

Original papers

Research on automatic 3D reconstruction of plant phenotype based on Multi-View images

Danni Yang^a, Huijun Yang^{a,c,d,*}, Dongfeng Liu^b, Xianlin Wang^b^a College of Information Engineering, Northwest A&F University, Yangling 712100, Shaanxi, China^b Shenzhen Agricultural Science and Technology Promotion Center, Nanshan 518000, Guangdong, China^c Key Laboratory of Agricultural Internet of Things, Ministry of Agriculture and Rural Affairs, Yangling, 712100, Shaanxi, China^d Shanxi Engineering Research Center of Agricultural Information Intelligent Perception and Analysis, Yangling, 712100, Shaanxi, China

ARTICLE INFO

ABSTRACT

Keywords:

Plant phenotyping
3D reconstructed point cloud
Structure from motion
Phenotyping platform

Three-dimensional reconstruction plays a crucial role in quantifying crop phenotypes and exploring crop physiological structures. This paper presents a phenotyping platform designed for the 3D reconstruction of complex plants, utilizing multi-view images and introducing a joint evaluation criterion for both the reconstruction algorithm and the platform. Initially, a device composed of Raspberry Pi camera, SSH protocol, USB-TTL, motorized turntable, and shadow booth is built for automated image acquisition and transmission. Then, a dataset containing carex cabbage and kale is created and trained based on U^2 -net to achieve precise image segmentation. After that, an improved structure from motion algorithm, named IVOP & AKAZE-SFM, and multi-view stereo algorithm are utilized for the fine-scale reconstruction of plants. Next, by combining color filtering and Euclidean clustering, a denoising algorithm is proposed to obtain clean point clouds of plants. Finally, a method for calibrating the size of the plant point cloud based on priori condition is proposed to solve the problem of point cloud deformation in reconstruction. The evaluation of image segmentation model resulted in a precision of 0.91, a recall of 0.972, an IOU of 0.943, and a $\max F_\beta$ of 0.099. The proposed IVOP&AKAZE-SFM is assessed against mainstream algorithms, the results show that our algorithm has the minimum average track length, minimum average reprojection error and generated the most points. The correlation coefficient (R^2) between the extracted traits and measured phenotype, such as plant height and plant width, are 0.999 and 1.000, while the root mean squared errors (RMSE) are 0.298 cm and 0.338 cm. Consequently, the platform offers a cost-effective, automated, and integrated solution for fine-scale plant 3D reconstruction.

1. Introduction

Carex cabbage and kale play significant roles in China's agriculture. Phenotypic traits of Carex cabbage and kale, such as plant height and width, reflect their growth and robustness, providing valuable insights for selecting cultivars with high yield, good quality, and disease resistance (Pieruschka and Schurr, 2019). Traditional phenotyping methods are time-consuming and labor-intensive, and also cause harm to the plants. With the rapid development of computer vision, plant phenotyping platforms based on vision technology have become an effective alternatives. These platforms can be categorized into two-dimensional platforms that utilize visible light imaging techniques (Crowell et al., 2014; Minervini et al., 2017; Parent et al., 2015; Tausen et al., 2020) and three-dimensional platforms that rely on 3D reconstruction techniques

(Sandhu et al., 2019; Thapa et al., 2018). However, in situations where the plant structure is complex and the leaves are heavily sheltered, the 2D platform can only obtain simple phenotypic traits, such as leaf length, leaf area, and leaf angle, due to the limitation of two-dimensional projection. In contrast, 3D platforms can accurately reconstruct point cloud models for nondestructively analyzing crop structure and obtaining more comprehensive and accurate phenotypic traits (Zhao et al., 2019).

Phenotyping platforms based on 3D reconstruction techniques can be categorized into three types based on the method of acquiring point clouds: laser scanner (Forero et al., 2022; Paulus et al., 2014), depth camera (Teng et al., 2021; Wang et al., 2022; Wang and Chen, 2020; Xu et al., 2023), and multi-view image (Martinez-Guanter et al., 2019; Rossi et al., 2020; Sunvittayakul et al., 2022; Yang and Han, 2020).

* Corresponding author.

E-mail addresses: dannyyang828@163.com (D. Yang), yhj740225@163.com (H. Yang), liudfh@163.com (D. Liu), 369883149@qq.com (X. Wang).

Nevertheless, these 3D reconstruction methods have individual applicable scenarios and limitations. Laser scanning methods can reconstruct dense point clouds of equal size to real objects, whereas high equipment prices limit their promotion and application. The depth camera, such as Kinect, can rapidly capture point clouds, yet, the quality of reconstructed point clouds is poor, and the robustness to the lighting is also poor. The 3D reconstruction method based on multi-view images primarily applies the structure from motion algorithm (Moulon et al., 2013) (SFM) and multiple view stereo algorithm (Furukawa and Ponce, 2010) (MVS) to reconstruct point cloud, offering the advantages of affordable price, convenient operation, and high quality. Therefore, it has progressively become the optimal solution for phenotyping platforms.

The 3D phenotyping platforms based on multi-view images can be categorized into two modes depending on the relative motion between the camera sensor and the target plant (Fiorani and Schurr, 2013): “camera to plant” and “plant to camera”. In the “camera to plant” mode, the target plant remains stationary while one or more cameras rotate around it to capture multi-view images. Wu et al. (Wu et al., 2022; Wu et al., 2020) developed MVS-Pheno and MVS-Pheno V2 phenotyping platforms for maize and wheat based on this model. The “camera to plant” mode enables high-quality image acquisition but necessitates a dedicated camera rotation device and control system. Additionally, it requires space depending on plant size to adjust the distance between the camera and the plant. In the “plant to camera” mode, the target plant is just placed on a rotating turntable and cameras remain fixed to capture multi-view images of the plant. Liu et al. (Liu et al., 2017) utilized the “plant to camera” technique for achieving arabidopsis phenotype, allowing the quick data acquisition and offer the minimal space requirements and little labor. The “plant to camera” model has been applied to several plants, such as wheat, pepper, Carex cabbage, and other specie (Duan et al., 2016; He et al., 2017; Wang et al., 2022b). Indeed, the platform based on “plant to camera” has the following drawbacks: (1) It is only suitable for early growth stages when plant architectures are simple, and occlusions are minimal. (2) The quality of the point cloud is poor. (3) The processing framework is separate. Thus, a high-precision, automated, and integrated 3D reconstruction platform for complex plants based on “plant to camera” mode is urgently needed.

In this paper, a new 3D reconstruction platform based on the “plant to camera” mode is presented for complex plants such as kale and carex cabbage, which have multiple tillers, leaves, and compact structures. The platform can automatically generate high-precision point clouds of plants for observing phenotypic traits, with its performance is evaluated in terms of reconstruction accuracy and phenotype accuracy. The main contributions are as follows:

- (1) By combining an automatic image acquisition-transmission system, the image segmentation model trained on U^2 -net, and an improved structure from motion algorithm (IVOP & AKAZE-SFM), a 3D reconstruction method is proposed for obtaining high-precision point cloud of complex plant automatically.
- (2) A scale restoration method, including color filtering, Euclidean clustering, and a size calibration algorithm based on prior information, is proposed to obtain a clean and proportional point cloud of complex plants.
- (3) A plant phenotyping platform based on the multi-view images to reconstruct high-precision and proportional point clouds for complex plants is constructed, integrating hardware components such as the Raspberry Pi camera and motorized turntable, as well as algorithms for image acquisition, image segmentation, 3D reconstruction, and point cloud processing.

2. Materials and methods

2.1. Environmental conditions of plant material

The data are collected at the experimental field of Shen Zhen Agricultural Science and Technology Promotion Center ($114^{\circ}36'6403''N, 22^{\circ}75'6488''E$). As shown in Fig. 1, five cultivars of carex cabbage with different shoot structures (Xiong Xi Yi Hao (XXYH), He Dou Bai (HDB), WT-8 (WT), Nai Bai Yi Hao (NBYH), and Nai Bai Cai (NBC)) and one kale cultivar (Long Tian Yi Hao (LTYH)) are selected as experimental materials to verify the stability of the algorithm and the versatility of the platform. The experimental plants are cultivated in a greenhouse with facilities, including water curtains and fans. Additionally, the seedlings are initially grown in a seedbed and subsequently transplanted into pots with a row spacing of 15 cm and a plant spacing of 16–22 cm.

During the growth stage of Carex cabbage, ternary compound fertilizer served as the base fertilizer, while urea is applied later by water and fertilizer integration. For the growth stage of kale, urea, and potassium chloride are applied after planting, and potassium chloride compound fertilizer is applied at the shoots stage.

The plants are sampled at the shooting and flowering stages to demonstrate the complex structural characteristics of the leafy vegetable plants. At the shooting stage in November 2022, the XXYH cultivars are sampled (3 replications). The Kale cultivars are sampled at the flowering stage in December 2022 (10 replications). In February 2023, the HDB, NBC, NBYH, and WT cultivars are sampled at the flowering stage (3–5 replicates). The diameter and height of the pots, as well as the plant height and plant width, are initially recorded by manual measurements, as presented in Table 1. Subsequently, an image acquisition system obtained multi-view sequence images of the plants, serving as the data source for subsequent experiments.

2.2. Pipeline of 3D reconstruction method

The pipeline of 3D reconstruction method proposed in this paper is shown in Fig. 2. Firstly, the Zhengyou Zhang calibration method (Zhengyou, 1999) is used to obtain the parameters of the HQ Camera-based Raspberry Pi camera. Moreover, an image acquisition system automatically captures images of plants by integrating the Raspberry Pi camera, motorized turntable, and USB-TTL. Then the image sequences are automatically transmitted to the computer via wireless SSH protocol. Secondly, the image background is removed by the image segmentation mode trained on U^2 -net, which is employed to reconstruct the point cloud mode using the proposed IVOP&AKAZE-SFM and multi-view stereo algorithms. Thirdly, the point cloud is denoised by the color filtering and the Euclidean clustering algorithm. Finally, to calibrate the size of the point cloud, the proposed algorithm starts with color filtering and statistical filtering to extract the clean pot. Next, it calibrates the orientation of plant point cloud using the RANSAC algorithm and Rodrigues' rotation algorithm. In addition, the scale factor is obtained by the pot point cloud measured with the Axis-Aligned Bounding Box (AABB) and utilized to construct the transformation matrix for calibration.

2.3. IVOP & AKAZE-SFM based 3D plant reconstruction

2.3.1. Establishment of Raspberry Pi camera

In point cloud phenotyping platforms, the conventional method of image acquisition involves using USB cables to synchronize signals and images across multiple cameras and computers. However, this approach is vulnerable to cable entanglement, leading to disruptions in the seamless and automated acquisition process. To address this challenge, we devised a Raspberry Pi camera system utilizing the SSH protocol, the HQcamera (IMX477), and a Raspberry Pi and automated data transfer through a pre-established network connection. This strategic

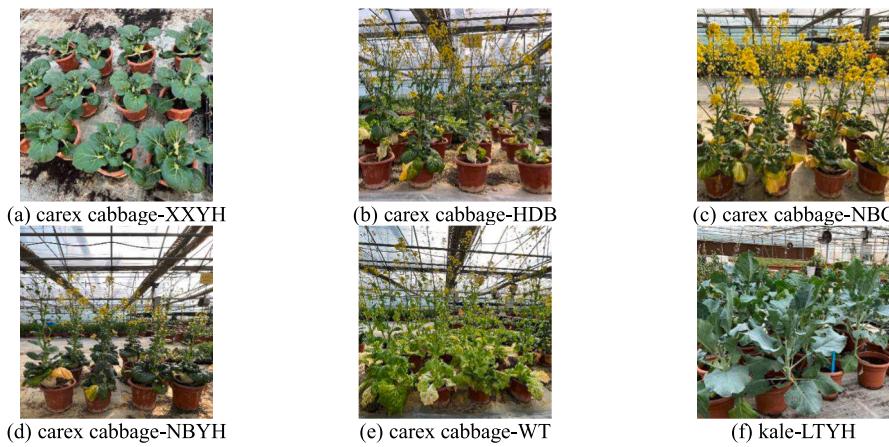


Fig. 1. carex cabbage and kale cultivars.

Table 1
pot and plant data by manual measurements.

Cultivar	Plant height (cm)	Plant width (cm)	Pot diameter (cm)	Pot height (cm)
LTYH0	49	52	22	18
LTYH1	48.5	48.5	22	18
LTYH2	44	48	22	18
LTYH3	44	44	22	18
LTYH4	52.5	49	22	18
LTYH5	47.5	54	22	18
LTYH6	58	46	22	18
LTYH7	59.5	56	22	18
LTYH8	60	58	22	18
LTYH9	49	48.5	22	18
HDB0	61	33	16	12.7
HDB1	64.5	32	16	12.7
HDB2	61.5	30	16	12.7
HDB3	65	36	16	12.7
WT0	67.5	42	19	14.2
WT1	50	32	19	14.2
WT2	53	22	16	12.3
NBYH0	38	24	18	14.5
NBYH1	34	23	18	14.5
NBYH2	40	22	18	14.5
NBYH3	53	27	18	14.5
NBC0	43	24	16	12.3
NBC1	53	29	16	12.3
NBC2	55	33	16	12.3
NBC3	55	36	16	12.3
NBC4	54	26	16	12.3

implementation effectively mitigated the cable entanglement issue, thereby enhancing the overall automation efficiency of the system. The integration of the HQcamera and high-resolution industrial lens contributed to the acquisition of high-quality images. Consequently, the Raspberry Pi camera system presents significant advantages in terms of both automation and high resolution.

The Raspberry Pi camera units, as depicted in Fig. 3, comprise a Raspberry Pi (4B), HQcamera (IMX477), and an 8 mm industrial prime lens. The Raspberry Pi is a card-based computer that runs on the LINUX-based Raspbian operating system with kernel 4.1.20-v7+. The HQcamera is a high-definition camera accessory equipped with Sony IMX477 CMOS, enabling the acquisition of high-resolution images. The relevant specifications of the device are provided in Table 2.

The proposed 3D reconstruction method utilizes the Raspberry Pi camera as the image acquisition equipment and the computer terminal as the data analysis equipment. The image acquisition equipment has functions for image acquisition, storage, and transmission, while the data analysis equipment has functions for image reception, image segmentation, point cloud reconstruction, and calibration. The image

acquisition program runs on the Raspberry Pi and utilizes the Picamera module to capture images regularly, collecting 64 image samples for each plant. Additionally, the image transmission program runs on the computer and receives the images from the Raspberry Pi via the wireless SSH protocol.

2.3.2. Raspberry Pi camera parameter extraction

The camera projection model serves as the foundation and critical element in 3D reconstruction. The classical pinhole camera model defines the relationship between pixel points in 2D images and 3D points in the real world using the projection equation (1):

$$P = M \bullet V \bullet P_w \quad (1)$$

P represents the coordinates on the image plane, M is the projection matrix that delineates the camera's internal parameters. V is the view matrix describing the camera's external parameters. P_w denotes a point in world coordinates, P, M, V are 4×4 matrices, while P_w is presented as a 4-dimensional column vector.

Aberrations caused by lens manufacturing and assembly errors can deviate the spatial points from the ideal position of the pinhole camera projection model, leading to the distortion of the image and the 3D model. The main types of distortion are radial and tangential, with emphasis typically placed on the effects of radial distortion. Radial distortion is related to the distance between the image point and the main point, causing a radial shift of the imaging point along the radial direction. The radial distortion model is expressed by equation (2):

$$\begin{cases} u' = u(1 + k_1 r^2 + k_2 r^4 + k_3 r^6) \\ v' = v(1 + k_1 r^2 + k_2 r^4 + k_3 r^6) \end{cases} \quad (2)$$

Where $r^2 = x^2 + y^2$, k_1, k_2, k_3 are the radial distortion coefficients, and (u, v) and (u', v') are the ideal (no distortion) and actual (distorted) normalized image coordinates, respectively.

The values of K, k_1, k_2, k_3 , etc., which correct the distorted lens and establish an accurate camera projection model, are obtained by the Zhengyou Zhang's calibration method in this paper. The following steps are performed:

- (1) Prepare a chessboard with a grid size of 15 mm * 15 mm and dimensions of 6*9 corners, as illustrated in Fig. 4.
- (2) Detect the corner points of the chessboard and obtain their pixel coordinates. Calculate their physical coordinates based on the known grid size and the origin of the world coordinate system.
- (3) Determine the internal parameter matrix K of the Raspberry Pi camera.

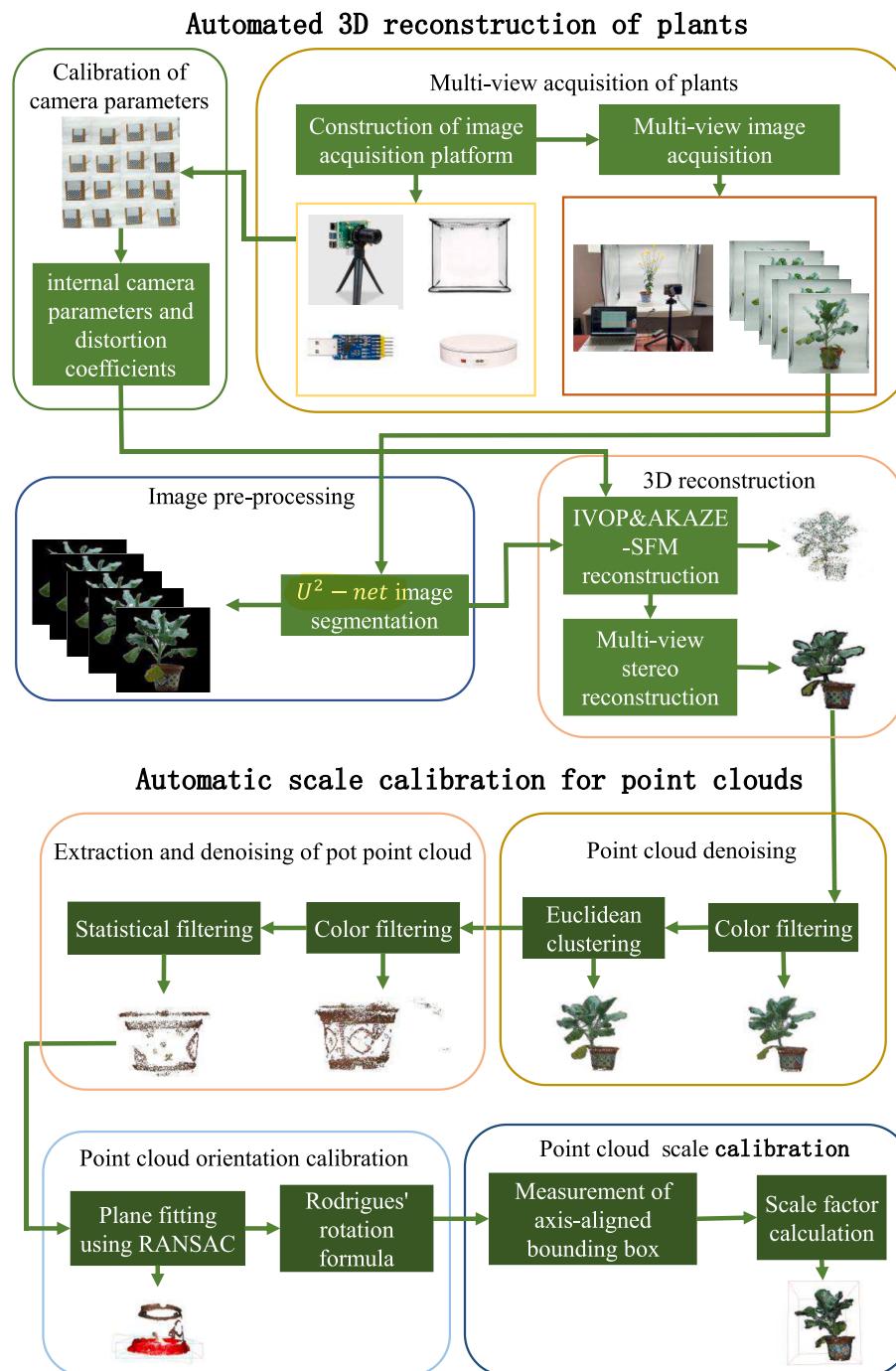


Fig. 2. The pipeline of 3D reconstruction method.



Fig. 3. The Raspberry Pi camera units.

Table 2
Raspberry Pi camera specification.

Parameters	Value
Sensor diagonal(mm)	7.9
Focal length (mm)	8
Image width(pixels)	1920
Image height (pixels)	1080
Close-up range (m)	1.35
Angle of view (°)	56.5°*43.9

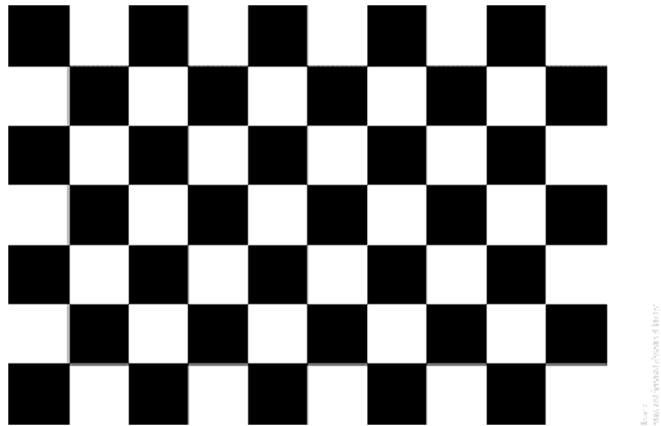


Fig. 4. chessboard board (6×9).

- (4) Determine the distortion coefficients k_1, k_2, k_3 of the Raspberry Pi camera.
- (5) Employ the L-M algorithm to optimize the aforementioned parameters.
- (6) Obtain the internal parameters $K = \begin{bmatrix} 2309.1 & 0 & 948.6 \\ 0 & 2311.0 & 472.6 \\ 0 & 0 & 1 \end{bmatrix}$ and distortion coefficients of the Raspberry Pi camera $k_1 = -0.083, k_2 = -0.38, k_3 = 1.9$.



(a)Studio



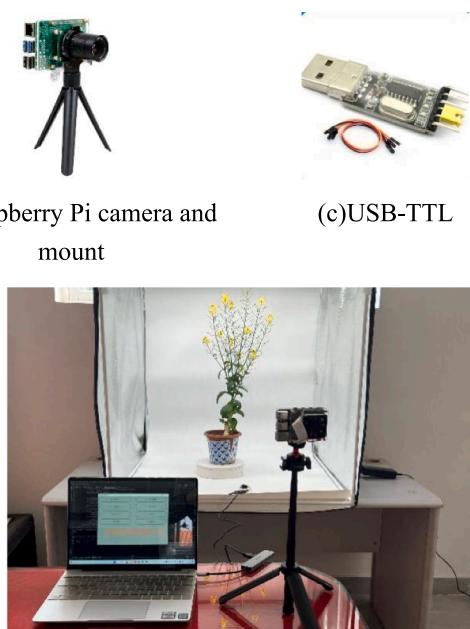
(b)Raspberry Pi camera and mount

2.3.3. Automated plant image acquisition

An automated system for image acquisition and wireless transmission is presented, as depicted in Fig. 5. The system is based on the studio, Raspberry Pi camera, camera mount, USB-TTL, motorized turntable, and computer. The studio is a 60 cm cube covered with shade cloth and equipped with LED light sources, providing a simple background, uniform illumination, and wind protection for image capture. The Raspberry Pi camera is used to capture and transmit high-resolution images. The camera mount adjusts the camera's height to accommodate plants of different heights in various growth stages. The USB-TTL sends infrared signals to control the rotation of the turntable. The motorized turntable with an inner diameter of 20 cm drives the target object to rotate. The computer is the receiver of the images.

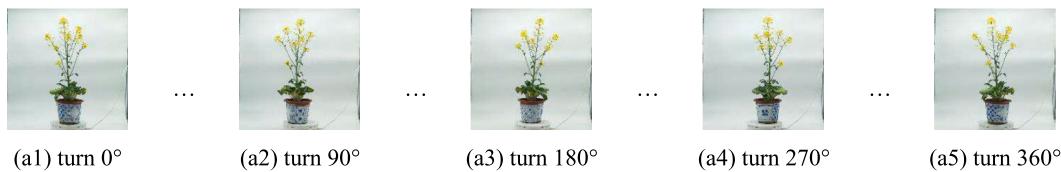
This paper proposes a linear procedure for image acquisition and wireless transmission as follows:

- (1) Sample preparation. The samples are selected in the greenhouse, and colored stickers are applied around the pots. Colored stickers provide additional correspondence, as the plants cannot offer enough pairs of image features for the 3D reconstruction due to their uniform color and irregular texture. In addition, applying distinctive stickers can improve the ability to completely extract the flowerpot point cloud when segmenting the point cloud.
- (2) Device setup. A turntable is placed in the center of the studio. The target plant is positioned on the turntable, and the lighting is adjusted accordingly. Following that, the Raspberry Pi camera is placed at a distance of 40 cm from the target plant, and its viewing angle is adjusted. Furthermore, the USB-TTL is connected to the computer and positioned near the turntable.
- (3) Image capture. The computer controls the turntable's rotation at a constant speed by sending infrared signals from USB-TTL, while the Raspberry Pi camera captures images from different viewpoints at a synchronized rate. An overlap of 80 %-90 % between consecutive images is necessary to ensure the quality of 3D reconstruction. Sixty-four images are captured in one cycle of turntable rotation, and some of the original images are shown in Fig. 6.



(e)Image acquisition system

Fig. 5. Hardware of image acquisition system.

**Fig. 6.** the image sequences (partial).

- (4) Image transmission. Once the image capture is completed, the images are wirelessly transferred to the computer terminal via wireless SSH protocol for further processing.

2.3.4. U^2 -netBased plant image segmentation

The image acquisition method, based on the “plant to camera” mode includes background information, such as the turntable and the shadow booth, which potentially interferes with the accuracy and efficiency of 3D reconstruction for the plant. Therefore, there is a need to remove the background of the images.

Traditional segmentation methods struggle with preserving intricate plant edges and accurately handling background shadows. The watershed algorithm, sensitive to noise and local extreme points, often produces unstable segmentation results. Additionally, these methods lack the capacity to capture high-level semantic information in complex plant images, relying mainly on local pixel grayscale values. Similarly, the k-means algorithm’s sensitivity to noise and outliers hinders its performance in achieving smooth segmentation boundaries. In contrast, the U^2 -net (Qin et al., 2020) excels by learning features from extensive data, allowing for end-to-end abstraction, leveraging global context, and exhibiting superior adaptability to complex plant structures and semantic nuances. Consequently, U^2 -net outperforms traditional methods in image segmentation tasks. Thus, plant image segmentation based on U^2 -net is proposed to automatically learn the multi-scale saliency features of images and achieve high-precision image segmentation.

Since no existing image segmentation model is available for complex plants such as carex cabbage and kale, this study utilizes images captured by Raspberry Pi cameras as datasets to train on U^2 -net, obtaining an image segmentation model capable of handling carex cabbage and kale plants with multiple tillers and complex morphology.

(1) Plant segmentation image dataset

Labelme, an annotation software (Torralba et al., 2010), is used for generating the image segmentation dataset. During the labeling process, the pixels in the images are categorized as either plant pixels or non-plant pixels. Plant pixels are assigned a value of 255 and labeled as white, whereas non-plant pixels are assigned a value of 0 and labeled as black. The dataset, comprising 560 images, is divided into train sets and

Table 4
Image Segmentation Dataset for Carex cabbage and Kale.

Cultivar	Training images	Testing images	Total number
XXYH	126	18	144
HDB	70	10	80
NBC	70	10	80
NBYH	70	10	80
WT	70	10	80
LTYH	84	12	96

test sets at a ratio of 8:1. The annotation results are presented in Table 3, and the dataset details are presented in Table 4.

(2) Plant image segmentation

The U^2 -net segmentation model, depicted in Fig. 7, is composed of three components: an encoder layer, a decoder layer, and a fully connected layer. In the U^2 -net, each encoder layer extracts features of a specific scale and enhances them using a U-shaped residual block. The output from each encoder layer is connected to the corresponding decoder layer, and a loss function is computed for each scale. Through iterative training of the samples, the model aims to minimize the loss function value and enhance the segmentation effect on plant images.

The training process comprises several steps. Firstly, the image size is resized to 320×320 and preprocessed by applying a random vertical flip and random crop of 288×288 . Subsequently, the AdamW optimizer is employed with a learning rate of 0.001 and weight decay of 1e-4. Then, set 2000 iterations to achieve convergence of the loss function. During the testing process, the network produced a prediction map size of 320×320 , which is subsequently resized to match the original size of the input image using bilinear interpolation.

2.3.5. 3D plants reconstruction

The IVOP&AKAZE-SFM algorithm is based on the OpenMVG framework (Moulon et al., 2017) and includes several improvements, such as adopting the AKAZE feature extraction algorithm (Alcantarilla and Solutions, 2011) and utilizing the internal camera parameters and distortion coefficients as the initial values for reconstruction. Subsequently, a high-quality point cloud model of the plant is reconstructed by combining the OpenMVS algorithm. The 3D reconstruction algorithm

Table 3
The annotation results of carex cabbage and kale.

Cultivar	HDB	NBC	NBYH	WT	XXYH	LTYH
Original image						
Label						

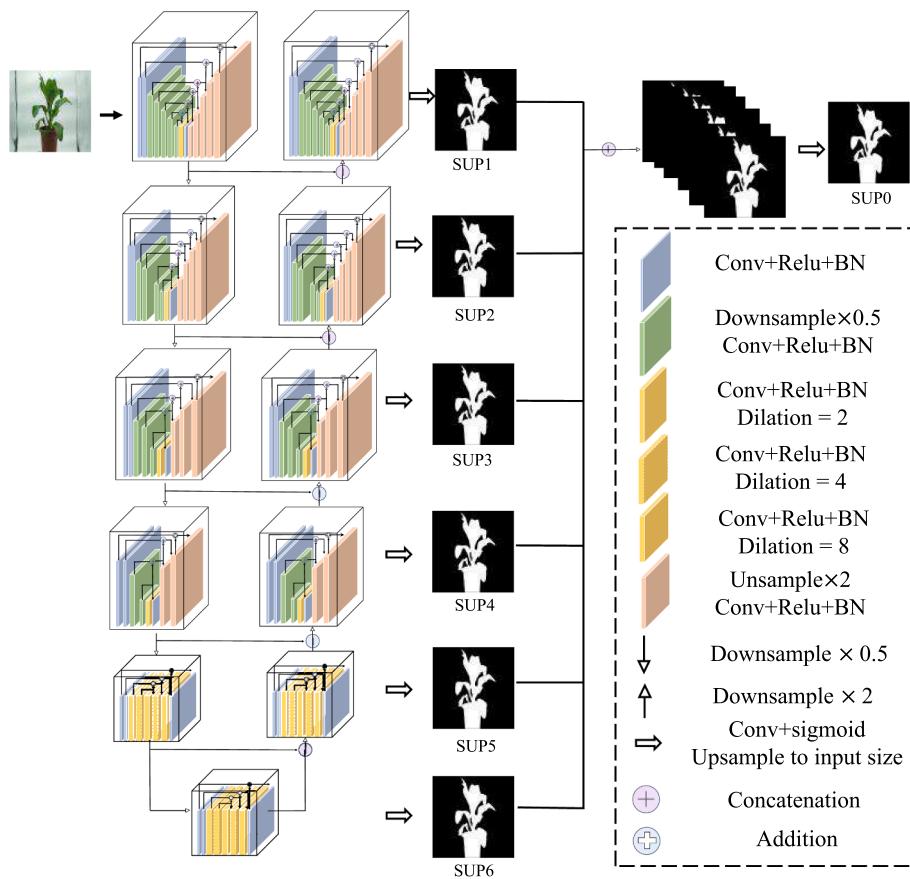


Fig. 7. U^2 -net Model.

flow is depicted in Fig. 8.

The 3D reconstruction process consists of several steps. Firstly, the input includes a set of plant image sequences, the camera's internal parameters, and distortion coefficients. Initializing these parameters and coefficients at the beginning of the 3D reconstruction process provides a robust initial estimate, which reduces calculated errors, accelerates algorithmic convergence, and enhances the reconstruction's quality. Secondly, the AKAZE algorithm is employed to extract features and descriptors from the images, and initial image pairs are generated based on video patterns. The AKAZE algorithm operates in nonlinear scale space and is particularly effective in capturing fine features on morphologically complex carex cabbage and kale plants. Thirdly, geometric validation of the fundamental matrix is performed to eliminate false matches and construct image connectivity maps. Fourthly, to complete the reconstruction initialization, a pair with the largest baseline is selected as the initial image pair from the matched image pairs. Camera poses, and 3D points are then estimated and optimized through linear triangulation and bundle adjustment methods. Fifthly, in the incremental reconstruction stage, the remaining images are gradually incorporated into the reconstruction process. For each newly added image, the PnP algorithm estimated its camera pose, and the triangulation method adds new 3D points. Ultimately, the bundle adjustment method is applied again to optimize camera poses and 3D points. Sixthly, the IVOP &AKAZE-SFM algorithm output includes a sparse 3D point cloud and camera poses. Finally, in the multi-view stereo reconstruction stage, the dense 3D point cloud is generated by the Patch-Match method.

2.4. Automated scale calibration of plant point cloud

2.4.1. Plant comprehensive denoising

Due to the leaf occlusion and matching feature errors, the plant's 3D point cloud model has two primary types of noise. One type is uniformly distributed along the edges of the plant stalks and leaves which is removed by color filter, while the other is scattered throughout the plant which is removed by Euclidean clustering. The original plant point cloud is shown in Fig. 9.

(1) Color filter for uniformly distribution noises

Given the position of the initial noise, occurring at the edges of plant stalks and leaves, presented as black and white, we employ color filtering (Watanabe et al., 2023) to eliminate the specific noise type. The precise denoising algorithmic flow is elucidated as follows:

Step one, the RGB color space is chosen, and the mask matrix B_W is generated utilizing boundary values derived from the color features of the plants.

Initially, the color values within the RGB color space are correlated with the boundary values representative of plant color features. Subsequently, by setting a threshold based on the mapped color values, the color values are partitioned into two distinct categories—one for plant color and the other for noise color. A mask matrix B_W is then generated, wherein the pixel value corresponding to plant color is set to 1, while the pixel value corresponding to noise color is set to 0.

Step two, the color matrix P_c derived from the point cloud data is employed for filtering, resulting in the generation of a new color matrix, denoted as R.

The color matrix P_c derived from the point cloud data undergoes element-wise multiplication with the mask matrix B_W . In the resultant

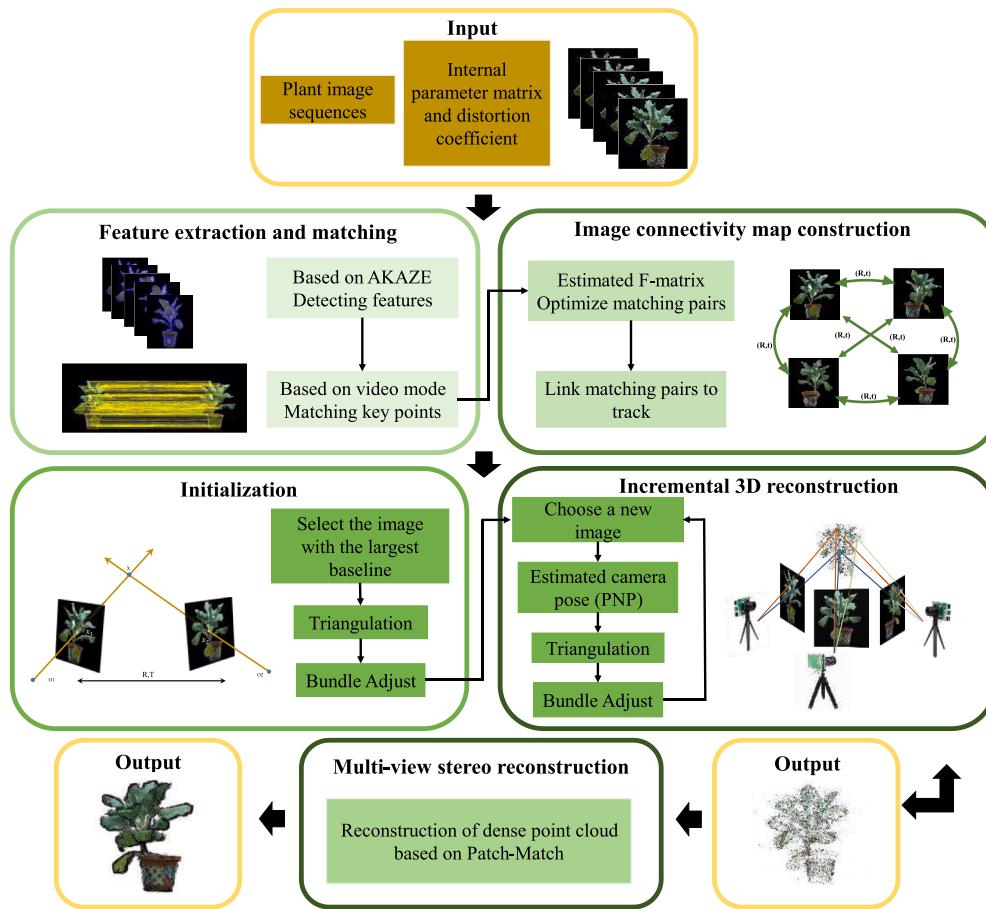


Fig. 8. 3D Reconstruction Based on IVOP &AKAZE-SFM and Multi-View stereo.



Fig. 9. Original plant point cloud.

product, elements corresponding to plant colors retain their original values, while elements corresponding to noise colors become 0. Consequently, a new color matrix R is obtained, exclusively preserving plant color information while effectively filtering out noise colors.

Step three, the denoised plant point cloud is derived by indexing the new color matrix R.

A denoised plant point cloud is extracted from the original point cloud data using the new color matrix R as an index. The resultant plant point cloud is free of noise, encompassing solely the points associated with the plant color.

(2) Euclidean distance clustering for scattered noises

Given that the second type of noise is distant from the target and exhibits an aggregated distribution, a KDTTree based Euclidean distance

clustering algorithm (Guo et al., 2023) is employed for noise removal. The specific denoising algorithmic flow is elucidated as follows:

First, identify neighboring points utilizing a hybrid search function based on KNN and RNN. For each point P_i within the original plant point cloud P, a hybrid KNN- RNN based search function is employed to establish the search radius R, determine the number of neighboring points k, and retrieve their corresponding indexes and distances.

Second, cluster P_i and its neighboring points into a group denoted as C_i . For each point P_i , group it into a cluster C_i along with the identified neighboring points.

Third, eliminate the noise cluster C from the original point cloud P. For all clusters C_i , identify the noisy clusters C within them, and subsequently remove the identified noisy clusters C from the original point cloud P. The resultant plant point cloud Q is obtained after the noise removal process.

2.4.2. Prior information-based plant size calibration

The method utilizes the actual size of the pot as prior information to calibrate the size of the point cloud. Moreover, the process of this method is as follows:

- (1) Extraction of the pot point cloud using color filtering. The boundary values of color in the RGB color space based on pots' color characteristics are constructed to create mask matrix C_W . Then C_W is applied to filter color on color matrix C of the plant point cloud. After that, a new color matrix R of the plant point cloud is generated. Lastly, R, as an index, extracts the pot point cloud.
- (2) Noise removal of the pot point cloud using statistical filtering. $P = \{P_i, i = 1, 2, \dots, n\}$ represents the set of point clouds before

denoising, and $P' = \{P'_i, i = 1, 2, \dots, m\}$ represents the set of point clouds after statistical filtering (Han et al., 2017). S_i represents the average distance from each point P_i in P to k points in its neighbourhood. Subsequently, the average distance from each point P'_i in P' to its corresponding neighbourhood S_i follows a Gaussian distribution with the μ (mean value) and the σ (standard deviation) determined by equation (3).

$$\mu = \frac{\sum_{i=1}^n S_i}{n}, \sigma = \frac{\sum_{i=1}^n (S_i - \mu)^2}{n} \quad (3)$$

Therefore, a point is retained if its distance from the adjacent point clouds falls within the range of $(\mu - std \cdot \sigma, \mu + std \cdot \sigma)$, and it is eliminated if it falls outside this range.

- (3) Orientation calibration of the point cloud using RANSAC and Rodrigues' rotation equation. The reconstructed original point cloud of the plant is tilted to the horizontal plane, requiring orientation calibration to ensure accurate extraction of the phenotypic traits. Additionally, the orientation calibration algorithm comprises the following steps: Initially, fit the pot point cloud to a plane using the RANSAC method. Secondly, estimate the plane equation and acquire the normal vector (a,b,c) of the ground plane. Thirdly, calculate the rotation matrix R based on the normal vector (a, b, c) of the ground plane and the normal vector (0, 0, 1) of the coordinate system using Rodrigues' rotation equation. Finally, apply the rotation matrix R to the original point cloud of the plant to obtain the horizontally calibrated point cloud.
- (4) Construction of transformation matrix using prior information. The scale factor is calculated based on the actual size of the pot to calibrate the size of the plant point cloud. Firstly, obtain the minimum bounding box of the pot point cloud using the AABB algorithm, and calculate the length L, width W, and height H of the bounding box. Secondly, divide the calculated value of the pot point cloud by the actual length L, width W, and height H of the pot to obtain the corrected scales in the x, y, and z directions, respectively. The algorithm steps are as follows:

The AABB is expressed by equation (4):

$$R = \{(X, Y, Z) | x_{min} < X < x_{max}, y_{min} < Y < y_{max}, z_{min} < Z < z_{max}\} \quad (4)$$

Where: (x_{min}, x_{max}) are the minimum and maximum values of the vertices in the AABB on the x-axis, respectively. (y_{min}, y_{max}) are the minimum and maximum values of the vertices in the AABB on the y-axis, respectively. (z_{min}, z_{max}) are the minimum and maximum values of the vertices in the AABB bounding box on the z-axis, respectively. The length L, width W and height H dimensions of the pot point cloud are expressed by equation (5)

$$L = x_{max} - x_{min}, W = y_{max} - y_{min}, H = z_{max} - z_{min} \quad (5)$$

To address potential missing point clouds in pot point clouds, we designed a strategy. When measuring the diameter and height of the point cloud, we opted AABB to measure the maximum value of each direction to approximate the true value, aiming to minimize errors arising from the absence of certain point clouds. Given the actual pot length L_{truth} , width W_{truth} and height H_{truth} the scaling factors of x-axis, y-axis and z-axis of the plant point cloud expressed by equation (6)

$$S_x = \frac{L}{L_{truth}}, S_y = \frac{W}{W_{truth}}, S_z = \frac{H}{H_{truth}} \quad (6)$$

Hence, the plant point cloud transformation matrix T_p and plant point cloud P after size calibration are expressed by equation (7):

$$T_p = \begin{bmatrix} S_x & 0 & 0 & 0 \\ 0 & S_y & 0 & 0 \\ 0 & 0 & S_z & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}, P = P \bullet T_p \quad (7)$$

Where P is plant point cloud before size calibration.

2.5. Joint evaluation method

Presenting a joint evaluation scheme to comprehensively assess the performance of proposed 3D reconstruction methods for complex plants, considering both the accuracy of the point cloud and the accuracy of phenotypic traits.

Evaluating the accuracy of the point cloud model: The IVOP&AKAZE-SFM algorithm undergoes testing and comparison with other open-source algorithms. Additionally, the reconstruction quality of the point cloud model undergoes quantitative analysis through four indicators: the number of points, average track length, number of registered images, average error reprojection and time efficiency. Primarily, the number of points reflects the information in the reconstructed scene. A more considerable value indicates that the algorithm extracts richer features and more details on the reconstructed 3D structure. Notably, the average track length reflects the number of 2D points corresponding to each 3D point. A smaller track length indicates that the algorithm extracts richer features. Especially the number of registered images refers to the number of images used in the reconstruction process. The higher the value, the more information is utilized during reconstruction. Significantly, the average reprojection error is the average distance error between the position of each 3D point projected onto each frame using the estimated pose and the actual detected 2D points. A minor reprojection error indicates higher accuracy of the point cloud. Time efficiency is a crucial metric, typically denoting the duration an algorithm requires to accomplish a task, where smaller durations signify more efficient algorithms.

Evaluation of accuracy of phenotypic traits: Phenotypic traits, including plant height and width, get extracted from the point cloud model using the AABB algorithm. Additionally, the correlation coefficient R^2 is calculated between the extracted and measured phenotypic traits. Furthermore, the root mean squared error (RMSE) is employed to assess errors in phenotypic traits. The RMSE is determined by equation (8):

$$RMSE = \sqrt{\sum_{i=1}^n (\hat{y}_i - y_i)^2} \quad (8)$$

where n is the number of samples, i is the current sample, \hat{y}_i is the predicted value, and y_i is the true value.

3. Result

3.1. Experimental environment

All experiments used Microsoft Windows 10 as the operating system.

Table 5
Hardware and software configuration for the experiments.

Configuration items	Item Value
Operating System	Microsoft Windows 10 Professional
CPU	AMD Ryzen 7 5800H with Radeon Graphics
GPU	GeForce RTX 3090
Available memory	24 GB
Programming Languages	Python(3.6.15)
Programming IDEs	Pycharm
Deep Learning Framework	Pytorch
Graphical interface framework	PYQT5

Table 5 shows the hardware and software configuration of the machines that ran the experiments.

3.2. Evaluation of plant image segmentation

3.2.1. Image background removal effect based on U^2 -net

Segmenting plant images is a pixel-level classification task that aims to separate plant pixels from background pixels. All pixels can be categorized into four groups based on the consistency between the predicted results and the true labels: TP represents the number of correctly predicted plant pixels, TN represents the number of correctly predicted background pixels, FP represents the number of background pixels wrongly predicted as plant pixels, and FN represents the number of plant pixels wrongly predicted as background pixels.

The performance of U^2 -net in segmenting plant images is assessed using four evaluation metrics: precision (Precision), recall (Recall), intersection over union (IoU), and maximum F measure ($\max F_\beta$).

- (1) The equation for the Precision is shown in equation (9), indicating the proportion of predicted plant pixels that are true plant pixels:

$$\text{Precision} = \frac{TP}{TP + FP} \quad (9)$$

- (2) The recall equation is shown in equation (10), which represents the proportion of correctly predicted pixels in the actual plant:

$$\text{Recall} = \frac{TP}{TP + FN} \quad (10)$$

- (3) The IOU equation is shown in equation (11), which indicates the ratio of the overlapping area of predicted and true labels to the merged area:

$$\text{IoU} = \frac{TP}{TP + FP + FN} \quad (11)$$

- (4) $\max F_\beta$ is used to indicate the summed average of the accuracy and recall rate, reflecting the overall effect of the segmentation algorithm, with the equation as in (12), $\beta = 0.3$:

$$\max F_\beta = \frac{(1 + \beta^2) \times \text{Precision} \times \text{Recall}}{\beta^2 \times \text{Precision} \times \text{Recall}} \quad (12)$$

The complexity of the plant can lead to different segmentation accuracy. Specifically, the cultivars with multiple tillers, leaves, and compact structures, such as NBC and HDB, pose challenges for segmentation compared to other varieties. Conversely, cultivars with fewer tillers, leaves, and simple structures, such as XXYH, LTYH, NBYH, and WT, are more amenable to accurate segmentation. **Table 6** and **Fig. 10**

Table 6
Carex cabbage and kale image segmentation results.

	HDB	NBC	NBYH	WT	XXYH	LTYH
Original image						
Segmentation map						

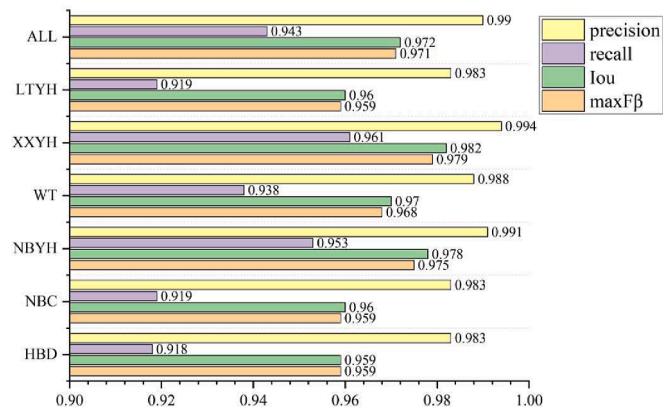


Fig. 10. The image segmentation model performance on different data sets.

present the final and average segmentation results of the U^2 -net model for various varieties on the test dataset. The final experimental results indicate that the U^2 -net model exhibits lower segmentation accuracy for NBC and HDB but higher segmentation accuracy for varieties such as XXYH, LTYH, NBYH, and WT. The average segmentation results demonstrate that the U^2 -net-based model performs as follows: precision is 0.99, recall is 0.943, IoU is 0.972, and $\max F_\beta$ is 0.971. These findings suggest that the segmentation model can accurately extract image features of carex cabbage and kale, resulting in more precise plant segmentation results.

3.2.2. Comparison of image background removal effects

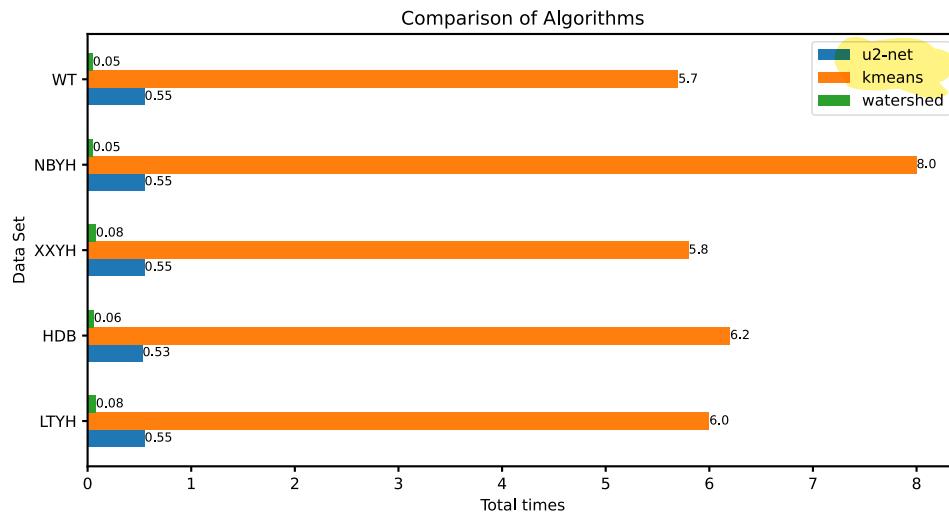
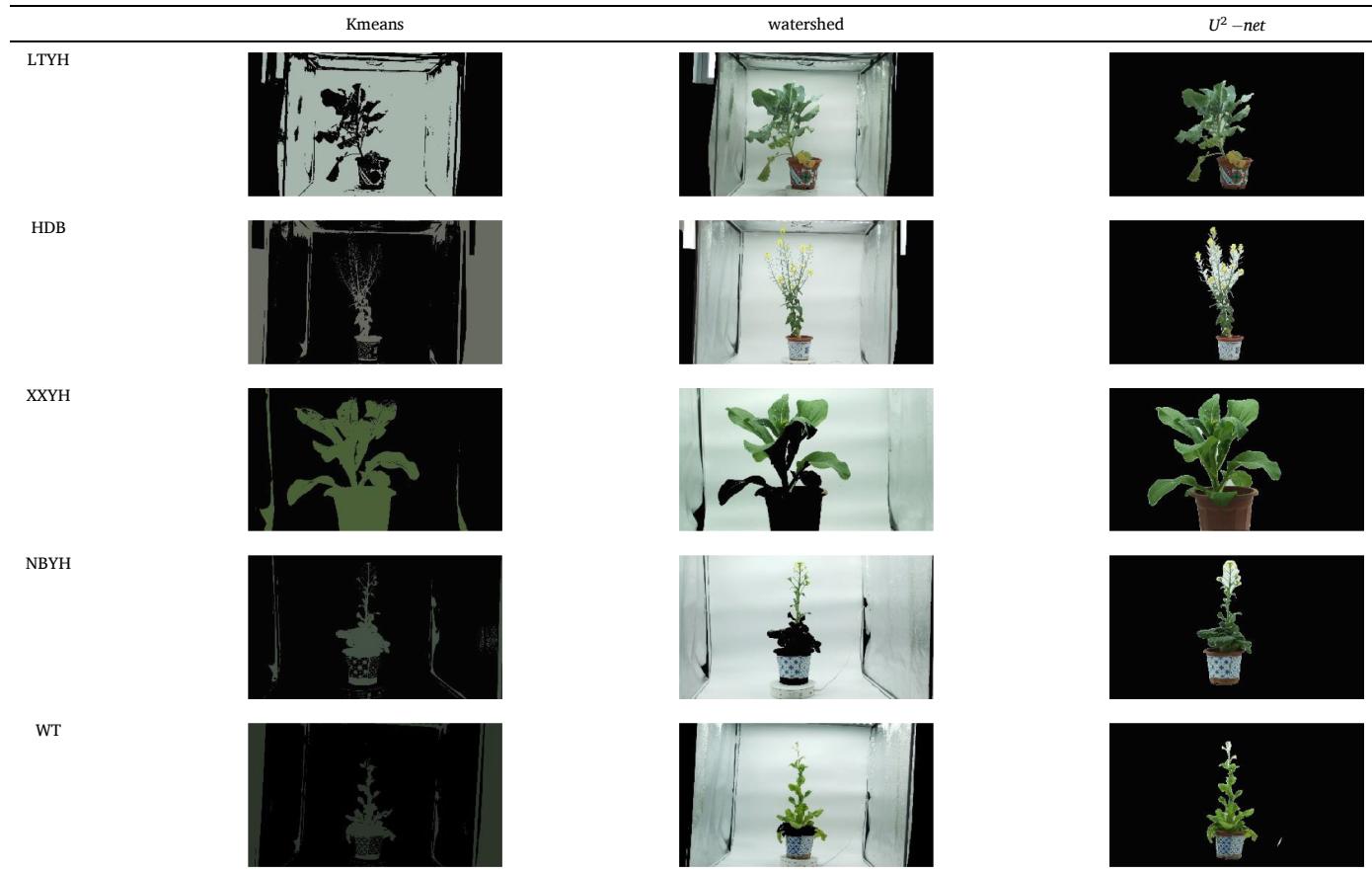
In order to validate the efficacy of image segmentation U^2 -net, five experimental materials are selected and subjected to comparison using the kmeans method, watershed method, and U^2 -net, respectively. The results of image segmentation are presented in **Table 7**.

In contrast to traditional methods, U^2 -net demonstrates efficient adaptability to diverse scenes through its ability to learn complex features and patterns. Unlike conventional approaches like watershed and k-means, U^2 -net obviates the need for manual parameter adjustments. It excels in capturing finer structures and delivering more precise segmentation results, particularly in handling complex textures and edges compared to traditional methods. Consequently, U^2 -net exhibits clear superiority in both efficiency and the outcomes of image processing.

Efficiency comparisons among different image segmentation methods are illustrated in **Fig. 11**. The image segmentation efficiency of U^2 -net falls between that of the k-means segmentation algorithm and the watershed algorithm, being lower than the latter and higher than the former. Notably, U^2 -net exhibits relatively consistent performance across all methods, with execution times hovering around 0.5 s. The algorithm's performance variation across different data levels is relatively minor, indicating good stability.

Table 7

Comparison of segmentation effect of different image segmentation algorithms.

**Fig. 11.** Comparison of efficiency of different image segmentation algorithms.

The occurrence of white noise in plant images resulting from $U^2 - net$ segmentation is attributed to the complex spatial structure and shading in leafy vegetable plants. Despite the presence of white dots, their quantity and spatial extent are rarity and isolation, thus most of white dots cannot be matched and will be removed by following comprehensive denoising in the 3D reconstruction process.

3.3. Visualization of plant point cloud

The reconstruction and denoising results of the plant point cloud are presented in Table 8. The experimental plants have different tiller and strain characteristics. Specifically, LTYH is expansive, HDB and NBC are loose, and NBYH and WT are compact. Moreover, the point cloud visualization demonstrated the proposed reconstruction algorithm's ability to obtain point clouds with clear structures and realistic colors for diverse plant models. In addition, the denoising method successfully

Table 8

Reconstruction and denoising results of carex cabbage and kale.

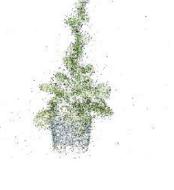
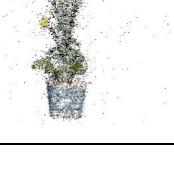
Cultivar	Sparse point cloud	Dense point cloud	Point cloud with color filtering (black)	Point cloud with color filtering (white)	Point cloud with Euclidean clustering
LTYH					
HDB					
WT					
NBYH					
NBC					

Table 9

Extraction and denoising results of pot point cloud.

Algorithm	LTYH	HDB	NBYH	NBC
Extraction of pot point cloud				
Denoising pot point cloud				

eliminated the black noise at the leaf edges, the white noise between stalks, and the scattered points around the plants, all while preserving the 3D information of the plants.

3.4. Size calibration of plant point cloud

3.4.1. Extraction and denoising of pot

The results of extracting the pot point cloud from the plant point cloud by color filtering and statistical filtering are presented in [Table 9](#). In essence, the figure demonstrates that color filtering accurately captures the general outline of the pot, and statistical filtering effectively denoises the discrete points in the point cloud.

3.4.2. Orientation calibration of plant

The results of fitting a horizontal plane to the pot point cloud by the RANSAC and Rodrigues' rotation matrix are presented in [Table 10](#). Indeed, the figure illustrates that the plane fitting based on RANSAC accurately identifies the plane of the pot point cloud. In addition, the orientation correction based on Rodrigues' formula effectively aligns the parallel of the plant point cloud to the horizontal plane.

3.4.3. Size calibration of point cloud

Due to the low survival rate of plants, only thirteen plants are selected for the size calibration experiments: one HDB plant, five NBC plants, three NBYH plants, and four LTYH plants. The experimental results are presented in [Table 11](#). Consequently, the dimensions of the pot point cloud varied significantly and deviated from the actual dimensions in [Table 1](#) under different acquisition scenarios. Furthermore, the scale factor also differed across scenes. After size calibration, the point cloud dimensions of the pots exhibited improved proximity to the actual values in [Table 1](#), indicating the effectiveness of the proposed calibration method.

3.5. Accuracy of 3D reconstruction algorithm

3.5.1. Accuracy of plant phenotypic traits

Phenotypic traits, including plant height and width, are extracted from the reconstructed point cloud, and the correlation is analyzed with the actual plant dimensions shown in [Fig. 12](#). Notably, the error analysis

results demonstrated that the plant height had an R^2 value of 0.999 and an RMSE of 0.298 cm, while the plant width had an R^2 value of 1.000 and an RMSE of 0.338 cm. Significantly, the R^2 values of plant height and width exceeded 0.99, indicating that the proposed 3D reconstruction method for complex plants accurately achieved the phenotypic information and strongly correlated with the actual data.

3.5.2. Accuracy of IVOP&AKAZE-SFM algorithm

The proposed IVOP&AKAZE-SFM algorithm is compared to three incremental structure from motion algorithms, including Colmap, OpenSfM, and OpenMVG, on Carex cabbage and kale plants to evaluate the effectiveness of reconstruction. Mainly, several metrics are tested to evaluate the performance of each algorithm from both subjective and objective perspectives.

Subjective indicators are employed to assess the integrity of the reconstructed point cloud, the distribution of outlier points, and the drift of the scene. The experimental results are presented in [Table 12](#).

The figure illustrates that the proposed IVOP&AKAZE-SFM algorithm reconstructs a more complete and detailed point cloud for all plants compared to other algorithms. Specifically, in the case of plants with a small number of tillers and simple morphology (e.g., LTYH, NBYH, and WT), all four algorithms exhibit reliable performance in reconstructing the contours and color of plants. Significantly, the proposed IVOP&AKAZE-SFM algorithm excels in extracting features of complex plant morphology, including more apparent flower contours during the flowering stage and more comprehensive reconstruction of tiny leaves. On the contrary, in the case of plants with complex morphology, slender stalks, and severe obstruction between blades (e.g., HDB and NBC), there are notable differences in the reconstruction results among the four algorithms. For instance, the OpenSfM and Colmap algorithms caused scene drift and point cloud distortion due to the accumulation of errors. Moreover, the point clouds reconstructed by the OpenMVG algorithm could not recognize the plant's stem structure because the point cloud is too sparse. In contrast, the IVOP&AKAZE-SFM algorithm is able to reconstruct rich, detailed features and reduce the accumulated errors, to maintain a better reconstruction quality, attributed to the ability to extract more plant features utilizing AKAZE features and minimize reconstruction errors through the precise initialization of internal camera parameters and distortion coefficients.

Table 10

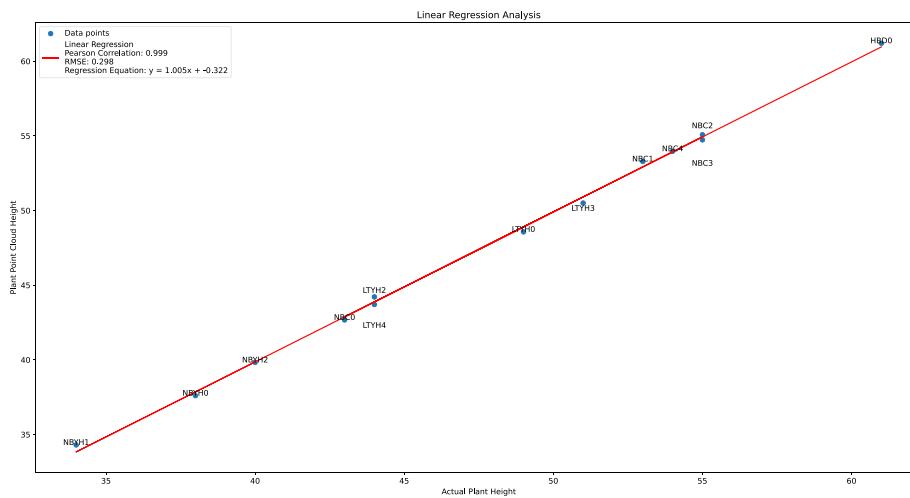
Orientation calibration results of plant point cloud.

	LTYH	HDB	NBC	NBYH
before Orientation calibration				
Plane fitting				
after Orientation calibration				

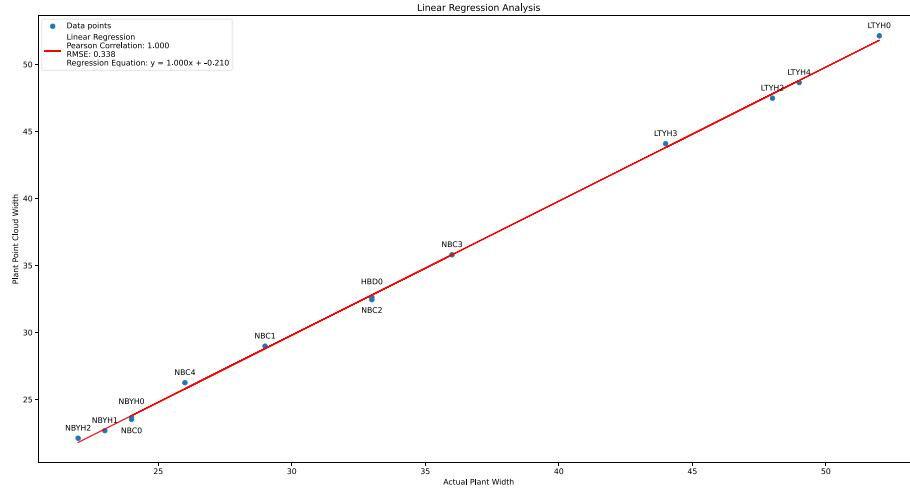
Table 11

Size calibration results of point cloud.

Types of pot point cloud	Dimensions of the pot point cloud before calibration			Scale factor of pot point cloud			Dimensions of the pot point cloud after calibration		
	X-axis	Y-axis	Z-axis	S_x	S_y	S_z	X-axis	Y-axis	Z-axis
HBD0	0.19	0.20	0.16	84.50	81.04	78.47	16	16	12.7
NBC0	0.09	0.10	0.07	172.04	169.29	161.45	16	16	12.3
NBC1	0.27	0.27	0.24	59.22	58.48	52.39	16	16	12.3
NBC2	0.26	0.27	0.23	60.75	59.86	54.27	16	16	12.3
NBC3	0.31	0.33	0.27	51.03	49.12	46.24	16	16	12.3
NBC4	0.20	0.21	0.18	80.38	78.08	67.42	16	16	12.3
NBYH0	0.26	0.26	0.21	69.86	70.28	69.82	18	18	14.5
NBYH1	0.27	0.26	0.21	67.98	69.43	68.42	18	18	14.5
NBYH2	0.34	0.29	0.24	52.84	62.33	61.50	18	18	14.5
LTYH0	0.32	0.35	0.27	67.95	62.07	66.73	22	22	18
LTYH2	0.28	0.27	0.21	80.02	80.92	84.44	22	22	18
LTYH3	0.30	0.29	0.23	72.96	75.94	78.83	22	22	18
LTYH4	0.35	0.33	0.30	63.80	66.61	59.67	22	22	18



(a) Correlation analysis between calculated and measured value of plant heights



(b) Correlation analysis between calculated and measured value of plant widths

Fig. 12. Correlation and RMSE analysis results.

In conclusion, the proposed IVOP&AKAZE-SFM algorithm performs superior reconstruction on various plants.

Objective metrics evaluate the performance of reconstruction algorithms by considering the number of points, the average track length, the number of registered images, the average reprojection error and time efficiency. Fig. 13 shows the result of the four algorithms' average

performance on these metrics.

The figure unequivocally illustrates the exceptional performance of the IVOP&AKAZE-SFM algorithm across a comprehensive range of evaluated metrics. Remarkably, the algorithm excels in point generation, surpassing its counterparts in both the extraction and reconstruction of features. As delineated in Fig. 13(a), IVOP&AKAZE-SFM

Table 12

Reconstruction results of different algorithms.

Algorithm	LTYH	HDB	NBC	NBYH	WT
OpenSfM					
Colmap					
OpenMVG					
IVOP&AKAZE-SFM					

outperforms OpenSfM by 169.19 %, Colmap by 386.11 %, and OpenMVG by an impressive 452.54 %.

Additionally, Fig. 13(b) highlights the algorithm's achievement of a minimum average track length, signifying its exceptional ability to track feature points with substantial depth variations. Concerning the number of registered images, as depicted in Fig. 13(c), the IVOP&AKAZE-SFM algorithm is comparable to OpenSfM and OpenMVG, yet superior to Colmap. This showcases its adeptness in effectively utilizing image information for reconstruction.

Of paramount importance, the IVOP&AKAZE-SFM algorithm exhibits a lower average reprojection error than its counterparts, as shown in Fig. 13(d), emphasizing its precision in reconstructing 3D point cloud models. Specifically, the percentage of error reduction for IVOP&AKAZE-SFM relative to OpenSfM is 58.05 %, compared to 55.94 % for Colmap, and 60.20 % for OpenMVG.

The reduction in time consumption achieved by the IVOP&AKAZE-SFM method, presented in Fig. 13(e), is substantial. Specifically, IVOP&AKAZE-SFM reduces time consumption by 44.77 % relative to OpenSfM, 48.40 % relative to Colmap, and 38.42 % relative to OpenMVG. These findings underscore the high time efficiency of the algorithm proposed in this paper.

In summary, the proposed IVOP&AKAZE-SFM algorithm demonstrates superior reconstruction performance across various plant cultivars, affirming its efficacy and efficiency in 3D reconstruction applications.

4. Discussion

In the reconstruction of point clouds for multi-stage plants, critical parameters will impacting the reconstruction accuracy, such as the number of cameras, lens angles, and the quantity of captured images, are carefully considered (Wang et al., 2020). To strike a balance between image quality and processing efficiency, we employ an adaptive strategy that adjusts the number of cameras and shooting angles based on the

plant's size and structure. For this study, a horizontal viewing angle lens is chosen to capture multi-view images of cabbage shoots and kale plants, effectively covering the entire plant. At the same time, to address the challenge of not obtaining clear images of the upward-stretching leaves, we propose the addition of an overhead 45° viewing angle to encompass a larger shot area and diminish occlusion between leaves, ultimately enhancing the quality of point cloud reconstruction.

Throughout the rotational multi-view image acquisition process of plants, jitter often raised by rotary table's drive, emerges as a primary error source in rotational multi-view point cloud reconstruction. This jitter may impede image alignment and matching, consequently diminishing the accuracy of point cloud reconstruction. To effectively mitigate jitter errors, this paper employs the strategy of capturing neighboring images with a 10-second interval following the cessation of plant rotation. This approach successfully circumvents errors produced by jitter and contributes to the generation of a highly accurate point cloud model of the plant. Also, the incorporation of image stabilization techniques, such as digital filtering or hardware stabilization devices, can be contemplated to further attenuate the impact of jitter on point cloud reconstruction.

For tall plants, the positional mitigation induced by the rotation of the turntable is effectively eliminated through image segmentation in sec2.3.4. Background removal enables the reconstruction algorithm to concentrate on the plant, thereby minimizing the influence of positional changes on accuracy. Furthermore, the inherent substantial plant weight contributes to its resistance to positional movement, resulting in minimal impact on reconstruction accuracy.

Comparing the 3D reconstruction algorithm proposed in this paper with deep learning-based counterparts in reconstructing high-precision plant point cloud models reveals the following advantages and disadvantages. In terms of generalization ability, the proposed algorithm outperforms deep learning algorithms when faced with unseen objects or fewer images, regardless of the dataset. Conversely, deep learning algorithms exhibit strong performance on training sets but encounter

challenges in generalization, particularly when reconstructing unseen objects accurately. Regarding finesse, our proposed algorithm performs well in reconstructing complex plant structures, capturing the overall structure effectively. However, there are some limitations in handling finer details. On the contrary, deep learning algorithms have made significant progress in improving the accuracy of 3D reconstruction, but

have lower performance in dealing with specific details. In terms of algorithmic efficiency, deep learning grapples with the trade-off between achieving high-resolution reconstruction accuracy in an end-to-end 3D reconstruction method and managing memory and time costs. Traditional algorithms, on the other hand, do not incur the time and space costs associated with pre-training. Therefore, in practical

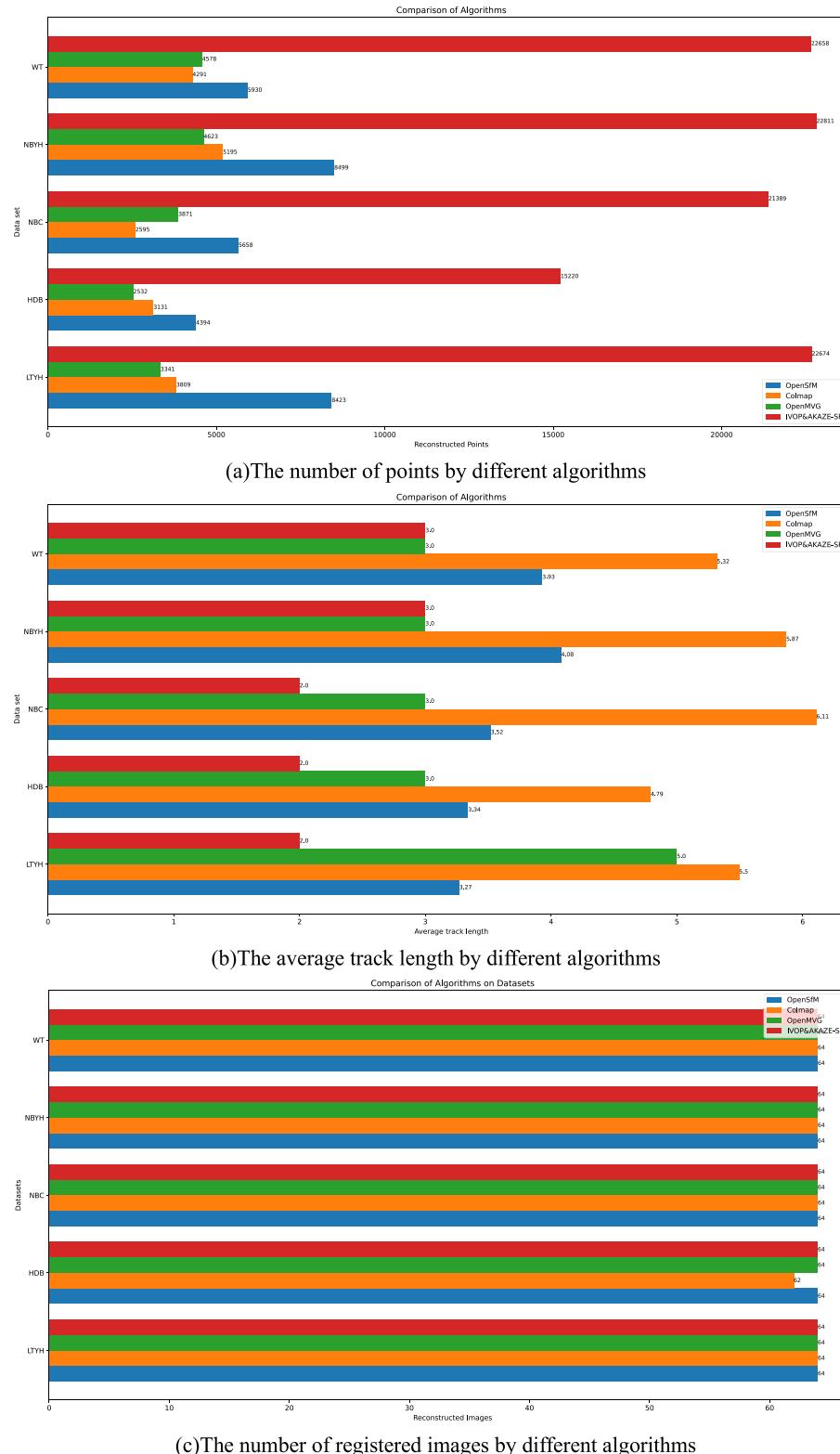
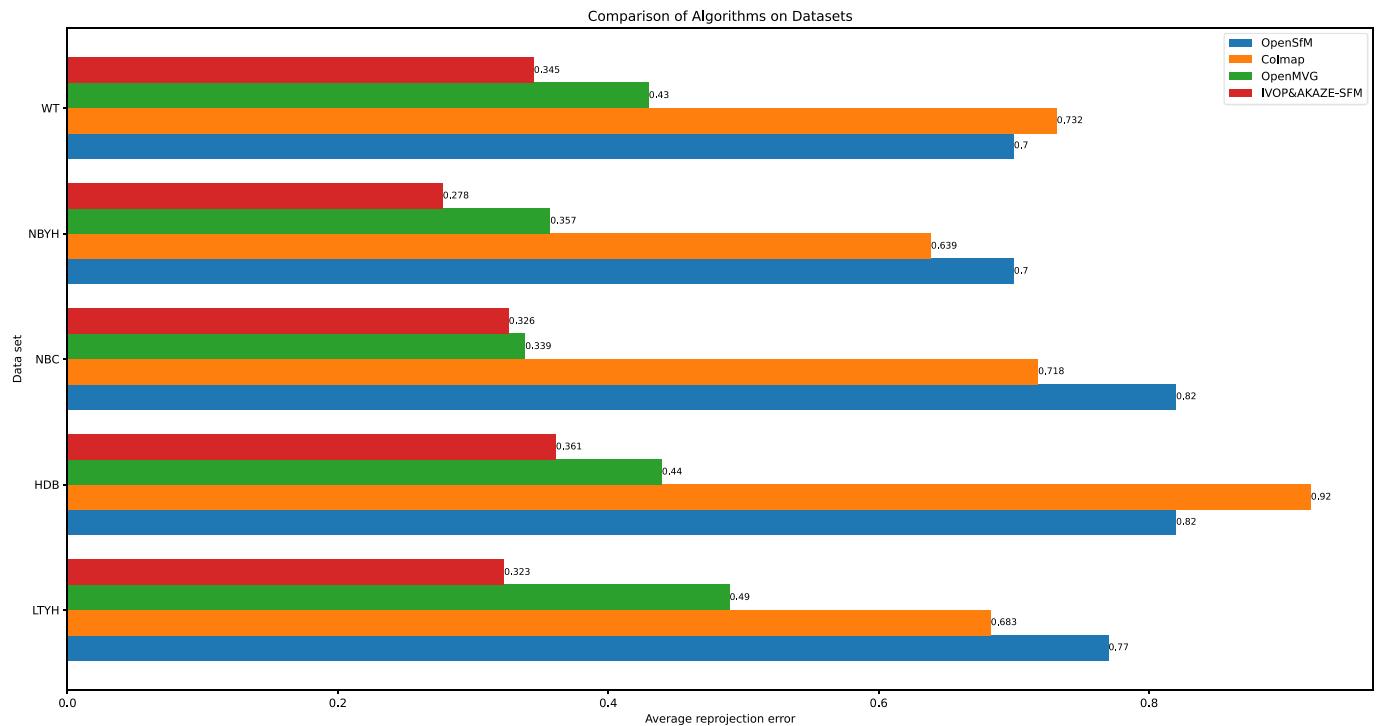
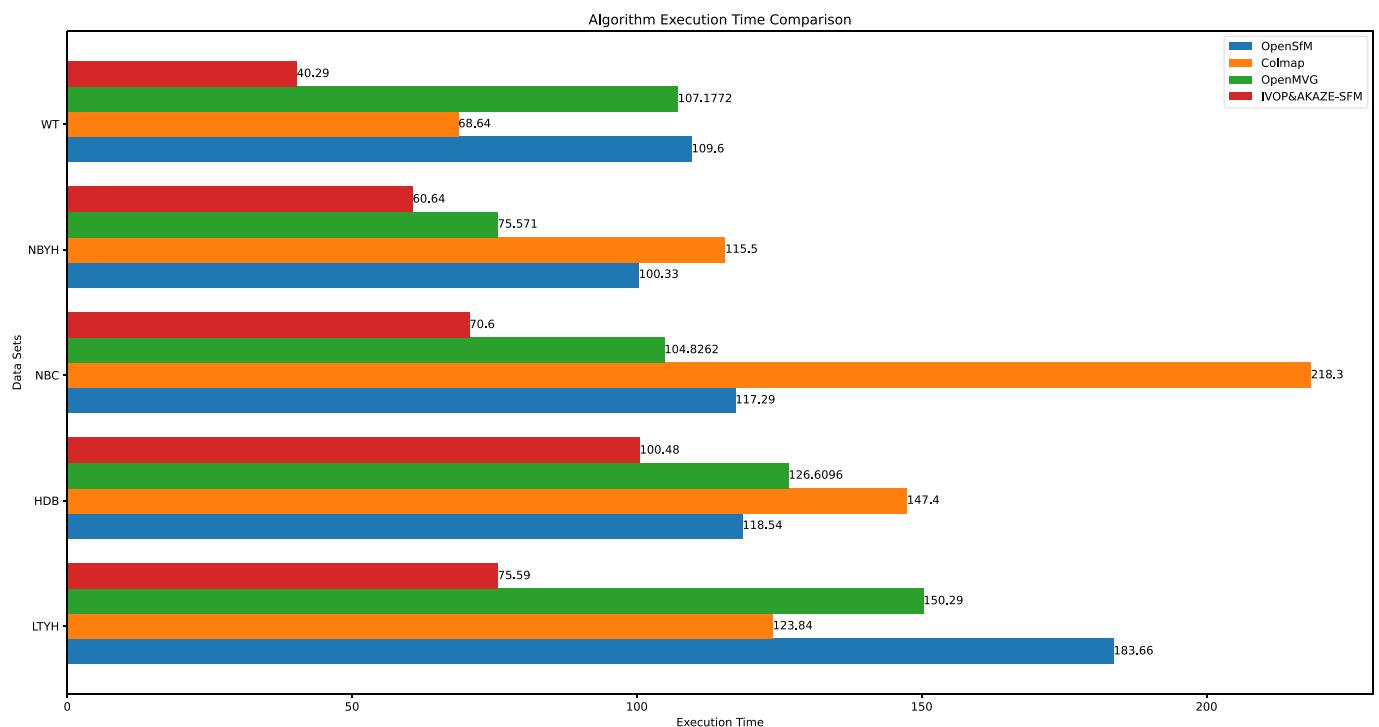


Fig. 13. Comparison of objective metrics for different algorithms on different data sets.



(d) The reprojection error by different algorithms



(e) Time efficiency of different algorithms

Fig. 13. (continued).

applications, our algorithm may present advantages in certain scenarios.

5. Conclusion

A 3D phenotyping platform based on multi-view images is proposed to reconstruct high-precision proportional and clean point cloud mode for complex plants, comprising five steps: image acquisition, image segmentation, 3D reconstruction, point cloud denoising, and calibration. The experiments on carex cabbage and kale show that, U^2 -net can segment plant images accurately. Furthermore, the proposed 3D reconstruction method for complex plants, based on IVOP&AKAZE-SFM and multi-view stereo, can reconstruct high-precision point clouds while avoiding scene drift and plant distortion, representing a superiority over the current incremental 3D reconstruction algorithms. Moreover, by combining color filtering and Euclidean clustering, noise can be effectively removed while preserving the integrity of the plant point cloud, thereby enhanced the quality and clarity of the point cloud. Also, by utilizing a priori information of the pot, the scale restoration method can accurately calibrate the size of plant point cloud, leading to a robust linear relationship between the extracted phenotype and measured traits, with a RMSE of 0.298 cm and 0.338 cm.

The proposed phenotyping platform can be utilized in high-throughput phenotypic studies of plants, including rice, sugar beet, and other crops. However, it cannot extract multiple phenotypic traits of carex cabbage and kale plants throughout their growth. In the future, the method combined with deep learning for extracting phenotypic traits throughout plant growth will be studied to enhance the comprehensiveness and intelligence of plant phenotyping technologies.

CRediT authorship contribution statement

Danni Yang: Conceptualization, Formal analysis, Methodology, Project administration, Validation, Writing – original draft, Writing – review & editing. **Huijun Yang:** Conceptualization, Funding acquisition, Validation, Writing – review & editing. **Dongfeng Liu:** Conceptualization, Writing – review & editing. **Xianlin Wang:** 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

Data will be made available on request.

Acknowledgement

Research supported by Foundation of Key Research and Development Program of Shaanxi province (2023-YBNY-229, 2024NC-YBXM-197), Undergraduate Training Program for Innovation and entrepreneurship plan (202210712189).

References

- Alcantarilla, P.F., Solutions, T.J.I.T.P.A.M.I., 2011. Fast explicit diffusion for accelerated features in nonlinear scale spaces. *J. Nonlinear Sci.* 21, 1281–1298. doi: 10.5244/C.27.13.
- Crowell, S., Falcao, A.X., Shah, A., Wilson, Z., Greenberg, A.J., McCouch, S.R., 2014. High-resolution inflorescence phenotyping using a novel image-analysis pipeline, PANorama. *Plant Physiol.* 165, 479–495. https://doi.org/10.1104/pp.114.238626.
- Duan, T., Chapman, S.C., Holland, E., Rebetzke, G.J., Guo, Y., Zheng, B., 2016. Dynamic quantification of canopy structure to characterize early plant vigour in wheat genotypes. *J. Exp. Bot.* 67, 4523–4534. https://doi.org/10.1093/jxb/erw227.
- Fiorani, F., Schurr, U., 2013. Future scenarios for plant phenotyping. *Annu. Rev. Plant Biol.* 64, 267–291. https://doi.org/10.1146/annurev-arplant-050312-120137.
- Forero, M.G., Murcia, H.F., Mendez, D., Betancourt-Lozano, J., 2022. LiDAR platform for acquisition of 3D plant phenotyping database. *Plants (Basel)* 11, 2199. https://doi.org/10.3390/plants11172199.
- Furukawa, Y., Ponce, J., 2010. Accurate, dense, and robust multiview stereopsis. *IEEE Trans. Pattern Anal. Mach. Intell.* 32, 1362–1376. https://doi.org/10.1109/tpami.2009.161.
- Guo, Z., Liu, H., Shi, H., Li, F., Guo, X., Cheng, B., 2023. KD-tree-based Euclidean clustering for tomographic SAR point cloud extraction and segmentation. *IEEE Geosci. Remote Sens. Lett.* 20, 1–5. https://doi.org/10.1109/lgrs.2023.3234406.
- Han, X.-F., Jin, J.S., Wang, M.-J., Jiang, W., Gao, L., Xiao, L., 2017. A review of algorithms for filtering the 3D point cloud. *Signal Process. Image Commun.* 57, 103–112. https://doi.org/10.1016/j.image.2017.05.009.
- He, J.Q., Harrison, R.J., Li, B., 2017. A novel 3D imaging system for strawberry phenotyping. *Plant Methods* 13, 93. https://doi.org/10.1186/s13007-017-0243-x.
- Liu, S., Acosta-Gamboa, L., Huang, X., Lorence, A., 2017. Novel low cost 3D surface model reconstruction system for plant phenotyping. *Journal of Imaging* 3, 39. https://doi.org/10.3390/jimaging3030039.
- Martinez-Guanter, J., Ribeiro, A., Peteinatos, G.G., Perez-Ruiz, M., Gerhards, R., Bengoechea-Guevara, J.M., Machleb, J., Andujar, D., 2019. Low-cost three-dimensional modeling of crop plants. *Sensors (Basel)* 19, 2883. https://doi.org/10.3390/s19132883.
- Minervini, M., Giuffrida, M.V., Perata, P., Tsafaris, S.A., 2017. Phenotiki: an open software and hardware platform for affordable and easy image-based phenotyping of rosette-shaped plants. *Plant J.* 90, 204–216. https://doi.org/10.1111/tpj.13472.
- Moulou, P., Monasse, P., Marlet, R., 2013. Adaptive structure from motion with a Contrario model estimation. Springer, Berlin Heidelberg, Berlin, Heidelberg, pp. 257–270. 10.1007/978-3-642-37447-0_20.
- Moulou, P., Monasse, P., Perrot, R., Marlet, R., 2017. OpenMVG: open multiple view geometry. Springer International Publishing, Cham, pp. 60–74. 10.1007/978-3-319-56414-2_5.
- Parent, B., Shahinnia, F., Maphosa, L., Berger, B., Rabie, H., Chalmers, K., Kovalchuk, A., Langridge, P., Fleury, D., 2015. Combining field performance with controlled environment plant imaging to identify the genetic control of growth and transpiration underlying yield response to water-deficit stress in wheat. *J. Exp. Bot.* 66, 5481–5492. https://doi.org/10.1093/jxb/erv320.
- Paulus, S., Schumann, H., Kuhlmann, H., Leon, J., 2014. High-precision laser scanning system for capturing 3D plant architecture and analysing growth of cereal plants. *Biosyst. Eng.* 121, 1–11. https://doi.org/10.1016/j.biosystemseng.2014.01.010.
- Piersuschka, R., Schurr, U., 2019. Plant phenotyping: past, present, and future. *Plant Phenomics* 2019, 7507131. https://doi.org/10.34133/2019/7507131.
- Qin, X., Zhang, Z., Huang, C., Dehghan, M., Zaiane, O.R., Jagersand, M., 2020. U2-net: going deeper with nested U-structure for salient object detection. *Pattern Recogn.* 106, 107404. https://doi.org/10.1016/j.patcog.2020.107404.
- Rossi, R., Leolini, C., Costafreda-Aumedes, S., Leolini, L., Bindu, M., Zaldei, A., Moriondo, M., 2020. Performances evaluation of a low-cost platform for high-resolution plant phenotyping. *Sensors (Basel)* 20, 3150. https://doi.org/10.3390/s20113150.
- Sandhu, J., Zhu, F., Paul, P., Gao, T., Dhatt, B.K., Ge, Y., Staswick, P., Yu, H., Walia, H., 2019. PI-plat: a high-resolution image-based 3D reconstruction method to estimate growth dynamics of rice inflorescence traits. *Plant Methods* 15, 162. https://doi.org/10.1186/s13007-019-0545-2.
- Sunvittayakul, P., Kittipadakul, P., Wonnapinij, P., Chanchay, P., Wannitkul, P., Sathithnaitham, S., Phanthanong, P., Changwitsukarn, K., Suttangkakul, A., Ceballos, H., Vuttipongchaikij, S., 2022. Cassava root crown phenotyping using three-dimension (3D) multi-view stereo reconstruction. *Sci. Rep.* 12, 10030. https://doi.org/10.1038/s41598-022-14325-4.
- Tausen, M., Clausen, M., Moeskjaer, S., Shihavuddin, A., Dahl, A.B., Janss, L., Andersen, S.U., 2020. Greenotyper: image-based plant phenotyping using distributed computing and deep learning. *Front. Plant Sci.* 11, 1181. https://doi.org/10.3389/fpls.2020.01181.
- Teng, X., Zhou, G., Wu, Y., Huang, C., Dong, W., Xu, S., 2021. Three-dimensional reconstruction method of rapeseed plants in the whole growth period using RGB-D Camera. *Sensors (Basel)* 21, 4628. https://doi.org/10.3390/s21144628.
- Thapa, S., Zhu, F., Walia, H., Yu, H., Ge, Y., 2018. A novel LiDAR-based instrument for high-throughput, 3D measurement of morphological traits in maize and Sorghum. *Sensors (Basel)* 18, 1187. https://doi.org/10.3390/s18041187.
- Torrabal, A., Russell, B.C., Yuen, J., 2010. LabelMe: online image annotation and applications. *Proc. IEEE* 98, 1467–1484. https://doi.org/10.1109/JPROC.2010.2050290.
- Wang, Y., Chen, Y., 2020. Non-destructive measurement of three-dimensional plants based on point cloud. *Plants (Basel)* 9, 571. https://doi.org/10.3390/plants9050571.
- Wang, R., Liu, D., Wang, X., Yang, H., 2022b. Multi-view geometry-based cabbage shoots segmentation and key phenotype measurement. *Transactions of the Chinese Society of Agricultural Engineering* 38, 243. https://doi.org/10.11975/j.issn.1002-6819.2022.16.027 (In Chinese with English abstract).
- Wang, F., Ma, X., Liu, M., Wei, B., 2022a. Three-dimensional reconstruction of soybean canopy based on multivision technology for calculation of phenotypic traits. *Agronomy* 12, 692. https://doi.org/10.3390/agronomy12030692.
- Wang, J., Zhang, Y., Gu, R., 2020. Research status and prospects on plant canopy structure measurement using visual sensors based on three-dimensional reconstruction. *Agriculture* 10, 462. https://doi.org/10.3390/agriculture10100462.
- Watanabe, R., Nonaka, K., Pavez, E., Kobayashi, T., Ortega, A., 2023. Graph-based point cloud color denoising with 3-dimensional patch-based similarity. In: ICASSP 2023–2023 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), pp. 1–5.

- Wu, S., Wen, W., Wang, Y., Fan, J., Wang, C., Gou, W., Guo, X., 2020. MVS-pheno: a portable and low-cost phenotyping platform for maize shoots using multiview stereo 3D reconstruction. *Plant Phenomics* 2020, 1848437. <https://doi.org/10.34133/2020/1848437>.
- Wu, S., Wen, W., Gou, W., Lu, X., Zhang, W., Zheng, C., Xiang, Z., Chen, L., Guo, X., 2022. A miniaturized phenotyping platform for individual plants using multi-view stereo 3D reconstruction. *Front Plant Sci* 13, 897746. <https://doi.org/10.3389/fpls.2022.897746>.
- Xu, N., Sun, G., Bai, Y., Zhou, X., Cai, J., Huang, Y., 2023. Global reconstruction method of maize population at seedling stage based on kinect sensor. *Agriculture* 13, 348. <https://doi.org/10.3390/agriculture13020348>.
- Yang, Z., Han, Y., 2020. A low-cost 3D phenotype measurement method of leafy vegetables using video recordings from Smartphones. *Sensors (basel)* 20, 6068. <https://doi.org/10.3390/s20216068>.
- Zhao, C., Zhang, Y., Du, J., Guo, X., Wen, W., Gu, S., Wang, J., Fan, J., 2019. Crop phenomics: current status and perspectives. *Front Plant Sci* 10, 714. <https://doi.org/10.3389/fpls.2019.00714>.
- Zhengyou, Z., 1999. Flexible camera calibration by viewing a plane from unknown orientations. *Proceedings of the Seventh IEEE International Conference on Computer Vision* 661, 666–673. <https://doi.org/10.1109/ICCV.1999.791289>.