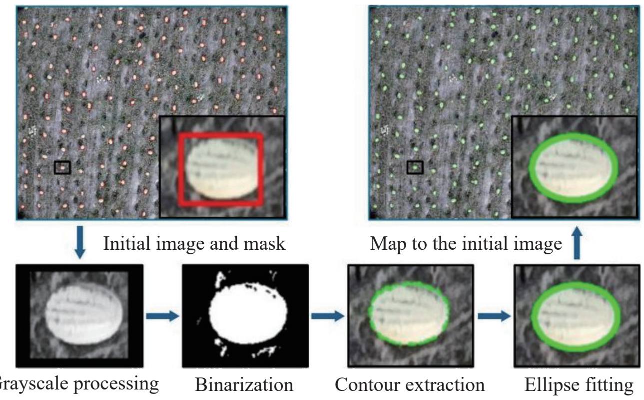


Figure 3 Pipeline of contour extraction and ellipse fitting

2.5 Influence of watermelon pose on weight estimation

A systematic investigation was implemented to explore the impact of the pose and altitude of Xisha watermelons on the yield estimation. There was great variation in the poses and elevations of Xisha watermelons in the field, due to the differences in altitude across various regions of the plot. The undulating surfaces were also found in the planting modes, such as drip irrigation[23] and ridge cultivation[24]. The altitude and tilt angles were then measured for the marked 120 Xisha watermelons.

In altitude, the Centimeter-accuracy handheld Real-Time Kinematic (RTK) positioning was used to measure the altitude data of the top of each marked Xisha watermelon. Specifically, the take-off points of the drone shared an altitude of 1654.72 m, with a flight height of 12.00 m relative to this point. Figure 4 shows the distribution of the data with the altitudes ranging from 1653.90 to 1655.40 m, and an average of 1654.86 m. The average difference in GSD was 0.038 mm after calculation. The GSD error was only 1.12% at an aerial photography altitude of 12 m, indicating a minor impact.

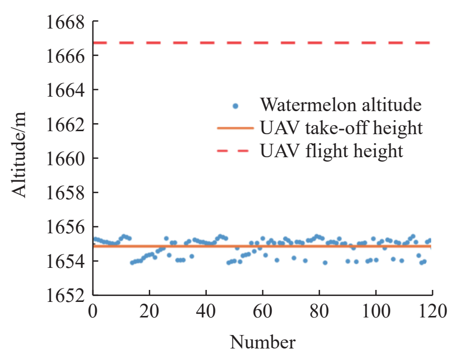


Figure 4 Altitude of individual Xisha watermelons in field after UAV measurement

Various poses were found in the naturally grown Xisha watermelons in the field. The volume estimation relied mainly on the projected lengths using images. A tilted pose also reduced the length under orthographic projection. As shown in Figure 5, the solid-line ellipse represents the pose of a watermelon in the field, with its long axis forming an angle α with the horizontal line, called the tilt angle. The arrow represents the orthographic perspective from which the drone captured the image, while the yellow dashed line is the projected length of the long axis when the watermelon is tilted. The blue line represents the actual long axis length of the watermelon. The difference between the blue and yellow dashed line is set as the projection error of length that is caused by the tilt.

As indicated in the red rectangle in Figure 5, the length L of the minimum bounding rectangle around the solid-line ellipse corresponded to the projected length of the long axis. The length was rapidly obtained using “cv2.boundingRect (points)”, indicating the non-rotated, horizontal bounding rectangle. The tilt angles of the major and minor axes were measured and then recorded for the 120 marked Xisha watermelons using a digital inclinometer with an accuracy of 0.10°. Among them, 98.5% of the watermelons shared the minor axis tilt angle of less than 16.00°. More than 91.5% showed a major axis tilt angle of less than 16.00°. The average tilt angles of the minor and major axis were 5.63° and 5.60°, respectively.

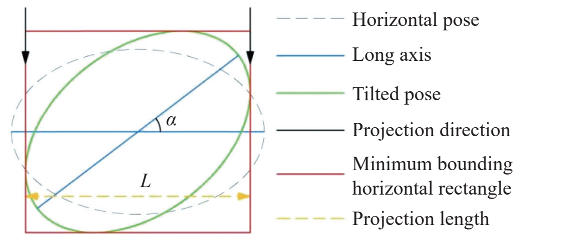
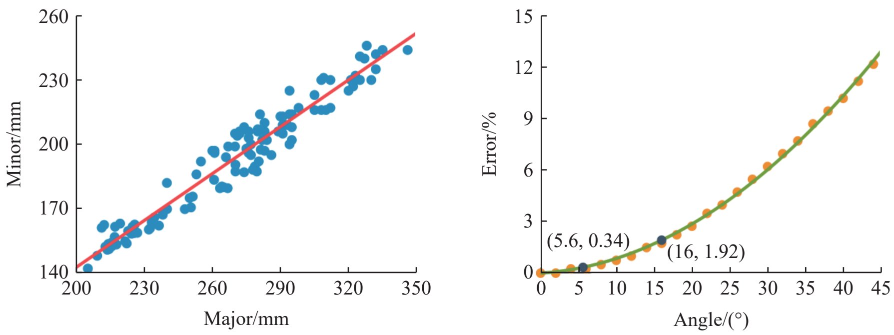


Figure 5 Projection error of major axis in the tilt angle of Xisha watermelons

Figure 6 presents the relationship between the major and minor axis lengths of Xisha watermelons and the corresponding projection error of the ellipse model under different tilt angles. Among them, Figure 6a illustrates the collected data for the major and minor axes of the Xisha watermelons. The ratio of the minor to major axis length was nearly linear. The minor axis was approximately 0.72 times the length of the major axis. The goodness of fit reached 0.9.



a. Ratio of major-minor axis length

b. Error and fitting at different angles

Figure 6 Relationship between the major and minor axis lengths of Xisha watermelons and the corresponding major axis projection error of the ellipse model under different tilt angles

A test ellipse was constructed with a minor-to-major axis ratio of 0.72. The midpoint of the ellipse’s major axis was selected as the origin of rotation, and the angle between the major axis and the horizontal line as the tilt angle α. The ellipse was rotated within a range from 0° to 90°. The horizontal bounding rectangle of the ellipse was drawn in each counterclockwise rotation by an inclination angular unit. The horizontal length of the rectangle was recorded to calculate the error value for that angle. Figure 6b shows the relationship between the error value and the rotation tilt angle. The quadratic fitting demonstrated that the error values were 0.34% and 1.92%, respectively, when the tilt angle reached the average field tilt angle of 5.6° and 16.0°, respectively. Therefore, the tilt angle of the Xisha watermelons in the field shared a relatively small impact on the weight estimation error using image analysis.

3 Yield estimation model

3.1 Yield estimation by sampling

The sample pictures were executed on yield estimation, as illustrated in the left red area of Figure 7. Firstly, a number of sample images were randomly collected from the target plot using drones. Secondly, the YOLOv8n object detection model with trained weights was used to identify the targets in the sample images after data collection. The coordinates of pixels were stored for all bounding boxes. Thirdly, ellipse fitting was performed on the pixel coordinates of the detection boxes. The sub-images with the watermelons were extracted from the much larger image. A series of operations was then conducted on the images, such as the grayscale processing, binarization, contour extraction, and ellipse fitting. The elliptical feature of the watermelons was obtained after operations. Fourthly, the volume-to-weight linear relationship was derived from the weight estimation model of individual Xisha watermelons. The volume was then applied to estimate the weight. The target of watermelon was continued to detect the next after recording the data. The estimation steps were repeated until all targets were processed in one image. Once the information was fully extracted from one sample image, the procedure was moved to the next image until completion. Finally, the average yield of the sampled area was calculated as the unit yield. The planting area of the target plot was then combined to estimate the total yield of the plot.

Figure 8 shows the typical sample images of watermelons with the dark areas and boundary during the estimation of average yield. Serious distortion and unclear boundary were also found around the edges of sample images and dark corners. A boundary box of yield estimation was defined to minimize the impact of these factors on the yield estimation. The sample images were drawn as indicated by the red rectangular box in Figure 8. The areas of yield estimation were then extracted in the boundary box from multiple sample images. The yield of watermelons was also estimated in this area. The ratio of the yield to the land area served as the unit yield data. Meanwhile, some watermelons were partially displayed near the boundary box of the estimation area, as shown in the red fitted ellipse in the magnified detail of Figure 8. The watermelons then failed to be detected beyond the weight estimation from the images. These partially visible watermelons led to underestimation of the unit yield. Typically, each image contained 2-15 partially visible watermelons. For example, Figure 8 shows that 10 watermelons were divided by the boundary box. Once these watermelons were excluded from the weight estimation, the unit yield decreased by approximately 0.232 kg/m^2. Given that the target field area was approximately 9000 m^2, the total yield was underestimated by 2088 kg. Therefore, it was necessary to consider the watermelons along the boundaries of the target areas, in order to improve the accuracy of the average yield estimation.

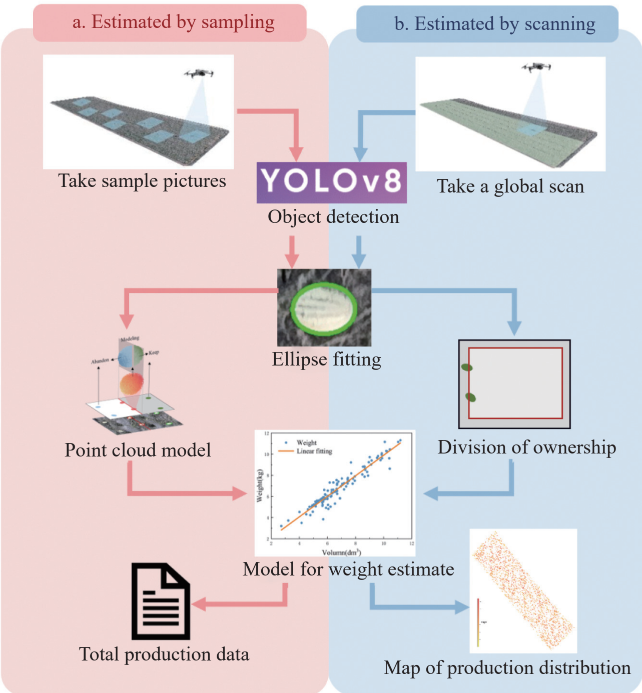


Figure 7 Constructing yield estimation model for Xisha watermelons using sampling and global scanning

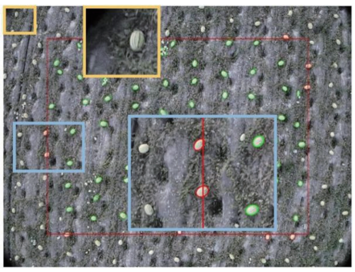


Figure 8 Sample images of watermelons with dark areas and boundary

Specifically, only the portion of the watermelon retained within the boundary box was used for weight estimation. The estimation model assumed that the watermelon was a uniformly dense ellipsoid, and a three-dimensional point cloud model was established to enable fast and accurate weight calculation, while also obtaining the partial volume of the watermelon within the yield estimation area. Compared to the traditional geometric method for calculating the partial volume of an ellipsoid, this approach is more straightforward, and compared to the two-dimensional area-based estimation model, it offers higher accuracy.

Ellipse fitting was used to obtain and then record the major and minor axes and positions of each watermelon after target detection. Then, there was the positional relationship between the detected targets and the boundary lines of areas. As shown in Figure 9, the gray cutting plane represents the map of the boundary box of yield estimation. Among them, the entire red three-dimensional point cloud of the ellipsoid model was divided into the green part on the right and the blue one on the left. The region of interest (ROI) was located in the green part of the estimation area. The red boundary line represents the rectangular box in Figure 8, where the right side is the region of yield estimation. Watermelon targets outside the ROI were discarded directly in the blue ellipse on the left in Figure 9. The green ellipse on the right is retained to record their weight data. The watermelon along the boundary represented by the red ellipse is shown in Figure 9. A three-dimensional point cloud model was then generated to map the red point cloud ellipsoid. The volume ratio was calculated as the number of points in the green and red parts of the point cloud three-dimensional array in the portion of the watermelon. The weight of ellipsoid in the ROI was then calculated using the Xisha watermelon volume-weight model.

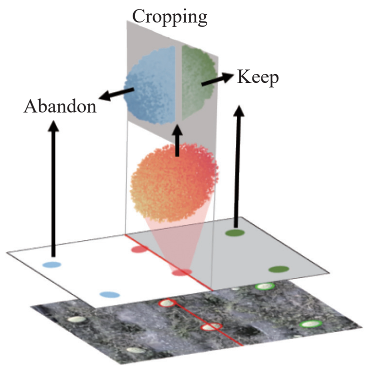
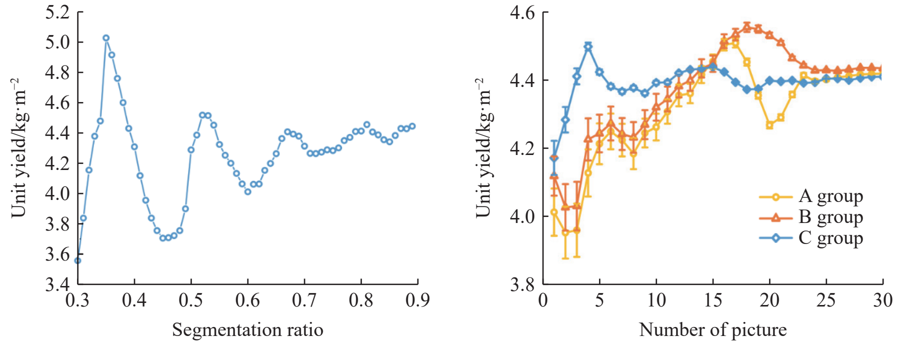


Figure 9 Schematic diagram of segmentation in 3D point cloud model

In a single image, the larger the boundary box of yield estimation was, the more useful information it contained. Thereby, the number of images was reduced to fully meet the requirements of yield estimation. Conversely, there were greater fluctuations in the average yield per image, if the estimation area was too small. It was very necessary to increase the sample size for the high accuracy of yield estimates. Nevertheless, the increasing number of sample images also led to longer sampling times, indicating the substantial consumption of time cost. However, the boundary box of yield estimation failed to enlarge the single image, due to image distortion and vignetting. Therefore, the cropping ratio of images was determined to acquire the yield estimation information.

The proportion of the boundary box was determined for a single image. One hundred sample images were selected to calculate the unit yield for each image under different segmentation ratios. The center of the image remained constant. The proportion of the boundary box ranged from 0.3x to 0.9x in the original image. An increment of 0.02 was also used in the proportion range for the unit yield. The average unit yield across all images was then computed for each segmentation ratio.

As shown in Figure 10a, the fluctuation of average unit yield data decreased as the segmentation ratio increased. The stable value after a critical point was reached before starting to fluctuate again, when the segmentation ratio exceeded 0.75, due to the distortions caused by the expansion of wide-angle parts of the image. A segmentation ratio of 0.75 was obtained, where the width and height of the yield estimation boundary box were 75% of the entire image.



a. Unit yield at different segmentation ratios b. Yield error of images in A, B and C groups

Figure 10 Unit yield error analysis

The number of sampling images was further identified to fully meet the requirements of sampling mode. Some assumptions were defined to improve the fairness and accuracy of the yield estimation. Specifically, there was no overlapping content among images during sampling. The ROI area was also required to reduce the empty fields or invalid walkways. The number of sampling images was then determined for yield estimation. In this case, 90 images were randomly selected and then divided into three groups for testing.

Each original image was segmented at the ratio of 0.75. The images were then numbered from 1 to 30 in each group. The unit yield was also calculated for the first n (n=1, 2, 3, …, 30) images, together with the relative errors between the unit yield and the overall average one. As shown in Figure 10b, the x-axis represents the number of images n, while the y-axis is the average unit yield from n images. Three lines represent the data from the A, B, and C groups, for each of which the scatter point indicates the error value relative to the overall average unit yield. It was found that the error value fell below 1.1%, 0.7%, and 1.1% from 24, 20, and 24 images, respectively, in A, B, and C groups. An additional 20.0% redundancy was applied to ensure the stable estimation during sampling, indicating at least 30 images.

3.2 Yield estimation by global scanning

The global scanning yield estimation is illustrated in the blue section of Figure 7b. 1) Data collection and preprocessing. The automatic shooting function of drone flight path was used to obtain the set of images of the target plot. These images were then processed using DJI Terra. A two-dimensional model was also established in the entire farmland, namely the Digital Orthophoto Map (DOM) of the plot. Python scripts were utilized to segment the DOM into sub-images with the same coordinates relative to the entire plot. 2) Object detection. The YOLOv8n model with trained weights was applied to identify the targets within the sub-images. The pixel coordinates were also stored for all bounding boxes. 3) Ellipse fitting. The elliptical feature of watermelons was extracted after object detection. 4) The volume-weight model was used to calculate the weight of each watermelon. The distribution of watermelon yield was then mapped, according to their position, weight, and elliptical features. The key information was obtained, such as the total number and the total yield of watermelons.

In global scanning, the DOM was cropped to fully utilize GPU memory and improve estimation efficiency. The size of image cropping was set as 5000×4000 pixels, close to the shot size of the pan-tilt-zoom drone camera. The reason was that the YOLOv8 model was used to train the original images, and then directly capture them as the dataset. The object detection of Xisha watermelons was readily applicable for both sampling and global scanning. The applicability of the detection model was maintained after segmentation.

When cropping the DOM of the plot into sub-images, it is usually necessary to set buffer zones to reduce duplicate counting or omissions of targets along the cropping lines, which can lead to errors in counting and weight estimation. Previous studies either ignored the fragments[24] or retained a 50-pixel buffer zone[25]. Wang et al.[26] proposed to add the buffer zones into the right and bottom edges of images during cropping and merging duplicate detections. Jiang et al.[20] also adopted the buffer zones around the images. This study also utilized the Intersection over Union (IoU) of object detection boxes to identify duplicate targets, ensuring the accuracy of the counting.

Figure 11 shows the schematic diagram and representative application of assignments. Among them, the watermelons were cut by cropping lines in Figure 11a. The incomplete area potentially led to missed or overcounted detections. Based on the practical harvesting experience of local watermelon growers and our field measurements, the maximum major axis length of a mature watermelon is typically within 400 mm, with only rare exceptions exceeding this value. A major axis length of 400 mm corresponds to approximately 122 pixels at the flight height of 12 m used in this study. To balance safety margin and computational cost, an empirical redundancy was introduced, and the buffer width was set to 1.2 times the maximum major axis length. After rounding, this value was approximately 150 pixels, corresponding to an actual watermelon major-axis length of about 491 mm. The red line represents the theoretical cropping line in Figure 11a, while the black line is the actual cropping line, and the gray area between them is the buffer zone. The watermelons were then intact on the cropping line. Furthermore, the specific step was also added to avoid double counting after object detection. The reason was that the watermelons on the theoretical cropping line were repeatedly detected in the buffer zones of adjacent sub-images. The center of the object was determined in the target region—the red line boundary. If the center was inside the target region, the object was assigned to the image. Otherwise, it was discarded. Figure 11a illustrates the assignment with the centers of two fitted ellipses. Ellipse ① was discarded with the center outside the target region in the buffer zone, while Ellipse ② was retained, weighed, and recorded with its center inside the target region.

Figure 11b shows the specific application. The DOM of the field was divided into the target regions, according to the theoretical cropping lines indicated by red dashed line frames. The actual division with the buffer zones is marked as the black solid lines. Two adjacent images were cropped after the division. According to the enlarged view, two watermelons were split by the red cropping line. The similar assignment of the watermelons is then followed in Figure 11a. The ellipses were also fitted for the watermelons in the ROI image. As such, two watermelons on the cropping line were assigned to the corresponding image areas in Figure 11b, according to the position of their centers.

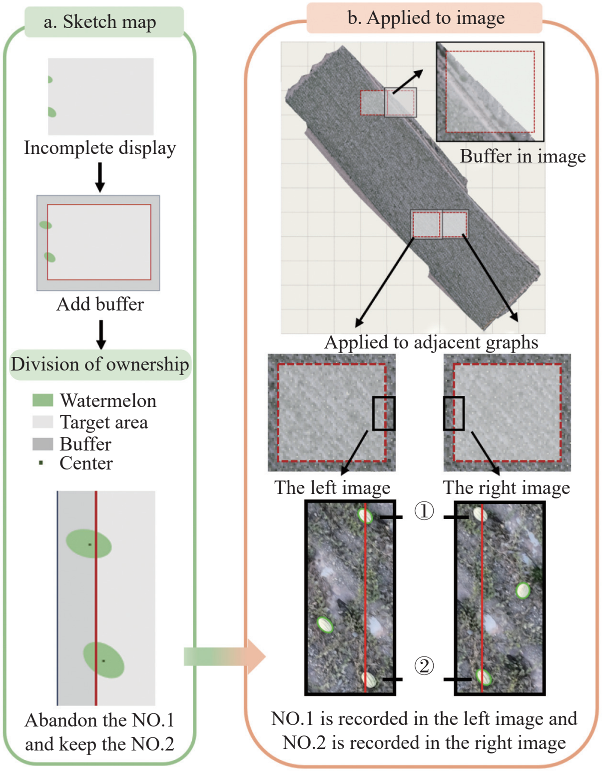


Figure 11 Schematic diagram and representative application of assignments

Manual verification was performed on all cropped images. It was found that a total of 209 watermelons were cut by the segmentation lines. Except for one watermelon missing, the rest was accurately counted without any omission or duplication. The single missed watermelon failed to be successfully detected during assignment because of its unclear appearance in the image. Anyway, the algorithm was reliable to count the watermelons cut by segmentation lines.

4 Results

Figure 12 shows the scatter plot of the volume and weight for the 120 marked Xisha watermelons. A linear fitting was also realized to obtain the following linear relationship between volume and weight:

$$W = 1.0424 \times V - 0.2456 \tag{3}$$

where, V is the volume of the Xisha watermelon calculated by Equation (1); W is the weight of the watermelon measured by an electronic scale, kg.

In the linear regression model, the goodness of fit R^2 was 0.916, indicating a strong linear relationship between the elliptical volume of the watermelon and the measured weight. This demonstrates the feasibility of estimating watermelon weight by constructing an ellipsoid model based on volume.

$$W = 4.3642 \times A^2 B - 0.2456 \tag{4}$$

where, A and B are the estimated real-world lengths of the minor and major axes of the watermelon ellipsoid (dm) respectively, obtained by multiplying the pixel dimensions of minor and major axes in the remote sensing image by the GSD.

The YOLOv8n-based object detection model trained in this study achieved the following results on the Xisha watermelon validation set: Precision (P) was 0.961, Recall (R) was 0.976, $F1$ score was 0.968, and the Average Precision (AP) was 0.986, demonstrating strong detection performance for Xisha watermelons.

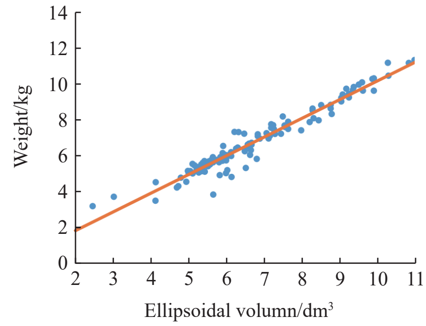


Figure 12 Linear fit relationship between the volume and weight of Xisha watermelons

Figure 13 shows the result of ellipse fitting on the watermelons using OpenCV. Green ellipses represent the fitted contours of watermelons, which were mostly well-aligned with the actual contours. However, due to factors such as light shadows, skin patterns of the watermelons, and irregular shapes of some watermelons, there were instances of overfitting and underfitting.

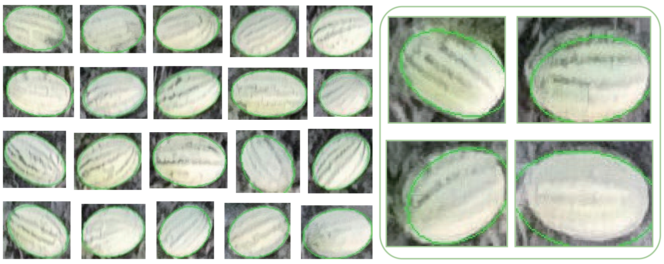


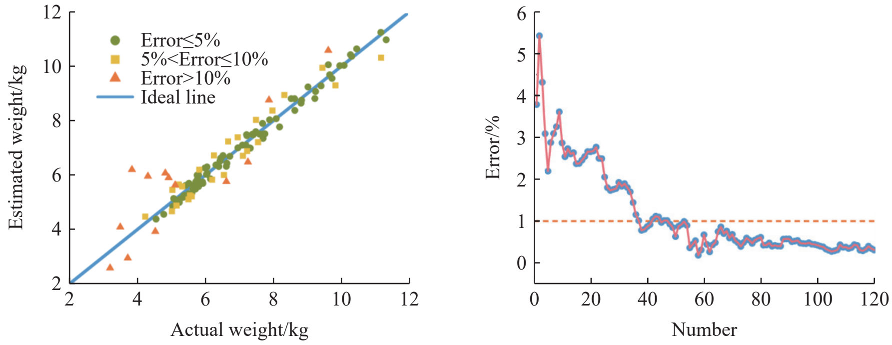
Figure 13 Good fitting (left) and inadequate fitting (right) visualization

A statistical analysis was conducted to quantify the combined impact of various factors, including the altitude, pose, and contour fitting of watermelons. The errors were also obtained for the individual watermelons and the total yield after weight estimation. The data was collected from 120 marked Xisha watermelons.

A specific procedure was utilized to estimate the weight error of individual watermelons. Firstly, the ellipse fitting algorithm was used to obtain the elliptical contours of the marked watermelons. Then, the pixel dimensions of the major and minor axes for each watermelon were calculated to convert into the actual ones using GSD, according to their shape characteristics. Finally, Equation (3) was applied into the weight of each watermelon, thus representing the algorithm-estimated weight.

Figure 14 shows the scatter plot with the errors between estimated and measured weight of the single and multiple watermelons. The x- and y-axes represent the measured and estimated weight, as shown in Figure 14a. The blue solid line represents the ideal scenario, where the estimated weight is equal to the measured one.

Different colored points indicate the varying ranges of errors between the estimated and actual weights. Orange triangles represent errors greater than 10.00%, accounting for approximately 10.83% of the total sample size. Yellow squares indicate errors between 5.00% and 10.00%, accounting for about 20.00% of the total, and green dots represent errors of 5.00% or less. Overall, the average estimation error of individual watermelon weight was 5.31%. Further analysis by fruit-weight categories showed that small watermelons (<5 kg) exhibited much larger estimation errors, with an average error of 20.88%, whereas medium (5-9 kg) and large (>9 kg) watermelons showed smaller errors, with average errors of 3.89% and 3.00%, respectively. It is worth noting that small watermelons are generally less likely to enter the market; therefore, their larger estimation errors have limited impact on the overall yield or economic value assessment.



a. A single watermelon b. Multiple watermelons

Figure 14 Error between estimated and measured weight

The cumulative error of weight estimation was calculated for the multiple watermelons. Firstly, the total estimated and measured weight were calculated for the first i (where i=1, 2, …, 120) watermelons out of the 120 marked Xisha watermelons. Then, the absolute value of the relative error between the total weights was computed using Equation (5). Figure 14b shows the line plot with the number of watermelons i on the x-axis and the total weight error on the y-axis. The relative error between the total estimated and the measured weight decreased significantly, as the number of watermelons increased. Eventually, the reasonable range of 1% was achieved after the number of watermelons reached 45.

$$E_i = \frac{\left|W_{e_i} - W_{a_i}\right|}{W_{a_i}} \times 100\% \tag{5}$$

where, E_i is the error between the algorithm-estimated total weight and the measured total weight of i watermelons, %; W_{e_i} is the algorithm-estimated total weight of i watermelons, kg; W_{a_i} is the actual measured total weight of i watermelons, kg.

In summary, the total yield error of watermelons consisted of the detection and weight estimation error. Among them, the detection error was caused by the watermelon counting after object detection. The YOLOv8n algorithm achieved an average precision of 0.986. Alternatively, the error of weight resulted from the cumulative estimation on the weights of individual watermelons, due mainly to the altitude, pose, and contour fitting.

The total weight error was within 1.00% from the 120 marked watermelons, as the number of watermelons increased. Although the average error of weight estimation on the individual watermelons was 5.31%, the over- or under-estimation were caused by overfitting or underfitting in the individual watermelons, as the number of targets increased. Thereby, the overall yield error reduced the offset of over- or under-estimation.

Consequently, the total yield error was expressed as Equation (6):

$$E_{\text{total}} = 1 - P \times (1 - E_i) \tag{6}$$

where, P is the average precision of the detection algorithm; E_i is the relative error between the total estimated and actual weight. The total estimation error E_{total} was less than 3.00%, indicating the feasibility of the detection and fitting algorithm on the yield estimation of Xisha watermelons.

5 Discussion

Two tables were output after data processing in the same field of Xisha watermelons using both sampling and global scanning. Thirty sample images were also selected from the aerial waypoints of the target field after sampling. One table documented the fundamental information for each watermelon after identification, including its position relative to the top-left corner of the image, the major and minor axes, angle, relationship with the cropping boundary, and estimated weight. Another table was used to record the statistical indicators, such as the number of watermelon targets per image, total estimated weight, average yield, and cumulative average total yield. After that, the unit yield was calculated as 4.42 kg/m^2 using sampling. The total estimated yield was 39 803.4 kg, given that the effective planting area of the field was approximately 9000 m^2.

In global scanning, all 796 images were captured along the flight paths of the target plot. The data table and yield distribution map were then obtained after global scanning. The table contained the fundamental indicators for all watermelon targets, including their position coordinates (center X, and center Y) relative to the top-left corner of the modeling image, shape features of the fitted ellipse (major axis, minor axis, and angle), and estimated weight. Additional statistics were also provided with the total number of watermelons, total estimated yield, and average weight per watermelon.

Figure 15 shows the yield distribution of the target field after estimation using global scanning. The table data was also obtained after processing. 4785 watermelons were found in the field with the total estimated yield of 37 873.65 kg. The average weight per watermelon was 7.96 kg. Each ellipse represents a watermelon in Figure 15, accurately reflecting its position, size, and orientation in the field. The color of the ellipses is determined by the weight attribute of the watermelon. The gradient is indicated from yellow to red in the weight range. Among them, the red colors represent the heavier watermelons.

A comparison was also made between the sampling and global scanning during yield estimation across multiple dimensions, as summarized in Table 1. These include the total time for each estimation, the data processing speed per unit area, and the final output content. Additionally, there are the relative differences in the number of watermelons and the total yield.

There was little difference in the total estimated yield of Xisha watermelons using sampling and global scanning, in terms of quantity and weight. However, the data processing of sampling was approximately 50 times faster than that of the global scanning. Additional expenditure was also reduced on commercial two-dimensional modeling software. Meanwhile, global scanning provided finer indicators on the distribution of watermelon yields.

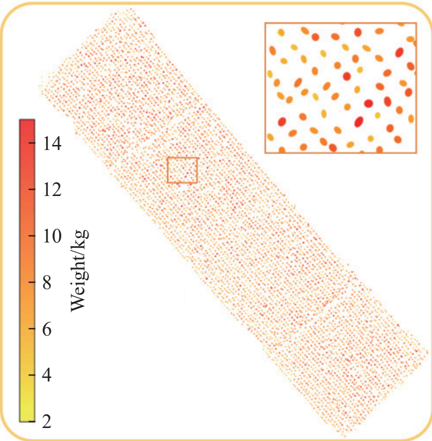


Figure 15 Yield distribution map after estimation using global scanning

Table 1 Comparison on sampling and global scanning after yield estimation

Items	Sampling	Global scanning	Notes
Modeling required (Yes/No)	No	Yes	Modeling required DJI Terra
Shooting duration	1-2 min	10 min	The total duration of the global scanning method was approximately 277 min.
Modeling duration	0	103 min 40.0 s	
Processing duration	3 min 28.8 s	163 min 40.0 s	
Overall processing speed	27.37 m²/s	0.541 m²/s	Approximately 50 times longer
Count of watermelon	4842	4758	Relative error of 1.77%
Unit yield	4.42 kg/m²	4.21 kg/m²	The field area was 9000 m²
Yield estimation	39 803.40 kg	37 873.65 kg	Relative error of 5.10%
Output content	Estimated total yield, Estimated number of watermelons, Watermelon indicators in sample images	Estimated total yield, A harvest report for each watermelon	

Therefore, sampling offered the higher and more efficient approach to obtain the total yield and number of watermelons in the target area, particularly for the harvesting, transportation, and decision-making on the multiple plots with the large areas of Xisha watermelons. In contrast, global scanning provided the comprehensive yield distribution to evaluate the different regions, periods, and plots. Agronomic practices can be expected to enhance watermelon yields with the benefit of such a comparison. Furthermore, the yield distribution visually presents all watermelons in the field, indicating a great contribution to the critical tasks, such as path planning of harvesting robots, in the future. As such, sampling and global scanning can be selected for yield estimation, according to the different needs in the agricultural applications.

6 Conclusions

For the task of yield estimation of Xisha watermelons in the field, a weight estimation model for individual watermelons was first established. This model estimates the weight of each watermelon based on the pixel lengths of the major and minor axes of ellipses in images. Subsequently, a recognition and ellipse-fitting algorithm was proposed. On the basis of watermelon detection using the YOLOv8 object detection algorithm, contour fitting was performed using OpenCV library functions to extract elliptical features. Finally, a comprehensive yield estimation model for Xisha watermelons was developed, incorporating two estimation methods. The first is a sampling method, which utilizes a 3D point cloud model to solve the problem of calculating the volume of sliced ellipsoids. This method is applied to sample images to obtain the average yield per unit area in the sampling region, and the total yield is then calculated based on the actual field area. The second is a global estimation method, in which a new buffer zone approach is proposed to address the issue of watermelons being split during image cropping, thus improving the accuracy of counting. By determining the center position of each fitted object relative to the cropping line, the method identifies the region to which each object belongs, thereby acquiring detailed information on the position, size, and weight of all watermelons within the plot.

Test results showed that the YOLOv8n-based Xisha watermelon detection model achieved an average detection precision of 0.986. The average weight estimation error for individual watermelons was approximately 5.31%, and the total yield estimation error for multiple watermelons was less than 3.00%. Comparative tests using both estimation methods on the same Xisha watermelon plot demonstrated that the data processing speed of the sampling-based method was 50 times faster than that of the global method. The difference in total yield estimation results between the two methods was around 5.10%, and the difference in counting results was 1.77%. If only the total number and total weight of Xisha watermelons in a planting plot are required, the efficient and rapid sampling method is recommended. If the spatial distribution of yield across the entire plot is also needed, the global estimation method should be adopted to obtain more detailed information. In summary, both estimation methods in the Xisha watermelon yield estimation model have their own advantages and should be selectively used according to different application scenarios.

In the future, the feasibility of the proposed model will be further validated under challenging conditions, such as varying illumination for watermelon detection and yield estimation and fields with large elevation differences (e.g., hilly areas), and the combined yield estimation model will be applied to multiple plots to further evaluate its stability. In addition, it is also planned to apply the model to other watermelon varieties and other crops such as cauliflower and Hami melons, with the aim of providing more practical yield estimation approaches and addressing yield estimation challenges in agricultural applications.

Acknowledgements

This study was financially supported by China's Central Guiding Local Technology Development Special Fund Project (Grant No. ZYYD2025CG21), the Zhejiang Science and Technology Plan Project-National Key Research and Development Program of China (Grant No. 2019C03075), "Leading Goose" Research and Development Program Subproject of Zhejiang (Grant

No. 2022C02021), and Xinjiang Boshiran Intelligent Agricultural Machinery Co., Ltd (Grant No. KYY-HX-20210639).

[References]

[1] Adesipo A, Fadeyi O, Kuca K, Krejcar O, Maresova P, Selamat A, et al. Smart and climate-smart agricultural trends as core aspects of smart village functions. Sensors, 2020; 20(21): 5977.

[2] Maddikunta P K R, Hakak S, Alazab M, Bhattacharya S, Gadekallu T R, Khan W Z, et al. Unmanned aerial vehicles in smart agriculture: Applications, requirements, and challenges. Sensors, 2021; 21(16): 17608–17619.

[3] Gatkal N R, Nalawade S M, Shelke M S, Sahni R K, Walunj A A, Kadam P B, et al. Review of cutting-edge weed management strategy in agricultural systems. Int J Agric & Biol Eng, 2025; 18(1): 25–42.

[4] Qi J, Li M, Zhang H, Zeng T. Detection of the yellow-leaf disease of rubber trees using low-altitude digital imagery from UAV. Int J Agric & Biol Eng, 2024; 17(6): 245–255.

[5] Cheng Q W, Wu B S, Ye H C, Liang Y Y, Che Y P, Guo A T, et al. Inversion of maize leaf nitrogen using UAV hyperspectral imagery in breeding fields. Int J Agric & Biol Eng, 2024; 17(3): 144–155.

[6] Xu X B, Teng C, Zhu H C, Feng H K, Zhao Y, Li Z H. Comparison of three models for winter wheat yield prediction based on UAV hyperspectral images. Int J Agric & Biol Eng, 2024; 17(2): 260–267.

[7] Rejeb A, Abdollahi A, Rejeb K, Treiblmaier H. Drones in agriculture: A review and bibliometric analysis. Computers and Electronics in Agriculture, 2022; 198: 107017.

[8] Leukel J, Zimpel T, Stumpe C. Machine learning technology for early prediction of grain yield at the field scale: A systematic review. Computers and Electronics in Agriculture, 2023; 207: 107721.

[9] Shen T T, Zhang X, Li L, Qi Y X, Ji H F, Yang G P, Zhang X X. Dynamic changes in rhizosphere microbial communities of watermelon during continuous monocropping with gravel mulch. Journal of Soil Science and Plant Nutrition, 2024; 24(1): 775–790.

[10] Farooq M R, Zhang Z Z, Yuan L X, Liu X D, Rehman A, Bañuelos G S, et al. Influencing factors on bioavailability and spatial distribution of soil selenium in dry semi-arid area. Agriculture, 2023; 13(3): 576.

[11] Koirala A, Walsh K B, Wang Z, McCarthy C. Deep learning for real-time fruit detection and orchard fruit load estimation: Benchmarking of 'MangoYOLO'. Precision Agriculture, 2019; 20(6): 1107–1135.

[12] Huynh T T M, TonThat L, Dao S V T. A vision-based method to estimate volume and mass of fruit/vegetable: Case study of sweet potato. International Journal of Food Properties, 2022; 25(1): 717–732.

[13] Sabouri A, Bakhshipour A, Poorsalehi M, Abouzari A. Machine learning techniques for non-destructive estimation of plum fruit weight. Scientific Reports, 2025; 15(1): 751.

[14] Kalantar A, Dashuta A, Edan Y, Dafna A, Gur A, Klapp I. Estimating melon yield for breeding processes by machine-vision processing of UAV images. Precision Agriculture '19. Wageningen Academic, 2019; pp.381–387. doi: 10.3920/978-90-8686-888-9_47

[15] Kalantar A, Edan Y, Gur A, Klapp I. A deep learning system for single and overall weight estimation of melons using unmanned aerial vehicle images. Computers and Electronics in Agriculture, 2020; 178: 105748.

[16] Wittstruck L, Kühling I, Trautz D, Kohlbrecher M, Jarmer T. UAV-based RGB imagery for Hokkaido pumpkin (*Cucurbita max.*) detection and yield estimation. Sensors, 2020; 21(1): 118.

[17] Ekiz A, Arıca S, Bozdogan A M. Classification and segmentation of watermelon in images obtained by unmanned aerial vehicle. International Conference on Electrical and Electronics Engineering (ELECO), IEEE, 2019; 619–622. doi: 10.23919/ELECO47770.2019.8990605

[18] Ekiz A, Arıca S. Detection of watermelon in RGB images via unmanned aerial vehicle by utilizing texture features for predicting yield. Pakistan Journal of Agricultural Sciences, 2022; 59(6): 873.

[19] Qiu J K, Xu X Y, Kang Y, Zang H, Ma K, Guo Z P. Research on watermelon fruit extraction from UAV images based on semantic segmentation. Journal of Chinese Agricultural Mechanization, 2024; 45(3): 182–188. (in Chinese)

[20] Jiang L G, Jiang H H, Jing X D, Dang H J, Li R, Chen J Y, et al. UAV-based field watermelon detection and counting using YOLOv8s with image panorama stitching and overlap partitioning. Artificial Intelligence in Agriculture, 2024; 13: 117–127.

[21] Yin H J, Wang B L, Jing Y G, Li J X, Wang P L, Quan G X, et al. Improved YOLOv7 method for counting watermelons in UAV aerial videos. Transactions of the CSAE, 2024; 40(19): 124–134. (in Chinese)

[22] Koc A B. Determination of watermelon volume using ellipsoid approximation and image processing. Postharvest Biology and Technology, 2007; 45(3): 366–371.

[23] Geng Y H, Cao G J, Wang L C, Wang M, Huang J X. Can drip irrigation under mulch be replaced with shallow - buried drip irrigation in spring maize production systems in semiarid areas of northern China? Journal of the Science of Food and Agriculture, 2021; 101(5): 1926–1934.

[24] Chen Y, Lee W S, Gan H, Peres N, Fraisse C, Zhang Y C, He Y. Strawberry yield prediction based on a deep neural network using high-resolution aerial orthoimages. Remote Sensing, 2019; 11(13): 1584.

[25] Lee C J, Yang M D, Tseng H H, Hsu Y C, Sung Y, Chen W L. Single-plant broccoli growth monitoring using deep learning with UAV imagery. Computers and Electronics in Agriculture, 2023; 207: 107739.

[26] Wang H Z, Li T, Nishida E, Kato Y, Fukano Y, Guo W. Drone-based harvest data prediction can reduce on-farm food loss and improve farmer income. Plant Phenomics, 2023; 5: 0086.