



Multi-phenotypic parameters extraction and biomass estimation for lettuce based on point clouds

Yu Zhang ^{a,d,1}, Maowei Li ^{a,d,1}, GuiXin Li ^{a,d}, Jinsong Li ^{a,d}, Lihua Zheng ^{a,b,d}, Man Zhang ^{a,b,d}, Minjuan Wang ^{a,b,c,*}

^a Key Laboratory of Smart Agriculture Systems, Ministry of Education, China Agricultural University, Beijing 100083, China

^b Yantai Institute of China Agricultural University, Yantai 264670, China

^c State Key Laboratory of Virtual Reality Technology and Systems, Beihang University, Beijing 100191, China

^d College of Information and Electrical Engineering, China Agricultural University, Beijing 100083, China

ARTICLE INFO

Keywords:

Plant phenotypic parameter
Point cloud
Lettuce
Biomass Estimation

ABSTRACT

Traditional phenotypic measurements were destructive, time-consuming and laborious. 3D computer vision technology collects plant information in a non-destructive manner and can work continuously. The purpose of this paper is to measure lettuce multi-phenotypic parameters and estimate biomass by point clouds. The result shows: (1) RGB information improved the segmentation effect of the point cloud. (2) The point cloud segmentation method proposed in this paper has a high accuracy (accuracy = 98.7 %) and keeps high spatial resolution. (3) On the test set, the R^2 of height, diameter, leaf area, fresh weight and dry weight are 0.935, 0.905, 0.969, 0.966 and 0.968, respectively, the cumulative NMSE is as low as 0.069. The Totalscore is 15 % better than the 1st of the online challenge held based on this data set. The results showed that the plant phenotypic extraction method proposed in this paper was accurate and efficient, and could be applied for high-throughput phenotyping.

1. Introduction

Lettuce is the most consumed and the most cultivated leafy vegetable worldwide with a large number of nutrient elements such as carotene, vitamin C and vitamin E [1]. Different horticultural types of lettuce exhibit tremendous morphological variation [2]. In recent years, the demand for lettuce has been increasing. How to grow lettuce scientifically and healthily has attracted widespread attention from researchers. The phenotype parameter of the plant is the information reflected by the crop during the growth process, including the color, size and shape of different organs, such as stems, leaves, coronary layers, roots, etc [3]. These phenotype parameters can reflect the growth level and health status of plants, which help producers to manage and make decisions. Traditional plant phenotype measurements such as physical observation and naked eye observation, are time-consuming, labour intensive, highly subjective and sometimes destructive [4]. Compared with traditional measurement methods, computer vision technologies collect plant information in a non-destructive, low-cost, and efficient way, which can work continuously. The 2D image analysis has good potential

in terms of quantified crop information and has been widely used to measure the basic shape characteristics, including plant height, leaf length and leaf width [5–8]. Three-dimensional point clouds refer to a set of points defined in 3D measurement space, usually containing three-dimensional space information, RGB color information, normal vector information, etc. Compared with two-dimensional image data, the three-dimensional point cloud data is more realistic and not vulnerable to object occlusion. The three-dimensional point cloud is widely used in crop phenotypic analysis. For example, subtract point cloud height coordinates to get crop height [9–11], calculate the surface area of mesh which is converted from point clouds to get leaf area [12,13], and predict the yield of crop [14,15].

In the process of calculating phenotypic parameters based on point clouds, the primary task is to segment crop point clouds clearly. The traditional point cloud segmentation mainly uses the color, coordinate, shape and other information of the point clouds to set different regional boundaries and then segment the point clouds [16–18]. It requires multiple parameters to ensure that the segmentation achieves the best results, with many restrictions. Using the deep learning model to

* Corresponding author at: Key Laboratory of Smart Agriculture Systems, Ministry of Education, China Agricultural University, Beijing 100083, China.
E-mail address: minjuan@cau.edu.cn (M. Wang).

¹ These authors contributed equally to this work and should be considered co-first authors.

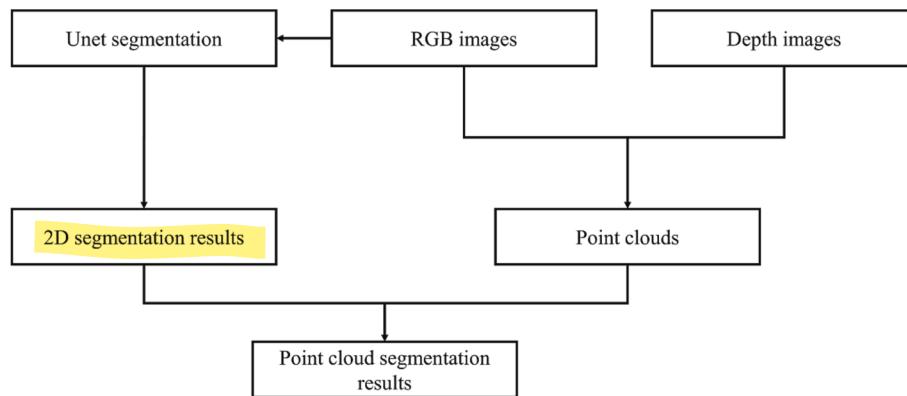


Fig. 1. Workflow diagram of the point cloud segmentation.



Fig. 2. Environment of data collection.

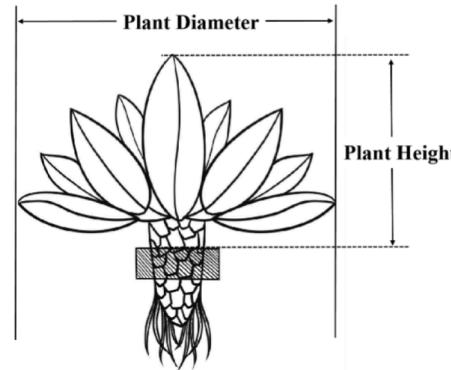


Fig. 3. Plant height and diameter of lettuce.

segment the point clouds requires the point clouds with specified resolution as the input [19–21]. The down-sampling operation or splitting up into multiple boxes will inevitably lose a lot of point cloud information. The resolution of the two-dimensional image is the same and constant. And it is easier and more mature to extract the feature information from the two-dimensional image than that of the three-dimensional point clouds in the process of model training. Shi et al. combined different viewpoints which were segmentation results on the 2D images to segment the 3D point cloud. This method can segment the point cloud of plant organs accurately and efficiently [22].

The primary goal of this paper was to estimate the major phenotypic parameters of lettuce and predict lettuce biomass with a RGB-D sensor. To achieve this goal, three specific objectives were defined and met: (1) mapping the segmentation results of the U-Net model to point clouds to get high-resolution point clouds segmentation, as shown in Fig. 1 [27]. (2) calculating the plant height and diameter of lettuce based on the high-resolution point cloud. (3) proposing prediction factors to calculate lettuce leaf area and predict the lettuce biomass.

2. Materials and methods

2.1. Data acquisition and processing

2.1.1. Data introduction

The data of this experiment is based on the lettuce dataset of the Autonomous Greenhouses International Challenge 3rd held by Wageningen University and Research Center and Tencent Lab [23]. This competition has 46 teams from 24 countries participating. The dataset contains four varieties of lettuce (Aphylion, Lugano, Salanova, Satine) grown by the company Rijk Zwaan. The data image was captured by the Intel D415 camera hanging about 0.90 m above the crop, which was shown in Fig. 2.

Table 1
Phenotypic data analysis of lettuce.

	Parameter	Min	Max	Mean	Standard deviation
Train	Height (cm)	4.300	25.000	12.301	4.744
	Diameter (cm)	9.500	42.000	22.631	6.300
	Leaf Area (cm ²)	57.600	6875.400	1849.350	1516.227
	Fresh Weight (g)	1.400	444.700	114.680	106.230
	Dry Weight (g)	0.090	32.000	5.397	4.600
	Height (cm)	5.000	25.000	12.684	5.314
Test	Diameter (cm)	8.200	37.000	23.030	6.229
	Leaf Area (cm ²)	67.200	5868.000	1848.246	1590.428
	Fresh Weight (g)	2.100	459.700	126.658	127.417
	Dry Weight (g)	0.100	20.100	5.952	5.481

The data collection period is from March 5, 2021 to April 16, 2021. About 70 plants were taken out from the experiment every week, and destructive measurement was carried out after photographing to obtain the relevant ground truth data of lettuce plant height, diameter, leaf area, fresh weight and dry weight. Plant height and diameter are shown in Fig. 3.

According to the competition situation, this experiment divided all the data sets into two categories: 341 training sets and 50 final test sets, including the data of each variety in each growth period. The final score of the competition depends on the prediction accuracy of the test sets. The analysis of train data and test data is shown in Table 1.

2.1.2. Data annotation and augmentation

LabelMe is used to annotate the outline of lettuce to create the dataset of two-dimensional image segmentation [24]. *Augmentor* is a pipelined

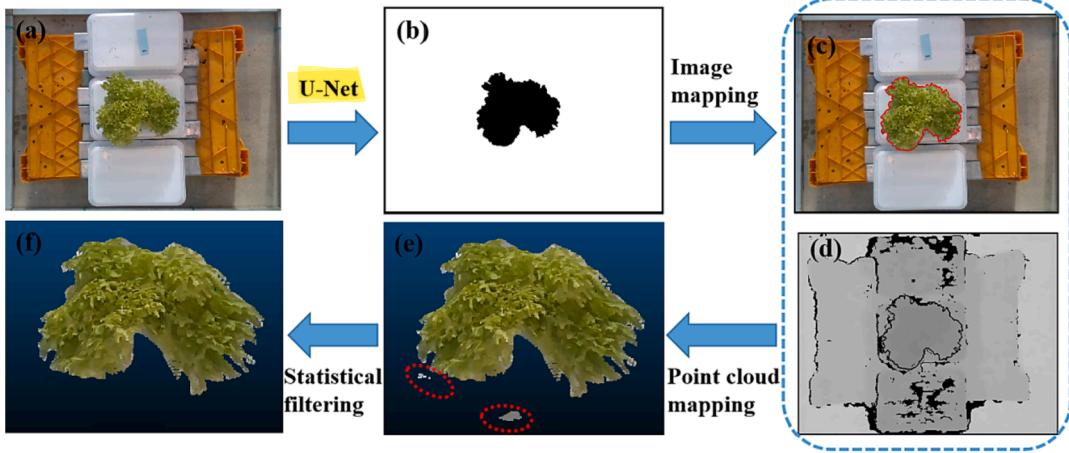


Fig. 4. The point cloud segment method based on image segmentation result of the U-Net mapping. (a): The RGB image of lettuce. (b): The segment result of U-Net. (c), (d), (e): The mapping results of RGB image, depth image and point cloud, respectively. (f): The segment result of statistical filtering. Note: Red circles in (e) representing the edge noise.

2D image stochastic augmentation library, which can enhance the original image and labeled mask at the same time[25]. In this paper, six data enhancement methods in *Augmentor*, including zoom, rotation, deformation, random erasure and brightness change, are randomly combined for data enhancement. The RGB images of 341 training sets are expanded to 3032 (including the original images), thus forming a data enhancement set.

Extracting the depth information in the **depth image** and **RGB** in the RGB image and combining them with the internal parameters of the camera can get the color 3D point cloud. The color point cloud data is annotated using the 3D point cloud annotation tool *CloudCompare* for subsequent segmentation network training[26].

2.2. Point cloud segmentation of lettuce

In this paper, the RGB images of lettuce were segmented by the U-Net model [27], and then the result of segmentation was mapped to the corresponding point cloud. First, using the U-Net network to segment the lettuce image, as shown in Fig. 4a. The segmentation results are shown in Fig. 4b. Then, mapping the segmentation results in the RGB images and the depth images. The mapping results are shown in Fig. 4c and Fig. 4d. Finally, the corresponding point cloud converted from RGB and depth images are shown in Fig. 4e.

However, since U-Net cannot accurately identify the lettuce edges, there is some error in the lettuce edges, thus there will be some edge noise in the mapping results of the point cloud, as shown by the red dashed line in Fig. 4e. Based on this problem, this paper proposes to use statistical filter to remove the edge noise mentioned above to ensure the point cloud segmentation accuracy, and the results are shown in Fig. 4f.

2.3. Evaluation of data segmentation

The segmentation effect was evaluated by Accuracy and mIoU. True positive (TP) is the number of points in which the true value is the lettuce, and the predicted value is the lettuce too. False positive (FP) is the number of points in which the true value is the lettuce, and the predicted value is the background. True negative (TN) is the number of points in which the true value is the background, and the predicted value is the background too. False negative (FN) is the number of points in which the true value is the background, and the predicted value is the lettuce.

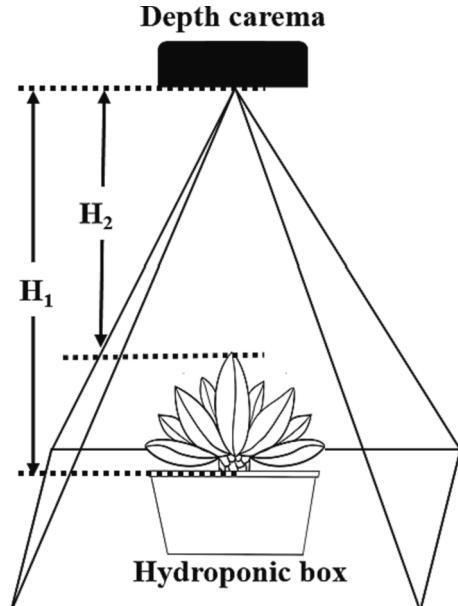


Fig. 5. The images are made with a RealSense D415, hanging about 0.900 m above the growing crop (H_1 equals 0.900 m).

2.4. Phenotypic parameter extraction and biomass prediction

2.4.1. Plant height

The plant height refers to the vertical distance between the highest point and the bottom of the leafy vegetable. Since the lettuce grows in the hydroponic box, the plane where the hydroponic box is located is set as the bottom of the lettuce. Plant height, as one of the most basic phenotypic parameters of plants, is a determinant of the light absorption capacity and it also has a high correlation with traits such as leaf area, leaf nitrogen content, and biomass, making it an essential object of study among plant traits. Since the Intel D415 camera imaging is based on the binocular infrared principle, the depth information corresponding to each pixel is stored in the depth images. By obtaining the depth value in the depth image and multiplying it by the corresponding depth coefficient, the distance value between the actual object corresponding to the pixel and the camera can be obtained. According to the above analysis, it can be known that the top point of the object corresponds to the point with the smallest depth value (closest to the camera), while the bottom

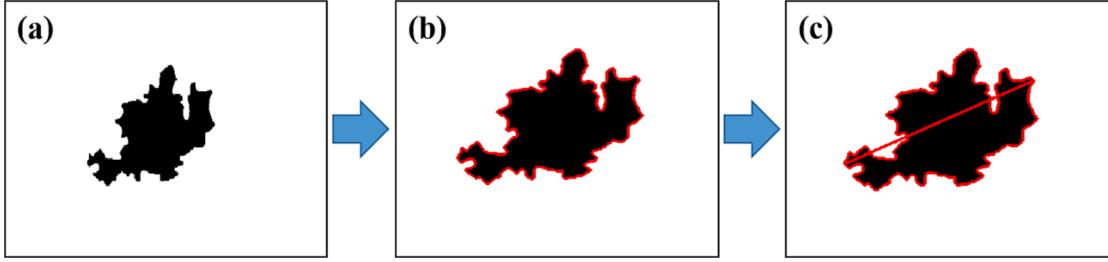


Fig. 6. Process of obtaining relative diameter. (a): The segment result of U-Net. (b): The edge pixels of segment result. (c): The farthest pixel distance from the edge pixels.

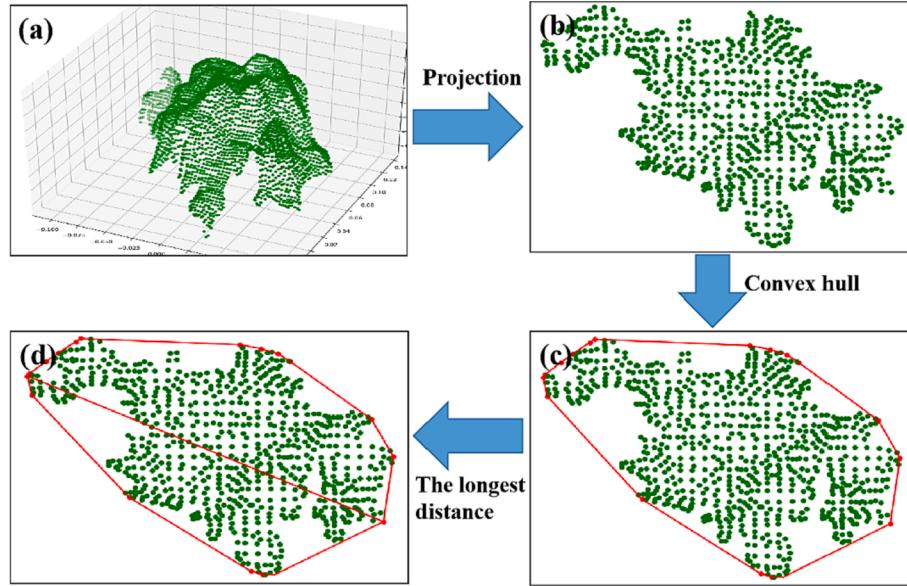


Fig. 7. Process of obtaining plant diameter. (a): The lettuce point clouds. (b): The projection result of point cloud. (c): The 2D convex hull. (d): The farthest distance from the 2D convex hull.

point is the point with the largest depth value (furthest away from the camera). The distance between the root and the top of the crop can be obtained by calculating the difference between the depth values of the two (Eq. (1)).

$$\text{Plant height} = H_1 - H_2 \quad (1)$$

The data acquisition apparatus for this experiment is shown in Fig. 5. Each lettuce plant was placed in a hydroponic box of equal size and the camera was fixed at the top position of the collection frame. The distance H_1 between the bottom of the crop and the camera was fixed with a depth value of 0.900 m, so the plant height of lettuce only depended on the value of H_2 .

2.4.2. Plant diameter

Lettuce diameter refers to the maximum distance between two leaf tips. The plant canopy diameter is one of the important indicators reflecting the extent of coverage of the plant canopy and the area receiving sunlight. It is closely related to the photosynthesis and respiration of plants [28]. This paper obtains lettuce diameter based on image segmentation results and point cloud segmentation results. By comparing the two results, a better method is used to obtain the diameter of the test set data.

(1) Diameter prediction based on image segmentation result.

According to the result segmented by U-Net, finding all edge pixels of the image through `cv2.findContours()` function of OpenCV, as shown in Fig. 6b. Then finding the pair of points which are with farthest pixel distance from the edge pixels, as shown in Fig. 6c. To facilitate

measurement, the relative diameter of the lettuce equals the distance multiplied by 0.01 (Eq. (2)).

$$\text{Relativediameter} = \text{MAX} \left(\sqrt{(x_i - x_j)^2 + (y_i - y_j)^2} \right) \times 0.01 \quad (2)$$

Where (x_i, y_i) and (x_j, y_j) are the pixel coordinates of the ith and jth points. Finally, the relationship of relative diameter and plant diameter can be fingered by the linear regression model [29].

(2) Diameter prediction based on point cloud segmentation result.

First, the 3D point cloud, as shown in Fig. 7a, is projected along the z-axis into a 2D image, as shown in Fig. 7b. Then finding the points in the 2D convex hull by Graham scan [30], as shown in Fig. 7c. Finally, finding the pair of points which are with the farthest Euclidean distance from the 2D convex hull, as shown in Fig. 7d. The distance is the plant diameter.

2.4.3. Leaf area, fresh weight and dry weight

Leaf area refers to the surface area of lettuce leaves, and its value is the cumulative sum of the area values of all lettuce leaves after they are flattened. Fresh weight is the weight of lettuce under normal living conditions of cells. Dry weight is the weight of the lettuce after the cells have removed all free water. The leaf is an important organ for the photosynthesis and the transpiration of plants and is highly correlated with crop yield. It has long been the focus and difficulty of agricultural research to measure leaf area, crop fresh weight, and crop dry weight simply and accurately. This paper obtained the predictors based on the image and point cloud segmentation results, and then predicted the leaf

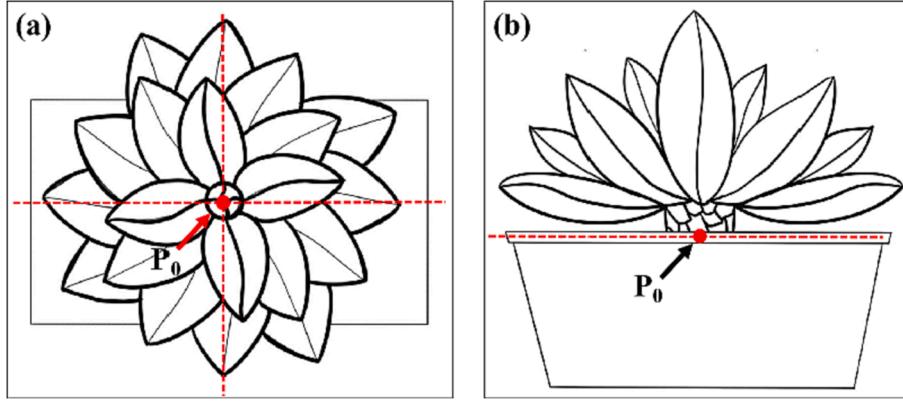


Fig. 8. Virtual point schematic diagram. (a) The position of the emergence point in the XY plane. (b): The position of the emergence point in the Z axis.

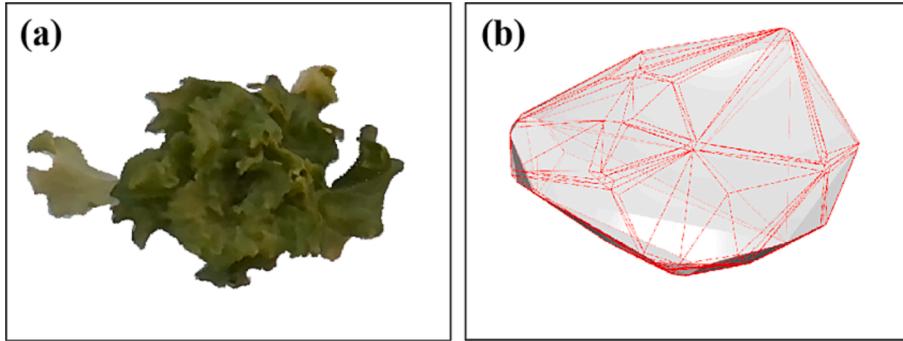


Fig. 9. Point cloud transformation diagram. (a) The point cloud of lettuce. (b): Triangulation result of convex hull.

area, fresh weight and dry weight of lettuce by these factors.

(1) Predictors based on the image segmentation result.

The projected area of a plant refers to the area surrounded by the outer contour line of the shadow formed by the projection of the plant to the bottom when the sun is directly above [31]. It can reflect the size of the area the plant receives sunlight and has a close relationship with the leaf area, fresh weight, and dry weight of the plant.

The number of pixels belonging to lettuce ($Pixl_{leaf}$) was calculated by the image segmentation result. As shown in equation (3), to overcome the differences between the magnitudes of predictors, $Pixl_{leaf}$ was multiplied by 10^{-4} as the relative projected area (PLA).

$$PLA = 10^{-4} \times Count of Pixl_{leaf} \quad (3)$$

(2) Predictors based on point cloud segmentation result.

There is only the information on lettuce in the XY plane can be obtained in the image, while the point cloud adds Z-axis spatial information. Therefore, we hope to extract more effective predictors from the point cloud. Due to the camera angle, the lower part of the lettuce is often obscured and therefore does not appear in the point cloud, which results in the point cloud only reflecting the morphology of the outer surface of the lettuce and cannot show the amount of lettuce inside. To solve this problem, we add a virtual point ($P_0(x_0, y_0, z_0)$) to the original point cloud to simulate the emergence position of lettuce. The point is at the center of gravity of the point cloud in the XY plane, as shown in Fig. 8 (a), and at the height of the top of the hydroponic box in the Z-axis position, as shown in Fig. 8(b).

Firstly, change the above point cloud to the 3D convex hull and curve it, as shown in Fig. 9. The triangular mesh area is obtained by equating the sum of all triangular areas. Suppose there exists an $\triangle ABC$ in the mesh, D_{AB} , D_{BC} , and D_{CA} are the Euclidean distances of AB, BC, and CA, respectively, as shown in Eq. (4). L is the half of the perimeter of $\triangle ABC$. As shown in equation (5), S_{ABC} is the area of $\triangle ABC$. The surface area of

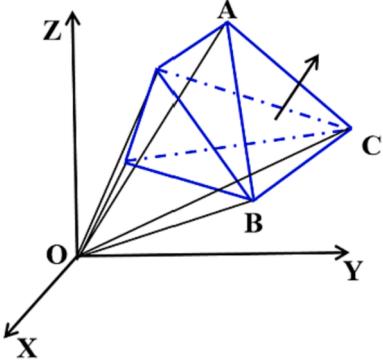


Fig. 10. Calculating volume of the polyhedron.

the mesh is shown in equation (6), where S_i is the area of the i th triangle in the mesh. To overcome the difference between the magnitudes of each predictive factor, the surface area is multiplied by 10^4 as the surface area of the relative 3D convex hull (SUR).

$$L = \frac{D_{AB} + D_{BC} + D_{CA}}{2} \quad (4)$$

$$S_{ABC} = \sqrt{L(L-D_{AB})(L-D_{BC})(L-D_{CA})} \quad (5)$$

$$SUR = 10^4 \times \sum_{i=1}^n S_i \quad (6)$$

The triangles in the mesh are connected to the coordinate origin O to form multiple tetrahedra. The volume of a curve is the algebraic sum of the signed volumes of all tetrahedra [32]. As shown in Fig. 10, the volume symbol of OABC is the symbol of the dot product of vector OA and

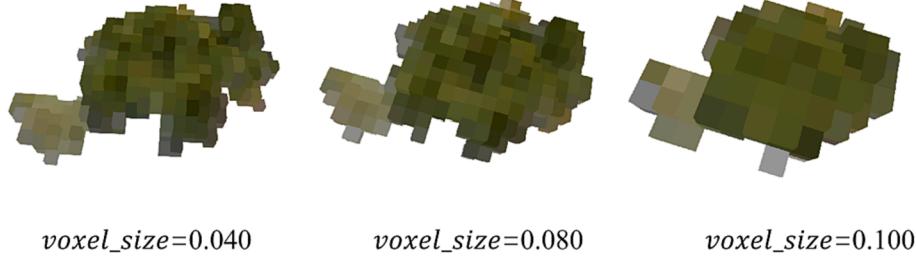
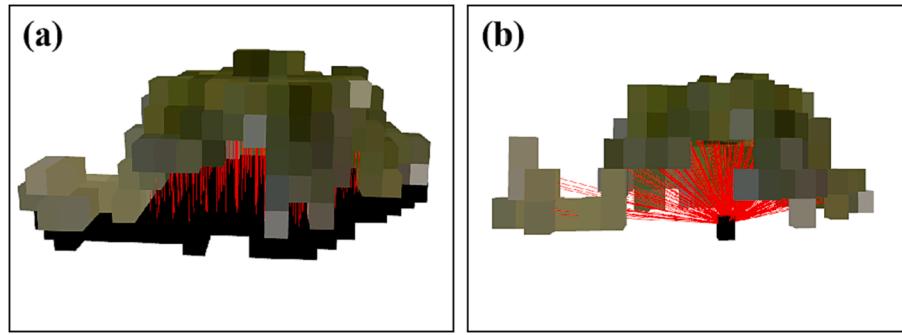
Fig. 11. Voxels of different *voxel_size* (cm).

Fig. 12. Predictors based on point cloud voxelization. (a): The black part is the XY plane where the point P is located, the red line is the distance ($Z_0 - Z_i$) between each voxel and the emergence point plane. (b): The black point is the point P, the red line is the distance between each voxel and the emergence point (D_{i0}). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

the normal vector of triangle ABC. Thus, the volume of OABC is calculated as shown in equation (7), where (X_{ij}, Y_{ij}, Z_{ij}) are the coordinate values of point j on the ith triangle and V_i represents the signed volume of a tetrahedron formed by the ith triangle and O. To overcome the difference between the magnitudes of each predictive factors, the mesh volume is multiplied by 10^4 as the relative 3D convex hull volume (VOL), as shown in equation (8).

$$V_i = \frac{1}{6} \begin{pmatrix} -X_{ib}Y_{ic}Z_{ia} + X_{ic}Y_{ib}Z_{ia} + X_{ib}Y_{ia}Z_{ic} \\ -X_{ia}Y_{ib}Z_{ic} - X_{ic}Y_{ia}Z_{ib} + X_{ia}Y_{ic}Z_{ib} \end{pmatrix} \quad (7)$$

$$VOL = 10^4 \times \sum_{i=0}^n V_i \quad (8)$$

As shown in Fig. 9, the 3D convex hull is the 3D point set that includes all the points, which loses the information of the concave part of the point cloud. While the concave hull [33] can represent the point cloud morphology more accurately compared with the convex hull. Different thresholds need to be set for different point clouds in the process of constructing the concave hull, and the concave hull curve often has hollows due to uneven point cloud distribution [34]. That's the reason of choosing the 3D convex hull rather than the concave hull in this paper.

(3) Predictor based on point cloud voxelization.

Based on the above analysis, the properties of voxel were proposed to solve the above problems. Point cloud voxelization is the transformation of an unstructured point cloud into a regular voxel grid. Compared with the convex hull, the voxel carries more information about the concave part in the point cloud; compared with concave hull, there is no problem that the parameter is hard to determine. Therefore, this paper obtained some predictors based on point cloud voxelization. Fig. 11 shows voxels in different *voxel_size* (cm).

Based on the results of point cloud voxelization, the number of voxels was multiplied by 10^{-2} was the relative voxel number (VOX_1), as shown in equation (9). Limited by the single-view point cloud, the number of voxels only shows the information of the outer surface of lettuce while it is hard to obtain the part being covered. The height of the voxel is the

distance between the voxel point and the plane where the emergence point is located. The area of each voxel multiplied by the height of the voxel is recorded as the volume of each voxel's projector. The sum of all these projectors' volume multiplied by 10^{-2} is recorded as relative voxel projector volume (VOX_2), as shown in Fig. 12(a) and equation (10). *Voxel_size* is the size of the voxel, Z_0 is the Z-axis coordinate value of the emergence point, and Z_i is the Z-axis coordinate value of the ith voxel. In order to fit the real shape of lettuce more efficiently, each voxel is projected on the emergence point. The area of voxel multiplied by the distance between the voxel and the emergence point is approximated as the volume of the projector formed by single voxel projecting on the emergence point, as shown in Fig. 12(b). The sum of above volume multiplied by 10^{-2} is recorded as the relative projector volume of the voxel projected to the emergence point (VOX_3), as shown in equation (11). Where (X_i, Y_i, Z_i) is the coordinate value of the ith voxel, (X_0, Y_0, Z_0) is the coordinate value of the emergence point, and D_{i0} is the Euclidean distance between voxel and germination point (Eq. (12)).

$$VOX_1 = 10^{-2} \times CountifVoxel \quad (9)$$

$$VOX_2 = 10^{-2} \times \sum_{i=0}^{i=0} voxel_size^2 \times (Z_0 - Z_i) \quad (10)$$

$$VOX_3 = 10^{-2} \times \sum_{i=0}^{i=0} voxel_size^2 \times D_{i0} \quad (11)$$

$$D_{i0} = \sqrt{(X_i - X_0)^2 + (Y_i - Y_0)^2 + (Z_i - Z_0)^2} \quad (12)$$

Different voxel sizes will obtain voxels in different morphology, which will have a direct influence on VOX_1 , VOX_2 , and VOX_3 . The following experiments were carried out to find the best voxel size. The minimum distance between adjacent points in the point cloud is about 0.040 cm, and the maximum distance is about 0.100 cm, so the voxel size range is 0.040 cm–0.100 cm, in increments of 0.100 cm. The correlation coefficients of VOX_1 , VOX_2 , and VOX_3 . With lettuce leaf area, fresh weight and dry weight was calculated, respectively. Then we take the average of the above correlation coefficients as *TotalCorrel*, as shown in equation (13), where X_i is VOX_1 , VOX_2 , and VOX_3 , Y_i is the lettuce leaf

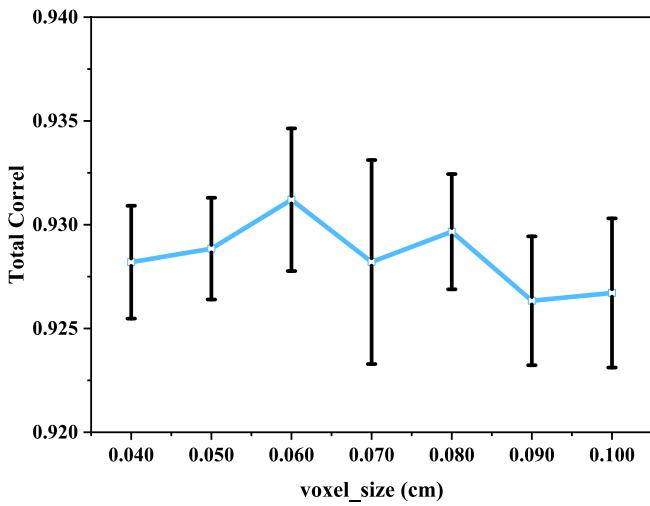


Fig. 13. Comprehensive correlation coefficients with different *voxel_size* (cm).

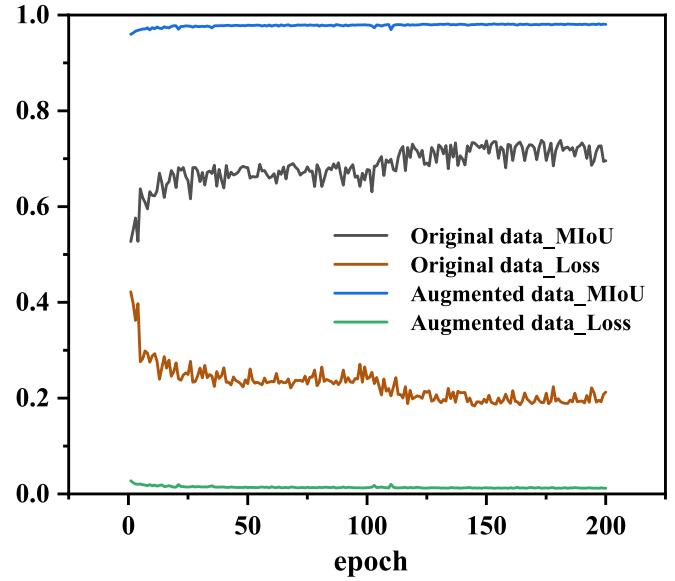


Fig. 14. mIoU and Loss changes in segmentation training.

Table 2
Detail for each factor.

Factors	Data Sources
PLA	Image segmentation results
SUR	3D convex hull based on point cloud segmentation results
VOL	3D convex hull based on point cloud segmentation results
VOX ₁ ,	Voxelization based on point cloud segmentation results
VOX ₂	Voxelization based on point cloud segmentation results
VOX ₃	Voxelization based on point cloud segmentation results
X × GP	The product of above predictor (X) and growth period

area, fresh and dry weight, *Correl* is a function of calculating correlation coefficients in Python's NumPy package. As Fig. 13 shows, 0.060 cm is the best *voxel_size*.

$$\text{TotalCorrel} = \frac{1}{9} \sum_{i=1}^{n=3} \text{Correl}(X_i, Y_i) \quad (13)$$

Obviously, there is a close relationship between the leaf area, fresh weight, dry weight, and the growth period (*GP*) of lettuce. To improve the effect of parameter prediction, this paper proposed to convert the single factor into a composite factor by multiplying the prediction factor and *GP*. In summary, the detailed information on each predictor was shown in Table 2.

2.5. Parameter extraction evaluation index

In the train dataset, the selection of the optimal predictor of leaf area, fresh weight, and dry weight is evaluated by *R*²(Eq. (14)), *RRMSE*(Eq. (15)), *MAPE*(Eq. (16)), and *Totalscore*. *Totalscore* is the comprehensive score under the comprehensive queuing scoring method. Sort the k different predictors according to *R*², *RRMSE* and *MAPE*, from the best to the worst. The first factors score is k, the last score is 1 point, and the middle factor is reduced by one point in order. If there are identical data, average them. Finally, the three index ranking scores are accumulated as *Totalscore*.

$$R^2 = 1 - \frac{\sum_{i=1}^n (y_i - f_i)^2}{\sum_{i=1}^n (y_i - \bar{y}_i)^2} \quad (14)$$

$$RRMSE = \sqrt{\frac{\sum_{i=1}^n (y_i - f_i)^2}{n}} \quad (15)$$

Table 3
Comparison of segmentation accuracy after data augmentation.

Data	Accuracy	mIoU
Original data	0.930	0.644
Augmented data	0.998	0.988

$$MAPE = \frac{\sum_{i=1}^n \left| \frac{y_i - f_i}{y_i} \right|}{n} \times 100\% \quad (16)$$

In the test dataset, on the basis of the above indicators, the *NMSE*(Eq. (17)), and *Totalerror* (Eq. (18)), were added to verify the result. Where *y_i* is the true value of the *i*th sample, *f_i* is the prediction result of the *i*th sample, *y_{ij}* is the true value of the *i*th sample of *j*th parameters, *f_{ij}* is the prediction result of the *i*th sample of *j*th parameters, *n* is the total number of samples, and *m* is the total number of phenotypic parameters and biomass of lettuce.

$$NMSE = \frac{\sum_{i=1}^n (y_i - f_i)^2}{\sum_{i=1}^n y_i^2} \quad (17)$$

$$Totalerror = \sum_{i=1}^m \frac{\sum_{j=1}^n (y_{ij} - f_{ij})^2}{\sum_{j=1}^n y_{ij}^2} \quad (18)$$

3. Results

3.1. 2D image segmentation results

The RGB images before and after data augmentation were input to U-Net for training. The result can be seen in Fig. 14 that data augmentation has a significant effect on lettuce image segmentation during the training process.

The experimental results of the test set show that the test results of the model after data augmentation are better. As shown in Table 3, its mIoU can reach 0.988, which is enough for point cloud segmentation.

3.2. Influence of RGB information on point cloud segmentation

To verify whether RGB information is helpful for point cloud segmentation, three point cloud segmentation networks (PointNet [19],

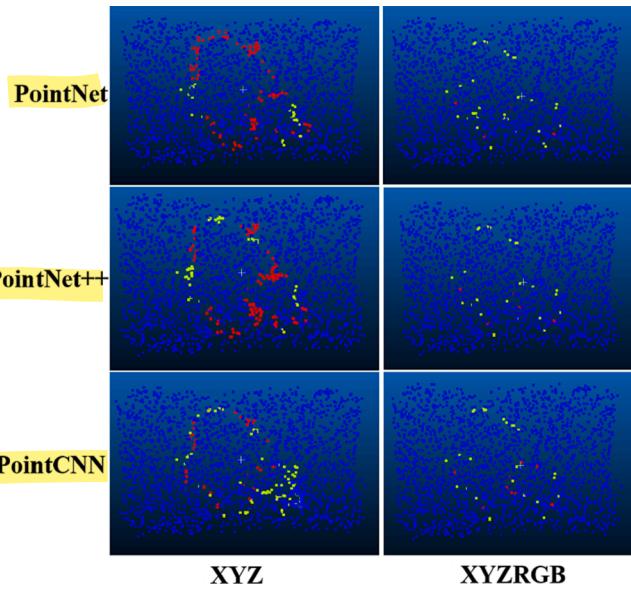


Fig. 15. The training results of models with different input.

Note: Blue points represent the points that are predicted correctly, red points represent the points that are wrongly predicted, and green points represent the points that are wrongly predicted as the background.

Table 4
Test precision on each model.

Model	mIoU	Accuracy
PointNet	0.884	0.946
PointNet++	0.895	0.952
PointCNN	0.890	0.951
PointNet_RGB	0.959	0.982
PointNet++_RGB	0.960	0.983
PointCNN_RGB	0.957	0.981

Note: The input data of the first to the third row model is XYZ, and the input data of other models is XYZRGB.

Table 5
Results comparison of each segmentation methods on the test set.

Segmentation methods	mIoU	Accuracy	Resolution
Image mapping (before statistical filtering)	0.960	0.981	>100000
Image mapping (after statistical filtering)	0.967	0.987	>100000
PointNet_RGB	0.962	0.984	4096
PointNet++_RGB	0.964	0.984	4096
PointCNN_RGB	0.963	0.984	4096

PointNet++ [35] and PointCNN [36]) are selected in this paper. The influence of color information on the point cloud segmentation effect is verified by inputting different dimensions (XYZ, XYZRGB). The inference results of the test set with different resolutions input on different models are shown in Fig. 15, and the test accuracy results are shown in Table 4.

As shown in Table 4, the segmentation accuracy of the three models increases after adding RGB information. In Fig. 15, blue points represent the points that are predicted correctly, red points represent the points that are wrongly predicted, and green points represent the points that are wrongly predicted as the background. It can be seen from Fig. 15 that the error points of model prediction are significantly reduced after adding RGB information. It shows that adding RGB information to model training can better solve the problem of difficult segmentation due to the adhesion of foreground and background point clouds.

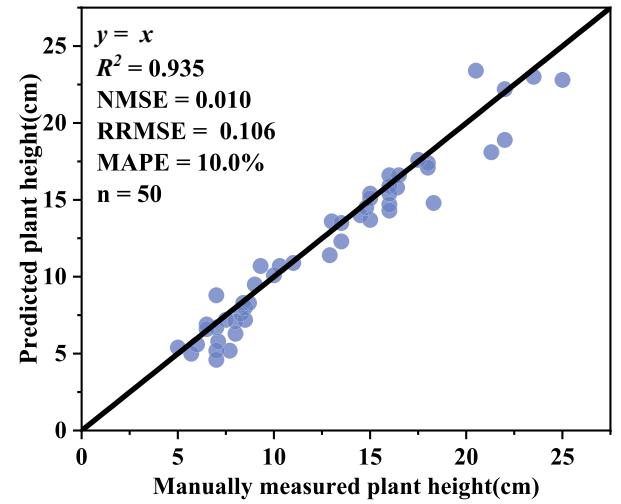


Fig. 16. Scatter plots of predicted plant height and manually measured plant height on test set. Note: n is the number of samples.

3.3. The point cloud segmentation results

The image segmentation results are mapped to the point cloud. The segmentation accuracy is shown in Table 5. The mIoU of point cloud segmentation through mapping is 0.960, and the Accuracy is 0.981. After statistical filtering of the mapping results, the mIoU is improved to 0.967, and the Accuracy is improved to 0.987. The resolution of the deep learning model is 4096, and the proposed method maintains the original resolution.

The entry dimension of the point cloud segmentation model is fixed, the resolution of the point cloud needs to be specified before the model is trained. However, there are many invalid points in the point cloud, and the resolution of each point cloud is different, which makes the sampling of the point cloud necessary for the model. Due to the specificity of the point cloud and the complexity of the model, high-resolution point clouds require higher arithmetic requirements and time costs. So the sampled point clouds generally have lower resolutions. However, low-resolution point clouds can lead to the loss of a lot of information, thus reducing the measuring accuracy of plant height and diameter. In contrast, the point cloud segmentation results obtained by the image-mapping-based method have no resolution limitation, so this segmentation method is more beneficial to the extraction of phenotypic parameters.

3.4. Extraction of plant height

Based on the point cloud segmentation results, the plant height is obtained by the coordinate difference method. As shown in Fig. 16, the R^2 was 0.935, the $RRMSE$, $MAPE$ and $NMSE$ were 0.106, 10.0 % and 0.010. The experimental results showed that the plant height extraction effect was good, indicating that the plant height prediction method proposed in this study was effective and relatively accurate.

3.5. Extraction of plant diameter

The relative diameter is obtained based on the two-dimensional image segmentation results, and the regression model is established to predict the real diameter using the relative diameter. As shown in Fig. 17a, $R^2 = 0.88$, and $RRMSE = 0.091$. The segmentation result based on the 3D point cloud is the direct extraction of lettuce diameter. As shown in Fig. 17b, R^2 is 0.882, and $RRMSE$ is 0.090. Therefore, the direct extraction of diameter based on point cloud segmentation results is better than the prediction of diameter based on image segmentation results. The diameter obtained in this way is closer to the actual value

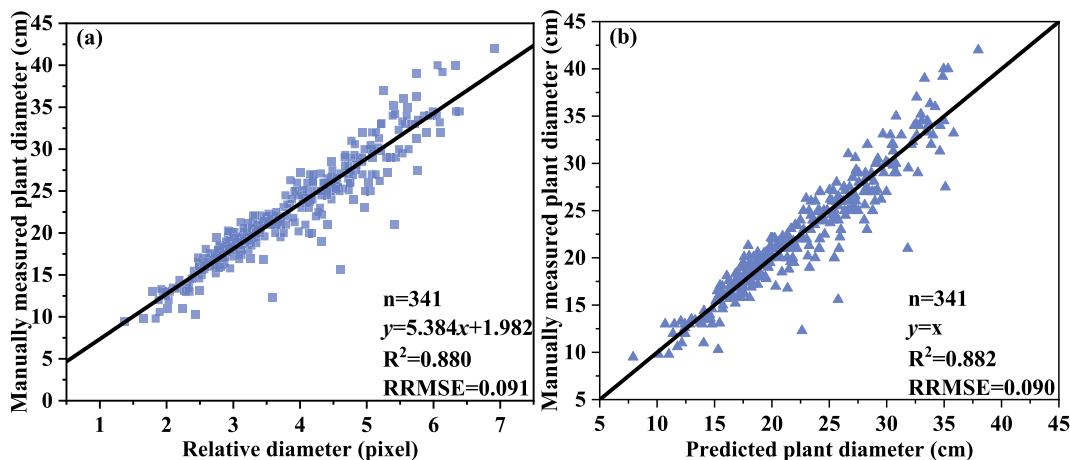


Fig. 17. Scatter plots of predicted plant diameter and manually measured plant diameter. (a): Diameter prediction by image segmentation results on train set. (b): Diameter prediction by point cloud segmentation results on train set.

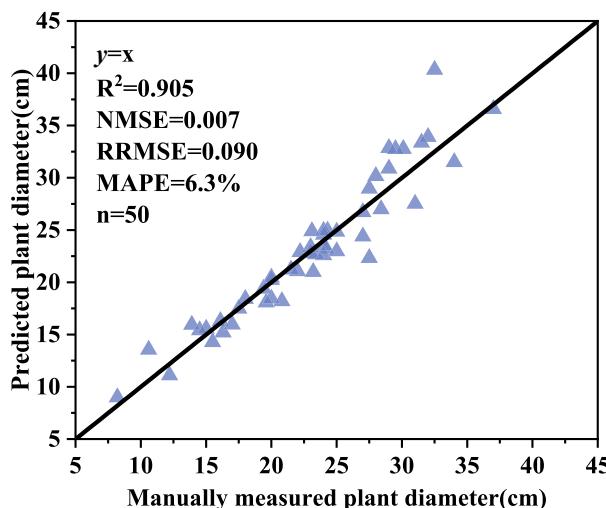


Fig. 18. Diameter prediction by point cloud segmentation results on test set.

without conversion. In the absence of measured diameter data, it is more appropriate to use existing data directly to obtain the diameter for practical production situations.

The method of obtaining lettuce diameter based on point cloud segmentation results is applied to the test set. As shown in Fig. 18, R^2 is 0.905, RRMSE is 0.090, MAPE is 6.3 %, and NMSE is 0.007. The results show: based on the point cloud segmentation results, the lettuce diameter obtained by the two-dimensional convex hull is reliable.

3.6. Prediction of leaf area, fresh weight and dry weight

Firstly, the modeling scores of different predictors are calculated. The highest score factor is considered the independent variable of modeling. Different varieties of lettuce have different forms. To better calculate the leaf area, fresh weight, and dry weight of lettuce, this paper established models for each variety of lettuce and applied them to the number of 50 test sets for effect verification. The predictive effects of different factors are shown in Table 6.

In the prediction of leaf area, when multiplied PLA , SUR , and VOX_1 by the growth period, the prediction effect was improved, thus proving the validity of the composite factor. VOX_3 has the best results with R^2 of 0.945, RRMSE of 0.194, MAPE of 26.9 %, and *Totalscore* of 33.700, followed by $VOX_1 \times GP$, which also proved the effectiveness of factor

Table 6
Prediction effect of different factors.

	Single Factors	R^2	RRMSE	MAPE(%)	<i>Totalscore</i>	Composite Factors	R^2	RRMSE	MAPE(%)	<i>Totalscore</i>
Leaf area	<i>PLA</i>	0.932	0.216	35.9	18.100	<i>PLA</i> \times <i>GP</i>	0.942	0.198	32.6	25.500
	<i>VOL</i>	0.924	0.229	29.1	15.500	<i>VOL</i> \times <i>GP</i>	0.922	0.231	56.1	7.500
	<i>SUR</i>	0.920	0.234	42.0	7.800	<i>SUR</i> \times <i>GP</i>	0.941	0.200	27.4	28.600
	<i>VOX</i> ₁	0.925	0.227	42.3	11.000	<i>VOX</i> ₁ \times <i>GP</i>	0.945	0.195	26.9	33.300
	<i>VOX</i> ₂	0.942	0.199	30.1	26.900	<i>VOX</i> ₂ \times <i>GP</i>	0.928	0.221	62.6	10.400
	<i>VOX</i> ₃	0.945	0.194	26.9	33.700	<i>VOX</i> ₃ \times <i>GP</i>	0.933	0.213	57.3	15.700
Fresh weight	<i>PLA</i>	0.910	0.286	118.2	10.800	<i>PLA</i> \times <i>GP</i>	0.953	0.208	40.8	32.200
	<i>VOL</i>	0.906	0.292	62.2	10.200	<i>VOL</i> \times <i>GP</i>	0.935	0.243	40.8	22.800
	<i>SUR</i>	0.886	0.321	130.2	4.000	<i>SUR</i> \times <i>GP</i>	0.942	0.231	51.2	22.100
	<i>VOX</i> ₁	0.894	0.310	134.8	5.000	<i>VOX</i> ₁ \times <i>GP</i>	0.949	0.215	44.2	26.900
	<i>VOX</i> ₂	0.933	0.248	40.1	21.900	<i>VOX</i> ₂ \times <i>GP</i>	0.951	0.212	52.0	25.400
	<i>VOX</i> ₃	0.935	0.243	51.1	19.100	<i>VOX</i> ₃ \times <i>GP</i>	0.955	0.203	42.0	33.600
Dry weight	<i>PLA</i>	0.917	0.239	57.0	14.700	<i>PLA</i> \times <i>GP</i>	0.952	0.181	29.3	31.900
	<i>VOL</i>	0.896	0.266	31.9	11.700	<i>VOL</i> \times <i>GP</i>	0.918	0.237	63.6	14.100
	<i>SUR</i>	0.894	0.269	66.1	5.900	<i>SUR</i> \times <i>GP</i>	0.942	0.198	25.0	31.000
	<i>VOX</i> ₁	0.908	0.251	68.3	8.000	<i>VOX</i> ₁ \times <i>GP</i>	0.940	0.203	64.2	22.100
	<i>VOX</i> ₂	0.922	0.231	37.3	18.700	<i>VOX</i> ₂ \times <i>GP</i>	0.932	0.216	72.8	17.400
	<i>VOX</i> ₃	0.928	0.221	30.9	22.900	<i>VOX</i> ₃ \times <i>GP</i>	0.954	0.177	24.0	35.600

Note: Bold and underline is the best score.

Table 7
Details of the optimal predictive models.

Parameter	Factor	Varieties	Predictive models	R ²	RRMSE	MAPE(%)
Leaf area	VOX ₃	Aphylion	y = 14.303x + 252.095	0.945	0.194	26.9
		Lugano	y = 19.651x + 17.383			
		Salanova	y = 20.107x - 107.838			
		Satine	y = 14.692x + 157.786			
Fresh weight	VOX ₃ × GP	Aphylion	y = 0.149x + 20.223	0.955	0.203	42.0
		Lugano	y = 0.248x + 6.148			
		Salanova	y = 0.146x + 2.668			
		Satine	y = 0.165x + 12.707			
Dry weight	VOX ₃ × GP	Aphylion	y = 0.00615x + 1.48196	0.954	0.177	24.0
		Lugano	y = 0.00876x + 1.013			
		Salanova	y = 0.00688x + 0.67433			
		Satine	y = 0.00702x + 0.93233			

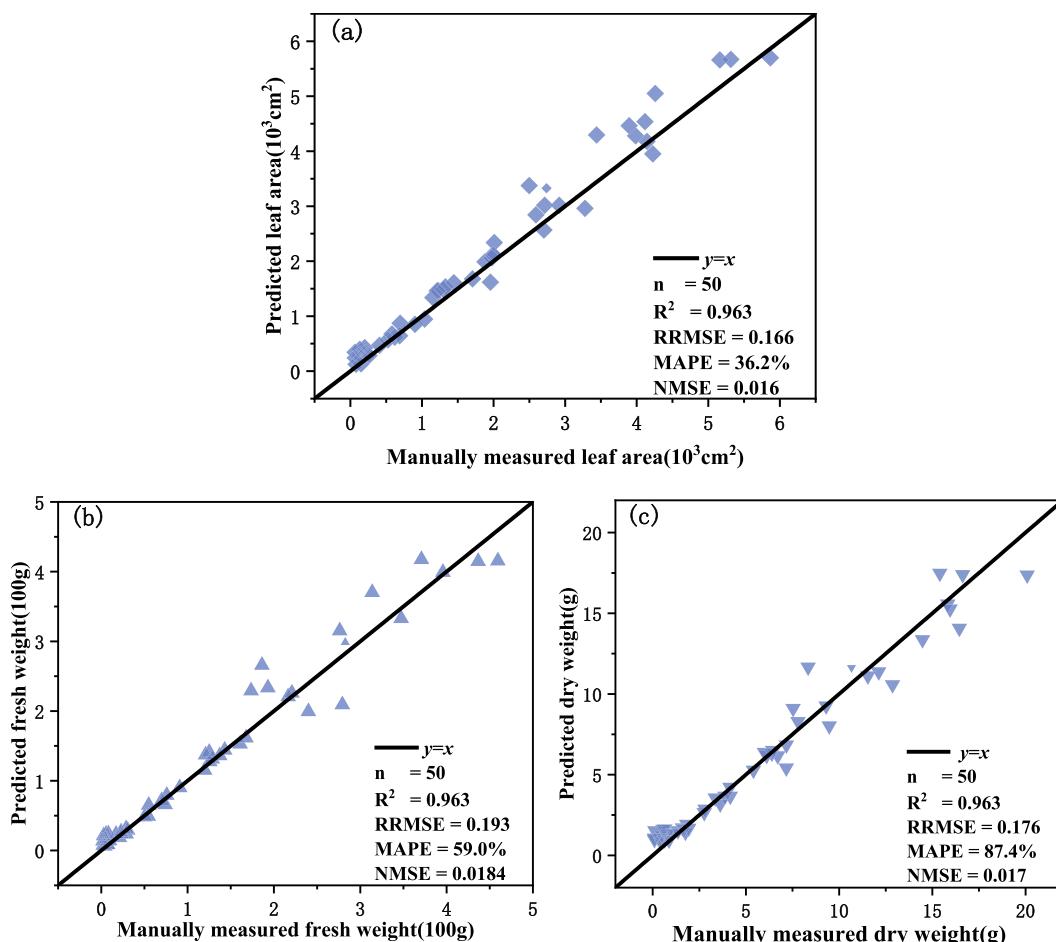


Fig. 19. Scatter plots of predicted parameters and manually measured parameters on test set. (a): Leaf area (b): Fresh weight (c): Dry weight. Note: n is the number of samples.

VOX₃ in predicting leaf area of lettuce.

In the prediction of fresh weight and dry weight, when multiplied all predictors by the growth period, the prediction effect was improved.

Table 8
Predicted results for each phenotypic parameter on test set.

Parameter	R ²	RRMSE	MAPE(%)	NMSE
Plant height	0.935	0.106	10.0	0.010
Plant diameter	0.905	0.090	6.5	0.007
Leaf area	0.963	0.166	36.2	0.016
Fresh weight	0.963	0.193	59.0	0.018
Dry weight	0.963	0.176	87.4	0.168

This also proves the validity of composite factor prediction. For fresh weight, VOX₃ × GP has the best results with R² of 0.955, RRMSE of 0.203, MAPE of 42.0 % and Totalscore of 33.600, followed by PLA × GP. For dry weight, VOX₃ × GP also has the best results with R² of 0.954, RRMSE of 0.177, MAPE of 24 % and Totalscore of 35.6, followed by PLA × GP.

The best predictors for leaf area, fresh weight, and dry weight were VOX₃, VOX₃ × GP and VOX₃ × GP, respectively. And their corresponding prediction models are shown in Table 7.

According to the models selected above, the test set is substituted for model testing. The predicted performance of leaf area is shown in Fig. 19 (a), where the R² was 0.963, the RRMSE and NMSE were as low as 0.166

and 0.016, respectively. The predicted performance of fresh weight is shown in Fig. 19(b), where the R^2 was 0.963, the RRMSE and NMSE were as low as 0.193 and 0.0184, respectively. The predicted performance of dry weight is shown in Fig. 19(c), where the R^2 was 0.963, the RRMSE and NMSE were as low as 0.176 and 0.017, respectively.

In summary, using a set of methods proposed in this paper to get the plant height, diameter, leaf area, fresh weight and dry weight of lettuce on the test set. As shown in Table 8, the Total error is as low as 0.069, which is 15 % better than the first-place score (0.081) of the online challenge based on this data set, indicating that the methods this paper proposed is reliable and accurate.

4. Conclusion

Plant phenotypic parameter extraction is one of the important basic tasks in the field of agriculture. To solve the problem of difficult and low accuracy of plant phenotypic parameters acquisition, this paper takes lettuce as the research object and proposes a set of high precision and low-cost lettuce phenotypic parameters acquisition methods.

Accurately separating crop point cloud from the environmental point cloud is the basis of point cloud research. For the resolution limitation of the point cloud segmentation model, a method based on the image segmentation result of U-Net mapping is proposed. The segmentation mIoU can reach 0.9668, and the original resolution of the segmented data can be guaranteed. Further research can try to combine 2D image segmentation and point cloud mapping to form a complete model.

Based on the segmentation results of the image and point cloud, this paper proposed multiple prediction factors to accurately predict the leaf area, fresh weight, and dry weight of lettuce. The R^2 of all predicted all above 0.90. This paper only uses linear model modeling, involving fewer models. In future research, it is proposed to use various functional forms (power function, exponential function, logarithmic function) and various model building methods, such as Decision Trees (DT), K-Nearest Neighbors (KNN), Support Vector Machines (SVM) to further improve the accuracy of parameter prediction. Because of the more complex lettuce phenotype and lettuce biomass, this experiment was modeled by varieties. There are great differences between lettuce seedlings and the mature period. Because this data is time-series, modeling by lettuce growth period classification can also be considered.

CRediT authorship contribution statement

Yu Zhang: Data curation, Software, Methodology, Visualization, Writing – original draft, Writing – review & editing. **Maowei Li:** Conceptualization, Investigation, Software, Writing – original draft, Writing – review & editing. **Guixin Li:** Validation, Data curation, Formal analysis. **Jinsong Li:** Software, Data curation, Formal analysis. **Lihua Zheng:** Resources, Writing – review & editing. **Man Zhang:** Resources, Writing – review & editing. **Minjuan Wang:** Resources, Supervision, Project administration, Writing – review & editing.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

I am using a public dataset, which is detailed in the article.

Acknowledgement

The authors thank Autonomous Greenhouses International Challenge 3rd held by Wageningen University & Research for the

contribution in supplying the RGB-D images and corresponding manual measurement data with different varieties.

Funding

This work was supported in part by National Natural Science Foundation of China (32201654), the Natural Science Foundation of Shandong Province (ZR2021MC021) and the Open Project Program of State Key Laboratory of Virtual Reality Technology and Systems, Beihang University (No. VRLAB2021A02).

References

- [1] I. Medina-Lozano, J.R. Bertolín, A. Diaz, Nutritional value of commercial and traditional lettuce (*Lactuca sativa L.*) and wild relatives: Vitamin C and anthocyanin content[J], Food Chem. 359 (2021), 129864, <https://doi.org/10.1016/j.foodchem.2021.129864>.
- [2] L. Zhang, W. Su, R. Tao, et al., RNA sequencing provides insights into the evolution of lettuce and the regulation of flavonoid biosynthesis[J], Nat. Commun. 8 (1) (2017) 1–12, <https://doi.org/10.1038/s41467-017-02445-9>.
- [3] D. Story, M. Kacira, C. Kubota, et al., Morphological and textural plant feature detection using machine vision for intelligent plant health, growth and quality monitoring[J], Acta Hortic. (2011).
- [4] M. Watt, F. Fiorani, B. Usadel, et al., Phenotyping: new windows into the plant for breeders[J], Annu. Rev. Plant Biol. 71 (2020) 689–712, <https://doi.org/10.1146/annurev-aplant-042916-041124>.
- [5] W. Yang, X. Xu, L. Duan, et al., High-throughput measurement of rice tillers using a conveyor equipped with x-ray computed tomography[J], Rev. Sci. Instrum. 82 (2) (2011), 025102, <https://doi.org/10.1063/1.3531980>.
- [6] A. Hartmann, T. Czauderna, R. Hoffmann, et al., HTPheno: an image analysis pipeline for high-throughput plant phenotyping[J], BMC Bioinf. 12 (1) (2011) 1–9, <https://doi.org/10.1186/1471-2105-12-148>.
- [7] A.S. Iyer-Pascuzzi, O. Symonova, Y. Mileyko, et al., Imaging and analysis platform for automatic phenotyping and trait ranking of plant root systems[J], Plant Physiol. 152 (3) (2010) 1148–1157, <https://doi.org/10.1104/pp.109.150748>.
- [8] T. Ishikawa, A. Hayashi, S. Nagamatsu, Classification of strawberry fruit shape by machine learning[J], et al., International Archives of the Photogrammetry, Remote Sensing & Spatial, Inf. Sci. 42 (2) (2018), <https://doi.org/10.5194/isprs-archives-xlii-2-463-2018>.
- [9] S.A. Bello, S. Yu, C. Wang, et al., Deep learning on 3D point clouds[J], Remote Sensing 12 (11) (2020) 1729, <https://doi.org/10.3390/rs12111729>.
- [10] M. Vázquez-Arellano, D.S. Parafos, D. Reiser, et al., Determination of stem position and height of reconstructed maize plants using a time-of-flight camera[J], Comput. Electron. Agric. 154 (2018) 276–288, <https://doi.org/10.1016/j.compag.2018.09.006>.
- [11] Y. Song, J. Wang, Winter wheat canopy height extraction from UAV-based point cloud data with a moving cuboid filter[J], Remote Sensing 11 (10) (2019) 1239, <https://doi.org/10.3390/rs11101239>.
- [12] Yidan Y, Chengda L, Ruifang Z, et al. Rapeseed 3D reconstruction and morphological parameter measurement based on laser point cloud[C]//2016 Fifth International Conference on Agro-Geoinformatics (Agro-Geoinformatics). IEEE, 2016: 1–6. 10.1109/agro-geoinformatics.2016.7577638.
- [13] Gérald W, Devy M, Herbolut A, et al. Model-based segmentation of 3D point clouds for phenotyping sunflower plants[C]//12. International Joint Conference on Computer Vision, Imaging and Computer Graphics Theory and Applications. 2017. 10.5220/0006126404590467.
- [14] N. Dube, B. Bryant, H. Sari-Sarraf, et al., Cotton boll distribution and yield estimation using three-dimensional point cloud data[J], Agron. J. 112 (6) (2020) 4976–4989, <https://doi.org/10.1002/agj2.20412>.
- [15] Y. Ohashi, Y. Ishigami, E. Goto, Monitoring the growth and yield of fruit vegetables in a greenhouse using a three-dimensional scanner[J], Sensors 20 (18) (2020) 5270, <https://doi.org/10.3390/s20185270>.
- [16] P.J. Besl, R.C. Jain, Segmentation through variable-order surface fitting[J], IEEE Trans. Pattern Anal. Mach. Intell. 10 (2) (1988) 167–192, <https://doi.org/10.1109/34.3881>.
- [17] T. Rabbani, F. Van Den Heuvel, G. Vosselmann, Segmentation of point clouds using smoothness constraint[J], International archives of photogrammetry, remote sensing and spatial information sciences 36 (5) (2006) 248–253.
- [18] N.K. Sallem, M. Devy, Extended grabcut for 3d and rgb-d point clouds[C]// International Conference on Advanced Concepts for Intelligent Vision Systems, Springer, Cham, 2013, pp. 354–365.
- [19] Qi C R, Su H, Mo K, et al. Pointnet: Deep learning on point sets for 3d classification and segmentation[C]//Proceedings of the IEEE conference on computer vision and pattern recognition. 2017: 652–660. 10.1109/cvpr.2017.16.
- [20] A.V. Pham, M. Le Nguyen, Y.L.H. Nguyen, et al., Dgcnn: A convolutional neural network over large-scale labeled graphs[J], Neural Networks 108 (2018) 533–543, <https://doi.org/10.1016/j.neunet.2018.09.001>.
- [21] Zhao H, Jiang L, Fu C W, et al. Pointweb: Enhancing local neighborhood features for point cloud processing[C]//Proceedings of the IEEE/CVF conference on computer vision and pattern recognition. 2019: 5565–5573. 10.1109/cvpr.2019.00571.

- [22] W. Shi, R. van de Zedde, H. Jiang, et al., Plant-part segmentation using deep learning and multi-view vision[J], Biosyst. Eng. 187 (2019) 81–95, <https://doi.org/10.1016/j.biosystemseng.2019.08.014>.
- [23] Hemming, S. (Silke); de Zwart, H.F. (Feije); Elings, A. (Anne); bijlaard, monique; Marrewijk, van, Bart; Petropoulou, Anna (2021): 3rd Autonomous Greenhouse Challenge: Online Challenge Lettuce Images. 4TU.ResearchData. Dataset. 10.4121/15023088.v1.
- [24] Torralba A, Russell B C, Yuen J. Labelme: Online image annotation and applications[J]. Proceedings of the IEEE, 2010, 98(8): 1467-1484. 10.1109/jproc.2010.2050290.
- [25] Bloice M D, Stocker C, Holzinger A. Augmentor: an image augmentation library for machine learning[J]. arXiv preprint arXiv:1708.04680, 2017. 10.21105/joss.00432.
- [26] D. Girardeau-Montaut, CloudCompare[J], EDF R&D Telecom ParisTech, France, 2016, p. 11.
- [27] O. Ronneberger, P. Fischer, T. Brox, U-net: Convolutional networks for biomedical image segmentation[C]//International Conference on Medical image computing and computer-assisted intervention, Springer, Cham, 2015, pp. 234–241.
- [28] T.S. Breure, A.E. Milne, R. Webster, et al., Predicting the growth of lettuce from soil infrared reflectance spectra: the potential for crop management[J], Precis. Agric. 22 (1) (2021) 226–248, <https://doi.org/10.1007/s11119-020-09739-x>.
- [29] S. Weisberg, Applied linear regression[M], John Wiley & Sons, 2005.
- [30] R.L. Graham, F.F. Yao, Finding the convex hull of a simple polygon[J], Journal of Algorithms 4 (4) (1983) 324–331, [https://doi.org/10.1016/0196-6774\(83\)90013-5](https://doi.org/10.1016/0196-6774(83)90013-5).
- [31] G.S. Campbell, J.M. Norman, The description and measurement of plant canopy structure[J], Plant canopies: their growth, form and function 1 (1989) 19, <https://doi.org/10.1017/cbo9780511752308.002>.
- [32] Zhang C, Chen T. Efficient feature extraction for 2D/3D objects in mesh representation[C]//Proceedings 2001 International Conference on Image Processing (Cat. No. 01CH37205). IEEE, 2001, 3: 935-938. 10.1109/icip.2001.958278.
- [33] S. Asaeedi, F. Didehvar, A. Mohades, α -Concave hull, a generalization of convex hull[J], Theoret. Comput. Sci. 702 (2017) 48–59, <https://doi.org/10.1016/j.tcs.2017.08.014>.
- [34] Y. He, Z. Hu, K. Wu, et al., A novel method for density analysis of repaired point cloud with holes based on image data[J], Remote Sensing 13 (17) (2021) 3417, <https://doi.org/10.3390/rs13173417>.
- [35] C.R. Qi, L. Yi, H. Su, et al., Pointnet++: Deep hierarchical feature learning on point sets in a metric space[J], Advances in neural information processing systems (2017) 30, <https://doi.org/10.1109/cvpr.2017.16>.
- [36] Y. Li, R. Bu, M. Sun, et al. Pointcnn: Convolution on x-transformed points[J], Adv. Neural Inform. Process. Syst., 2018, 31. 10.48550/arXiv.1801.07791.