

Three-dimensional reconstruction of cotton plant with internal canopy occluded structure recovery



Yang Li^{a,b,c,d}, Shuke Si^a, Xinghua Liu^{a,c,*}, Liangliang Zou^{a,c}, Wenqian Wu^a, Xuemei Liu^{a,c}, Li Zhang^{e,*}

^a College of Mechanical & Electronic Engineering, Shandong Agricultural University, Tai'an 271018, China

^b Key Laboratory of Modern Agricultural Equipment, Ministry of Agriculture and Rural Affairs, Nanjing 210014, China

^c Shandong Provincial Key Laboratory of Horticultural Machinery and Equipment, Tai'an 271018, China

^d Shandong Provincial Engineering Laboratory of Agricultural Equipment Intellectualisation, Tai'an 271018, China

^e College of Resources, Shandong University of Science and Technology, Tai'an 271018, China

ARTICLE INFO

Keywords:

Inner occluded structure
Instance segmentation
Generative adversarial network
Three-dimensional reconstruction of plants

ABSTRACT

The inner leaves of crop canopies are obscured by outer branches and leaves, leading to information loss and observation difficulty of the occluded canopy structure using modern crop monitoring techniques. It has restricted the development of phenotypic analysis and precision agriculture. In this paper, we propose a neural network approach to reconstruct the occluded structure of crop canopies with an RGB-D sensor. Taking the cotton plant as the object of study, we propose a novel Cascade Leaf Segmentation and Completion Network (CLSCN) to reconstruct the occluded leaf images and propose a Fragmental Leaf Point-cloud Reconstruction Algorithm (FLPRA) to complete the missing point clouds. By combining the Instance Segmentation Network (ISN), Generative Adversarial Network (GAN) and Point-cloud Reconstruction Algorithm (PRA), the three-dimensional models of cotton plants with both completed internal and external structures of the canopy are smoothly reconstructed. Firstly, we collect a large number of leaf images and point clouds of cotton plants using an RGB-D sensor with the top view and construct a manually labeled cotton leaf dataset for training and evaluation. Secondly, a network named CLSCN is cascading constructed with an Instance Segmentation Network (ISN) and a Generative Adversarial Network (GAN), and the two parts of CLSCN are separately trained with our constructed dataset to output complete cotton leaves. Thirdly, with the fusion of the completed RGB images output by cascaded network segmentation and the point clouds captured by RGB-D sensor, the proposed FLPRA is used to filter, reconstruct, fuse and register the cotton canopy leaf point clouds, and to obtain the whole cotton canopy point-clouds with inner occluded structure recovery. Finally, the CLSCN and FLPRA are validated using the validation dataset of cotton leaf. The test results indicate that the front-end ISN of the proposed CLSCN can generate high-quality cotton leaf masks, with FID scores less than 35 and mIoU up to 84.65%. Additionally, the back-end GAN of CLSCN can complete the occluded leaves with an accuracy of over 94%. The reconstruction accuracy of the final three-dimensional model of the cotton canopy is as high as 82.70%. Therefore, the proposed neural network and algorithm effectively solve the problem of incomplete canopy point cloud caused by the occlusion of outer leaves and provide an effective way to recover the complete three-dimensional structure of crop canopy with internal occlusion. It is a meaningful theoretical and technical support to realize real-time crop status observation and precise field management in agriculture production.

1. Introduction

The structure of the canopy can fully reflect the growth status of crops, and it has been the data basis for precise management in modern

agriculture (Nguyen et al., 2015; Wang et al., 2020; ZHU Rongsheng, 2021). Nowadays, the three-dimensional reconstruction technology of crop canopies has been a research hotspot in the field of agricultural information. Recent studies show that with the reconstruction methods

* Corresponding authors at: College of Mechanical & Electronic Engineering, Shandong Agricultural University, Tai'an 271018, China (X. Liu). College of Resources, Shandong University of Science and Technology, Tai'an 271018, China (L. Zhang).

E-mail addresses: lkh935@163.com (X. Liu), skdzl1984@163.com (L. Zhang).

such as Structure from Motion (Dugdale et al., 2019), Multiple View Fusion (Zhu et al., 2023b), Point-cloud Merging (Pagliai et al., 2022), etc., the surface information of crop canopy can be effectively reconstructed (Fang et al., 2016; Okura, 2022). However, due to the mutual occlusions among the branches and leaves in the canopy, these existing methods cannot directly reconstruct the occluded structure of the canopy from the outside. It causes the incomplete expression of the three-dimensional model of crops with internal feature loss and restricts the development of digital agriculture.

Currently, the information collection of crop canopies is mainly achieved by some non-contact sensors such as industrial cameras, depth cameras, structured light sensors, LiDARs, ultrasonic sensors, and others. Benefiting from the sustained efforts of relevant scholars, many valuable achievements have been made in the capture and reconstruction of external information on the crop canopy. K. Yin proposed an intrusive acquisition method to reconstruct canopies, by disassembling the plant into disjoint parts that can be accurately scanned and reconstructed offline, they used the reconstructed part meshes as three-dimensional proxies for the reconstruction of the complete plant and devise a global-to-local non-rigid registration technique that preserves specific plant characteristics (Yin et al., 2016). B.X. Wei constructed a multi-source image synchronous acquisition platform for soybean canopy based on Kinect 2.0 to obtain the canopy image data, proposed a calculation method of LAI for the soybean canopy based on three-dimensional reconstruction and established the dynamic simulation model of canopy LAI (Wei et al., 2023). T. Zhu used three Kinect 2.0 sensors to acquire a full range point cloud of tomato canopy at crucial growth stages and proposed a method for high-throughput detection of tomato canopy's phenotypic traits based on three-dimensional structure reconstruction with multi-views (Zhu et al., 2023a). Z.Z. Song adopted DeepLabV3 + to segment the fruit calyx, branch, and wire and developed a method of discrete wire pixels reconstruction on Progressive probabilistic Hough transform (PPHT) to help sense the distribution of the wire (Song et al., 2021). R. Giuliani designed a simulation study to provide the statistical information necessary to drive and control a foliage random sampling process on a real tree canopy and proposed a point-intercept leaf collection method simulated on a single walnut-tree crown by performing a Monte Carlo (MC) sampling from a data set obtained (Giuliani et al., 2005). D. Rapado-Rincón introduced a novel approach for building generic representations in occluded agro-food environments using multi-view perception and three-dimensional multi-object tracking, based on a detection algorithm that generates partial point clouds for each detected object, followed by a three-dimensional multi-object tracking algorithm that updates the representation over time (Rapado-Rincón et al., 2023). H. Cuevas-Velasquez proposed a method responsible for performing the segmentation of the branches and recovering their morphology in three-dimension, and the obtained reconstruction allows the manipulator of the robot to select the candidate branches to be pruned (Cuevas-Velasquez et al., 2020). J. A. Gibbs presented an active vision-based pipeline which aims to contribute to reducing the bottleneck associated with the phenotyping of architectural traits, providing a fully automated response to photometric data acquisition and the recovery of three-dimensional models of plants without the dependency on botanical expertise, and proposed a novel surface reconstruction algorithm (Gibbs et al., 2020). S. Das Choudhury introduced a novel algorithm to compute three-dimensional plant phenotypes from multi-view images using voxel-grid reconstruction of the plant and also proposed a novel method to reliably detect and separate the individual leaves and the stem from the three-dimensional voxel-grid of the plant using voxel overlapping consistency check and point cloud clustering techniques (Das Choudhury et al., 2020).

The studies mentioned above mainly focus on reconstructing the external structures of different crop canopies, rather than the inner occluded structures. However, existing research indicates that information of inner canopies is beneficial for excavating more data about crop growth, and carrying out more precise crop management in the

field, including the application of water, fertilizer, and pesticide. Therefore, improving the information collection of inner canopies is of great significance for efficient agricultural resource utilization and crop quality and yield improvement. In this work, we choose cotton plant as the research object and propose the three-dimensional reconstruction method that can recover internal occluded leaves. The main contributions of this study include: proposing a Cascade Leaf Segmentation and Completion Network (CLSCN) to reconstruct leaf images, and combining an instance segmentation network with a generative adversarial network to implement an end-to-end method to generate and complete the occluded leaves of the inner canopy; proposing a Fragmental Leaf Point-cloud Reconstruction Algorithm (FLPRA), based on the completed images of cotton leaves output by CLSCN, with the fusion of images and point clouds collected by RGB-D sensors to achieve efficient three-dimensional reconstruction of fragmental point-clouds of the inner cotton canopy; exploring and theoretically verifying a systematic approach that fuses GANs and ISNs to achieve three-dimensional reconstruction of cotton plants with internal canopy occluded structure recovery.

The rest of this work is structured as follows: Section 2 introduces the data acquisition method of cotton canopies, and illustrates the overall structure of the proposed CLSCN, and the implementation of the proposed FLPRA with the fusion of the images and point clouds; Section 3 illustrates the training process of the CLSCN with the established canopy database, analyzes the instance segmentation accuracy and the leaf completion effect of the CLSCN via experimental analysis, verifies the feasibility of the proposed FLPRA for the point-cloud reconstruction of cotton canopy, and finally discusses the performance and effectiveness of the reconstruction method for cotton canopies with the recovery of internal occluded structures; Section 4 summarizes this study and make prospects of future works.

2. Materials and methods

2.1. Overview

The fundamental concept of the canopy reconstruction method proposed in this work is as follows: based on the RGB-D information of the outer canopy from the top-view, implement instance segmentation of leaves on RGB images using a neural network to separate complete cotton leaves and occluded ones and generate the occluded parts of fragmental leaf images; perform a point-to-point fusion of the images and the point-clouds of cotton leaves, and for the inner fragmental point-clouds caused by the occlusion of outer branches and leaves, take the completed leaf images generated by a neural network as the vertical projection regions of the fragmental point-clouds, and design a novel point-cloud extension algorithm restricted by these projection regions, and finally reconstruct the fragmental point-cloud of inner cotton canopies. The flow chart of this work is shown in Fig. 1. The images and point clouds of cotton canopies are acquired with an RGB-D sensor from the top-view. The process of the three-dimensional reconstruction of the canopy with inner occluded leaves recovery is divided into two steps: Construct the CLSCN with two cascading sub-networks, namely an instance segmentation network and a generative adversarial network. With the acquired leaves images of canopies as the input of the front-end network, the leaf masks are outputted by the instance segmentation network to realize the separation of leaves. The separated leaves are transmitted to the back-end network one by one, and through the structure of generators and discriminators, this generative adversarial network can maintain the full images of the complete leaves, and auto-complete the fragmental images of the occluded leaves; With the proposed FLPRA, align the image and point-cloud point to point, filter the edge noises and outlier points, and using the masks outputted by the back-end network of CLSCN as a reference to separate the point-clouds of leaves one by one, then complete the fragmental point-clouds of inner occluded leaves through point-cloud extension with the masks as the

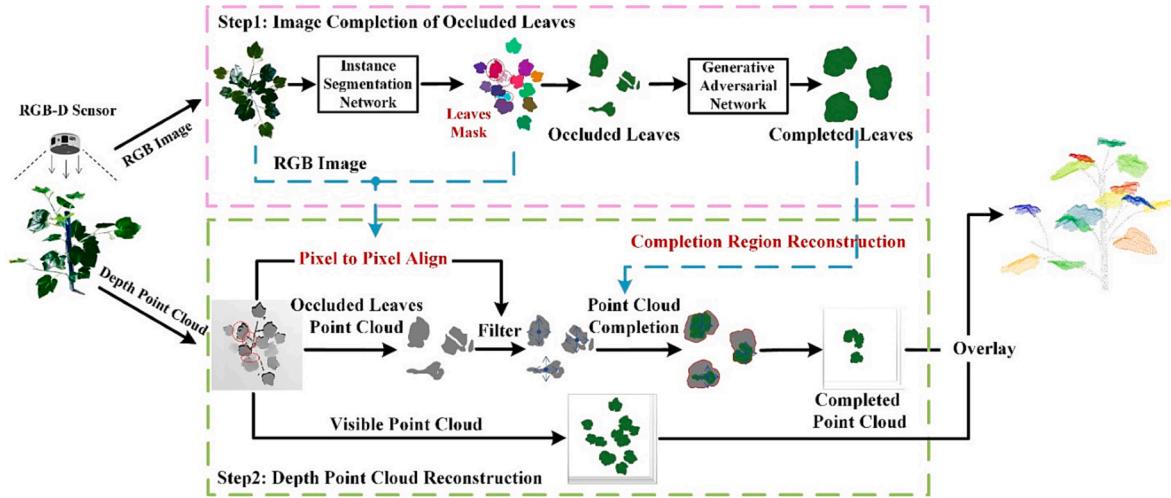


Fig. 1. Three-dimensional reconstruction of a cotton canopy with the recovery of inner occluded leaves.

projection areas, and finally construct the three-dimensional model of a cotton canopy by overlaying these point-clouds of leaves with the recovery of inner occluded leaves.

2.2. Data acquisition and augmentation

Artificial cotton plants are built instead of real plants for data acquisition, network training and test validation. Considering that the seasonal growth natures of crops make the experiments inefficient, in recent years, benefiting from the development of crop morphology and its related technologies, realistic artificial plants have often been built for the study and validation of some theoretical methods, which has significantly improved the research efficiency, especially in the field of three-dimensional reconstruction. Cotton morphology indicates that although the spatial distribution of the local branches and leaves in a canopy reflects the randomness of growing competition in nature, there are significant morphological features of the whole canopy structure in a global view: the leaves and branches are spirally distributed on the main stem, and it is called the “3/8” growth law of cotton (Liu et al., 2020), namely, eight branches and leaves spiral around the main stem by three times, and the ninth branch is exactly above the first one along the direction of stem growth. Based on the growth law, several artificial cotton plants were built using seedling cotton as a reference (Fig. 2ab).

The information on canopies was captured with a RealSense L515 RGB-D sensor produced by Intel Corporation. The RealSense L515 is a revolutionary solid-state LiDAR depth camera that utilizes a proprietary MEMS mirror scanning technology, enabling better laser power efficiency compared to other time-of-flight technologies. With a power consumption of less than 3.5 W for depth streaming, the Intel RealSense LiDAR camera L515 is the world's most power-efficient high-resolution

LiDAR camera. The basic parameters of the sensor are: Depth: resolution up to 1024×768 ; minimum distance of around 25 cm; depth accuracy of around 5 mm to 14 mm; frame rate up to 30fps; RGB: resolution up to 1920×1080 ; frame rate up to 30fps. The sensor supports the automatic alignment of RGB images to depth images point-to-point. Considering the depth Field of View (FOV) is $70^\circ \times 55^\circ (\pm 3^\circ)$, and the RGB FOV ($H \times V$) is $70^\circ \times 43^\circ (\pm 3^\circ)$, the device was fixed on a frame at a height of about 0.6 m above the artificial cotton plants (Fig. 2c), covering an area of about $1.0 \times 0.5 \text{ m}^2$ where it is able to cover the entire canopy of the artificial plants. The fixed height can be adjusted according to the size of the canopy, ensuring both the integrity and accuracy of the canopy information.

The original resolutions of RGB images and depth images were both set as 480×480 during data acquisition (Fig. 2de). We collected over 620 images of cotton canopies and converted them into a VOC dataset for the training of the proposed neural networks. Additionally, to enhance the generalization ability of the model and prevent overfitting during network training, this study further performed data augmentation on cotton canopy images. Data augmentation not only expands the dataset but also improves the quality of neural network training. The specific implementation process of the data augmentation method used in this study is as follows: (1) with a 50% probability, perform image affine transformations such as rotation, scaling, adding noise, cropping, and translation on the dataset to reduce the impact of different cotton leaf postures and sizes on instance segmentation results; (2) with a 50% probability, adjust the brightness of the dataset and randomly choose to increase the brightness by 1.2 times or decrease it by 0.8 times to reduce the impact of different light intensities on the accuracy of cotton leaf instance segmentation; (3) with a 50% probability, adjust the contrast of the dataset and perform 0.8 ~ 1.2 times attenuation or enhancement to

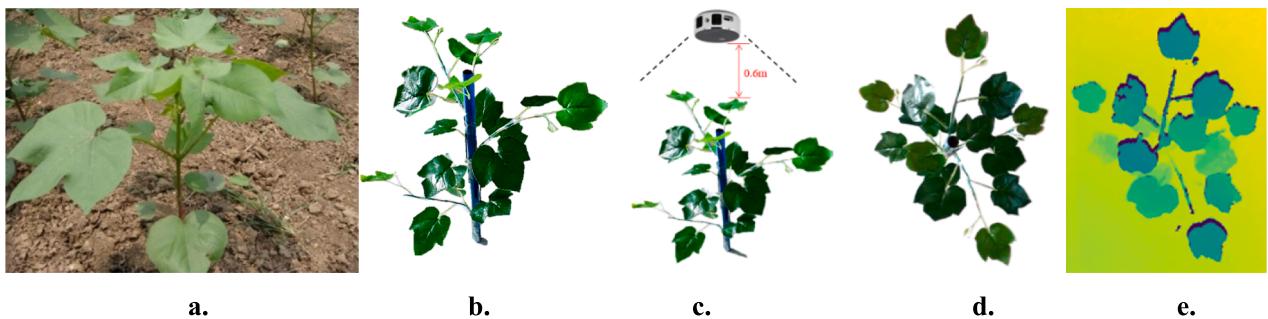


Fig. 2. Artificial cotton plant construction. a. cotton at seedling stage, b. artificial cotton plant, c. perspective of data acquisition, d. RGB image of cotton canopy, e. depth image of cotton canopy.

better enable the neural network to learn more details of cotton leaf features. Using the above data augmentation methods, the cotton canopy image dataset was expanded to over 3,400 images. Some augmented data are shown in Fig. 3.

2.3. Cascaded leaf segmentation and completion network

$$\left\{ \begin{array}{l} L_{RPN} = \frac{1}{N_{cls1}} \sum_i L_{cls}(p_i, p_i^*) + \lambda_1 \frac{1}{N_{reg1}} \sum_i p_i^* L_{reg}(t_i, t_i^*) \\ L_{Mul-Branch} = \frac{1}{N_{cls2}} \sum_i L_{cls}(p_i, p_i^*) + \lambda_2 \frac{1}{N_{reg2}} \sum_i p_i^* L_{reg}(t_i, t_i^*) + \gamma_1 \frac{1}{N_{mask}} \sum_i L_{mask}(s_i, s_i^*) \end{array} \right. \quad (1)$$

2.3.1. Front-end leaf segmentation network

The front-end leaf segmentation network is based on Mask-RCNN, which is an extension of Faster R-CNN (Li et al., 2021; Shi et al., 2019). By adding a parallel branch for outputting the masks of objects, Mask-RCNN has extended its applications to the field of instance segmentation. In this paper, the mask output module of the Mask-RCNN network is further improved. The data output mode was modified from the overlay mode to the single mode, so that the segmented leaf masks of canopies can be outputted separately. The improved leaf instance segmentation network mainly contains three parts: feature extraction and fusion, ROI proposal and ROI classification regression. For achieving fast extraction of cotton canopy leaf feature maps, A ResNet50 network is chosen as the backbone for feature extraction, which can greatly improve the convergence speed while ensuring the learning

ability of the network. The structure of the front-end leaf segmentation network is shown in Fig. 4.

Define the total loss of the instance segmentation network as $L_{sg-net} = L_{RPN} + L_{Mul-Branch}$, where L_{RPN} includes the anchor classification loss and the bounding-box regression loss in the RPN network, and $L_{Mul-Branch}$ includes the anchor classification loss, the bounding-box regression loss and the mask loss in multi-branch prediction, and

where, i represents the anchor index; N_{cls1} and N_{cls2} represent the number of anchors; N_{reg1} , N_{reg2} and N_{mask} represent the number of bounding boxes; λ_1 , λ_2 and γ_1 represent the hyperparameters to balance the total training losses; p_i represents the classification probability of the i -th anchor; p_i^* represents the corresponding probability of the i -th ground-truth label; t_i represents the difference between the i -th prediction bounding box and the i -th ground-truth box; t_i^* represents the difference between the i -th ground-truth label box and the positive anchor; s and s^* represent the mask matrices in binary format from the predictions and the ground-truth labels, and

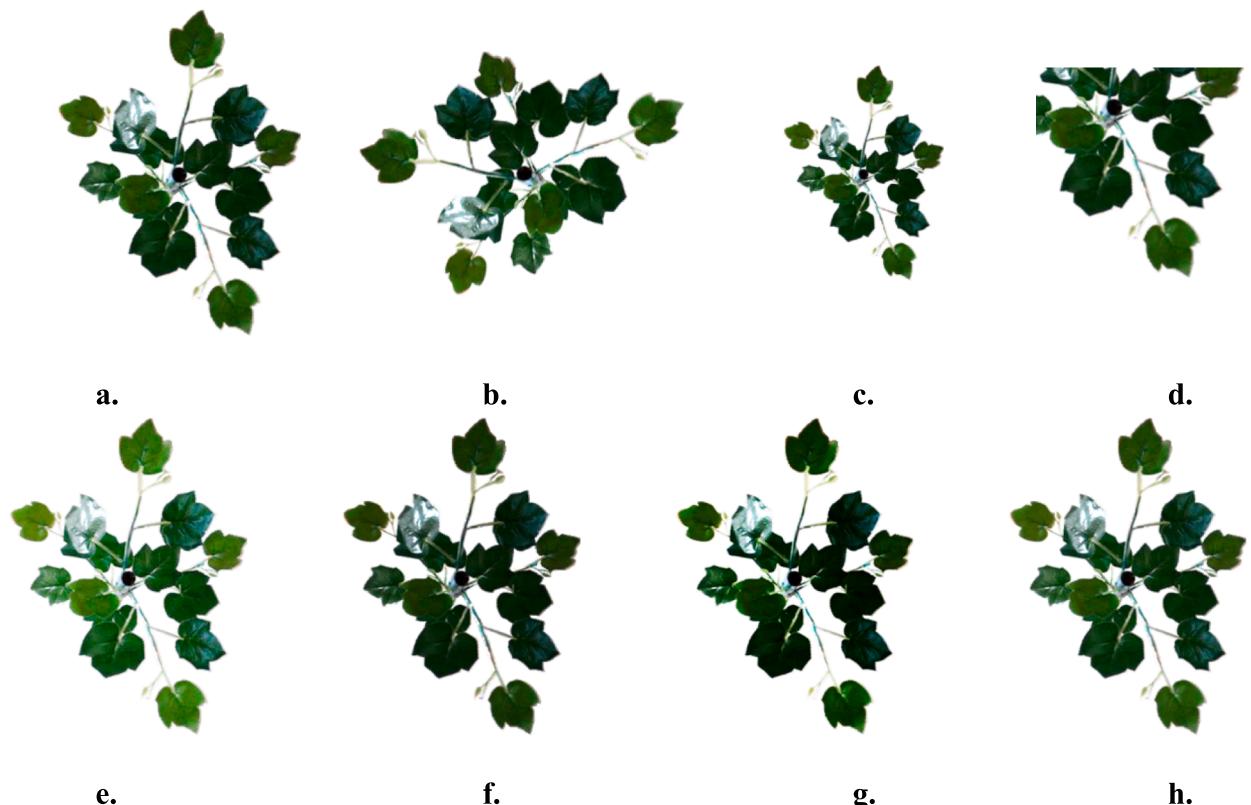


Fig. 3. Data augmentations. a. original image, b. random rotation, c. random scaling, d. random cropping, e. brightness enhancement, f. brightness reduction, g. contrast enhancement, h. contrast reduction.

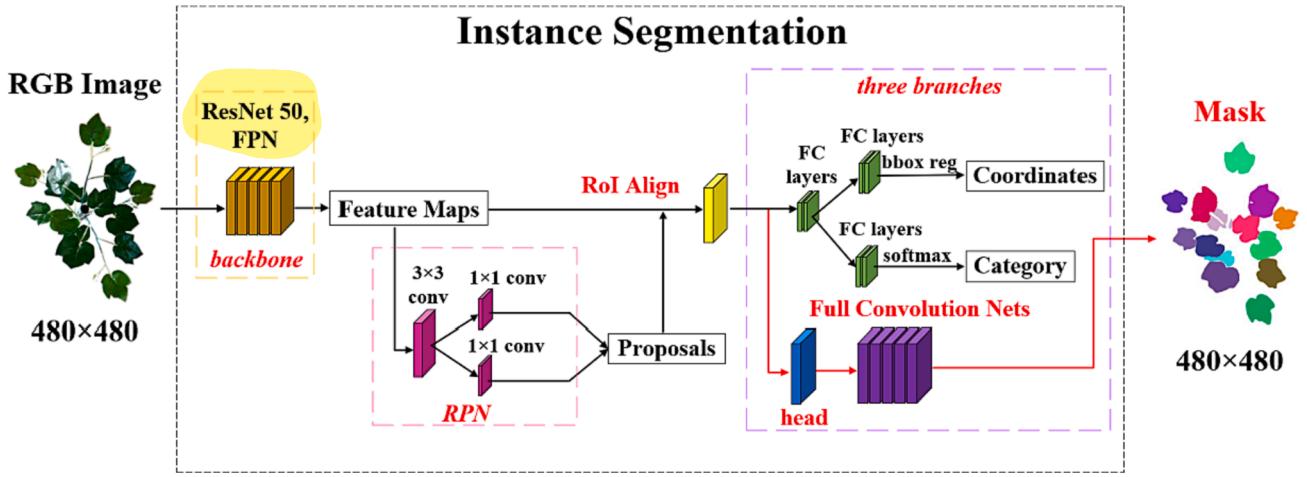


Fig. 4. Instance segmentation network for cotton leaves.

$$\left\{ \begin{array}{l} L_{cls}(p_i^*, p_i) = -\log p_i^* \\ L_{reg}(t_i^*, t_i) = smooth_{L1}(t_i^* - t_i), \text{smooth}_{L1}(x) = \begin{cases} 0.5x^2 & \text{if } |x| < 1 \\ |x| - 0.5 & \text{otherwise} \end{cases} \\ L_{mask}(s^*, s) = -(s^* \log(s) + (1 - s^*) \log(1 - s)) \end{array} \right. \quad (2)$$

The RGB images of cotton canopies are input to the ResNet50 backbone to extract features and export the feature maps. An FPN is used for the detection of cotton leaves with different sizes so that it can extract the multi-scale features of leaves. Then, the region of interest (RoIs) of cotton leaves can be extracted by the previous layer of the network with an RPN. These RoIs are adjusted to a unified size by an ROI-Align algorithm, which uses bilinear interpolation to guarantee the bounding-boxes well-aligned with the RoIs on the original images, so that more accurate segmentation results can be obtained. Thus, the leaf regions are predicted by regression and classification with the fixed-size RoIs, and the instance segmentation masks of cotton leaves are generated, whose size remained consistent with the original images. With this front-end leaf segmentation network, cotton leaves in the canopy can be segmented one by one. It provides high-quality mask images of cotton leaves inputted to the following generative adversarial network.

2.3.2. Back-end leaf generative network

The back-end generative network is based on an adversarial structure with context encoders integrated, and the overall structure is shown in Fig. 5. The adversarial structure consists of a generator and a discriminator, and the generator generates candidates while the discriminator evaluates them (Espejo-Garcia et al., 2021; Huang et al., 2023; Lu et al., 2022; Madsen et al., 2019). With this structure, the network can implement the high-quality completion of fragmental leaf images. Furthermore, the generator is constructed by an encoder and a decoder. The encoder is a multilayer convolution with a 3×3 convolution kernel and the decoder are a multilayer deconvolution with a 4×4 convolution kernel, and the two parts are connected by a convolution layer with 1×1 convolution kernel. Corresponding to the output size of the generator, we also use 4×4 convolution kernels to construct the multilayer convolution of the discriminator.

The total loss of the GAN includes generative loss and adversarial loss. The former is a Mean Squared Error Loss (MSE Loss) to quantify the ability of image generation, defined as

$$L_{rec}(x) = \|\hat{M} \odot (x - F((1 - \hat{M}) \odot x))\|_2^2 \quad (3)$$

where, x represents the ground-truth data; F represents a context encoder; \odot represents the multiplication operator for the corresponding

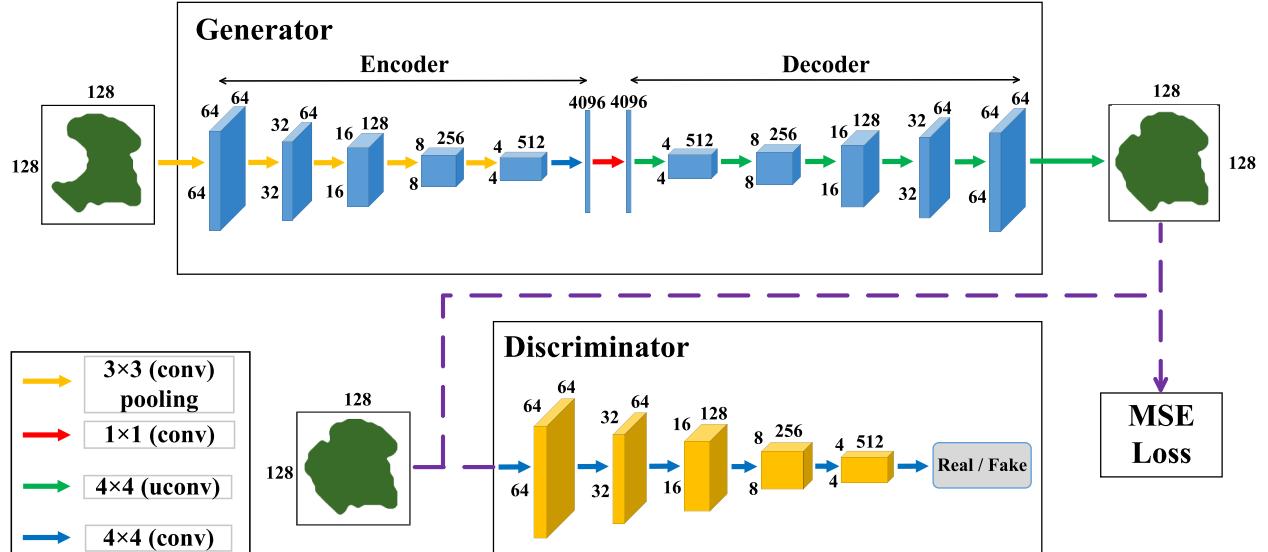


Fig. 5. Leaf generation network.

positions of a matrix; \hat{M} represents the binary mask of the missing region. Adversarial loss is to distinguish the original images from the fragmental ones, defined as

$$L_{adv} = \min_{G} \max_{D} \mathbb{E}_{x \in X} [\log(D(x))] + \mathbb{E}_{z \in Z} [\log(1 - D(G(z)))] \quad (4)$$

where z represents the input noise; Z represents a noise distribution; X represents a data distribution; D represents the discriminator; G represents the generator. Therefore, the joint loss L_{joint} of the back-end leaf generative network is defined as

$$L_{joint} = \lambda_{rec} L_{rec} + \lambda_{adv} L_{adv} \quad (5)$$

where, λ_{rec} and λ_{adv} represents the hyperparameters to balance the total training loss.

Before training, a method named Missing at Random Circles (Guo and Liu, 2019) is used to destroy the complete leaves for the supervised adversarial training. The method can erase pixels of a RGB image in a circle region with a random center and radius. With the fragmental leaf images inputted to the generator, image features are extracted by the convolutional layers of the encoder. An extracted feature map is converted to a high-dimensional latent vector with 4096 dimensions, and sequentially through a decoder to decode the latent vector with multiple deconvolution layers, and output a completed image of a cotton leaf in the same size. The completed image will be imported to the discriminator, which extracts and compares the features of the original image and the completed image, to discriminate the authenticity of the completed leaf image. Through the GAN, the acquired fragmental images of inner leaves occluded by the outer leaves can be well completed to obtain complete leaf images.

2.4. Fragmental leaf point-cloud reconstruction algorithm

2.4.1. Point-cloud filter of cotton leaves

The completed leaf images generated with GAN as the vertical projection regions of the fragmental leaf point clouds, the leaf point clouds extend within the restrictions of projection boundaries, and then the fragmental point clouds of leaves inside the cotton canopy can be finally reconstructed. However, the leaf point clouds directly acquired by RealSense L515 often include discontinuous values, noises and outliers (Fig. 6a), due to the factors such as object reflectivity, edge reflectivity and ambient light intensity. These original point clouds should be well-filtered. Considering a point cloud of cotton canopy can be separated into single leaves one by one reference to the mask of instance segmentation with point-to-point alignment, the filtering operation can be performed in leaf unit. The point clouds to be filtered are mainly considered as two categories. The first category is the outlier points within a leaf surface, and the second category is the discontinuity points on the outline edge of a leaf. As shown in Fig. 6b, the geometric center O_r of a leaf can be located by index searching both in a row and column direction. Reference to the depth value of the center point, the depth

smooth of the whole leaf surface can be performed in vertical slices. The S plane in Fig. 6b represents a vertical slice, and the area circled in red is the cross-section where the S intersects the leaf surface. For the former, the mean filtering method is used for the point-cloud filter in plane S . Search around the center O_r within a leaf surface, mark the point depth significantly changing as D_i and use it as a reference point with the depth value as d_i . Search for the two adjacent points D_{i-1} and D_{i+1} in a direction and record their depth values as d_{i-1} and d_{i+1} respectively. Then, the gradients k_{i-1} and k_{i+1} among the adjacent points are calculated, and the next adjacent point D_{i+2} can be continuously found in this direction. The gradient and average depth are calculated between adjacent points as.

$$k_i = \frac{d_{i-1} - d_i}{y_{i-1} - y_i} \quad (6)$$

$$\Delta \bar{d} = \frac{(d_{i-1} - d_i) + (d_{i+1} - d_{i+2})}{2} \quad (7)$$

where, y_i represents the coordinates on section S , d_i represents the depth value, and k_i represents the slope between two adjacent points. By comparing adjacent gradients, the depth-varying direction on a leaf surface within the neighborhood can be inferred. After determining the direction, the first category points can be replaced with the average value $\Delta \bar{d}$.

The second category is the points of depth faults on the edge of a cotton leaf, which often occur due to the large changes of reflectivity at the edges of the observed objects during data acquisition with an RGB-D sensor. For these points, the changes in depth values need to consider the gradients rather than just using the average depth as a substitute. Considering the directions of depth changing are generally along the directions of laser emission, it is necessary to adjust and assign depth for each segment in change separately, and each segment is

$$\Delta \bar{d} = (d_{i+1} - d_i) \frac{k_{i+1}}{k_i} \quad (8)$$

By filtering the point clouds of cotton leaves, the surface of cotton leaf point clouds can be smoother, the overall coherence of cotton leaf point clouds is better, and it is closer to real cotton leaves, thus improving the reconstruction accuracy of occluded leaves inside cotton plants.

2.4.2. Point-cloud reconstruction of cotton leaves

The point-cloud reconstruction is based on the comprehensive masks generated by the leaf GAN. The reconstruction process can be regarded as an extension of the fragmental point cloud within a restricted area (within the contour of a leaf) determined by the projection of the mask. As shown in Fig. 7, the dark green area represents a leaf mask (with the dark red area indicating the completed part generated by the GAN), and its vertical projection generates an area with a dashed contour. In this area, the green dots represent the fragmental point cloud, while the red dots represent the points that need to be reconstructed. The point-cloud

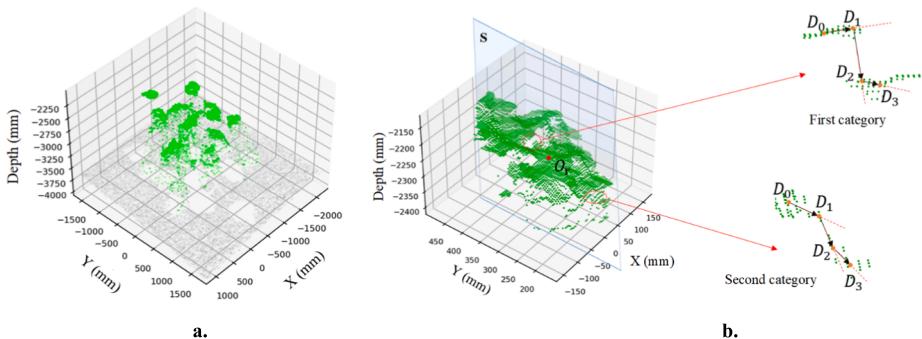


Fig. 6. Point-cloud processing of cotton canopy. a. cotton canopy point-cloud, b. point-cloud filter.

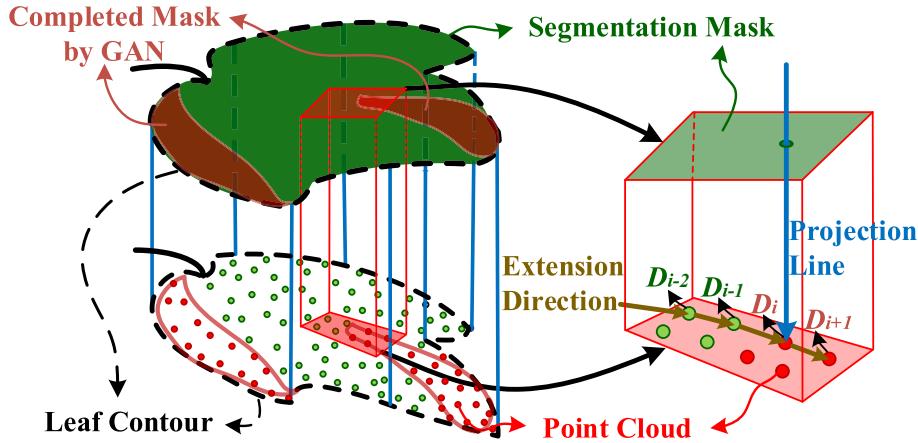


Fig. 7. Point-cloud reconstruction diagram.

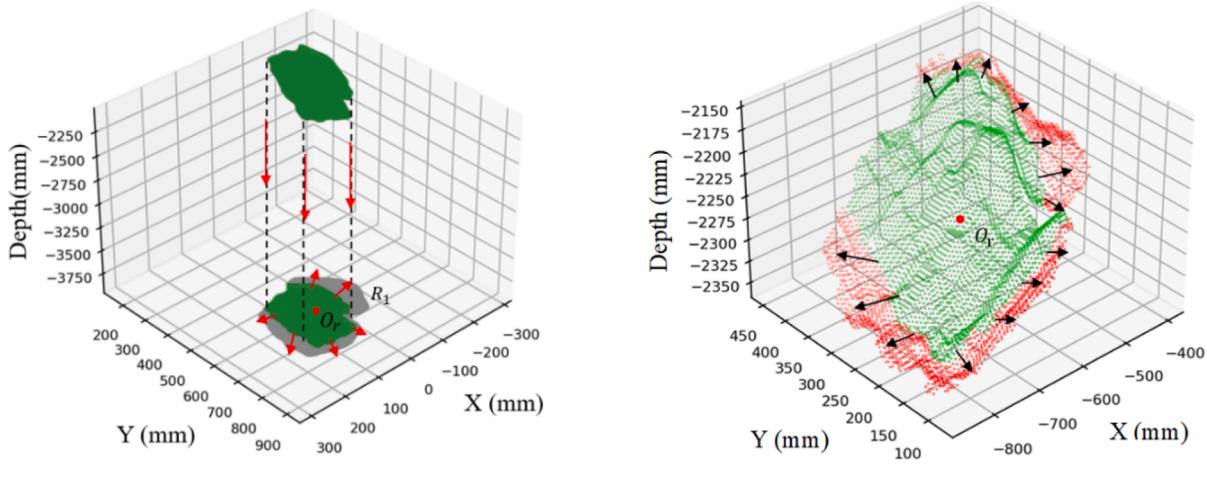


Fig. 8. Leaf point cloud completion. a. projection of point-cloud, b. extension of point-cloud.

reconstruction is based on the fundamental assumption that the curvatures among adjacent points are continuous. This means that the extension direction of a new point is determined by its adjacent points in the local neighborhood. As an example, a new point D_i is determined by the existed points D_{i-1} and D_{i-2} , and the next point D_{i+1} is determined by the neighboring point D_{i-1} and the new generated point D_i , and so on.

Using this method, a leaf point cloud can be completely reconstructed.

In detail, as shown in Fig. 8a, take the center point O_r of a fragmental point cloud as the original point and extend new points to the boundary, and the grey area is the restricted area of the projection area R_1 which is a generated mask by the GAN. Considering that the leaf point-cloud is three-dimensional, the point-cloud extension should maintain the

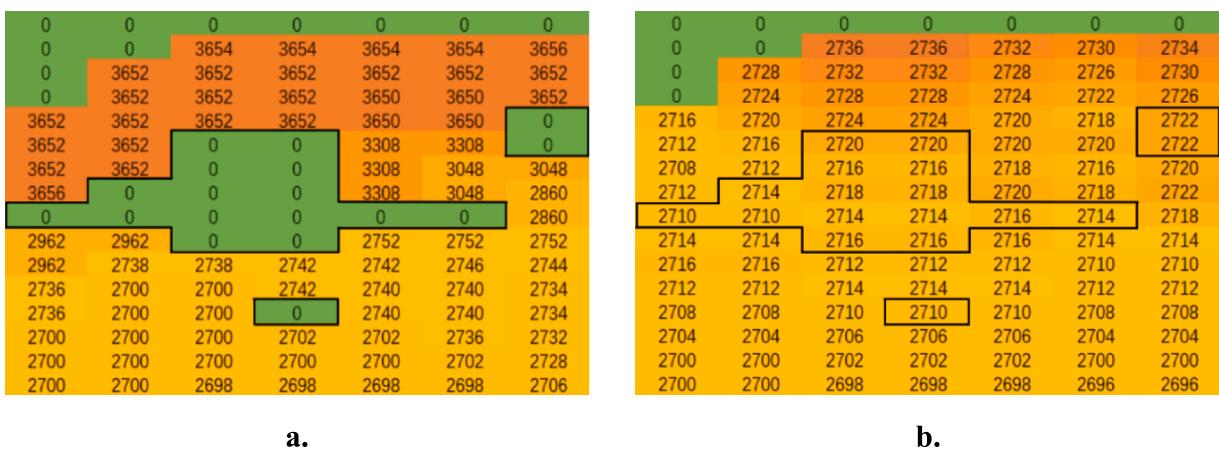


Fig. 9. Depth distribution in data structure. a. before and b. after point cloud extension.

curvature continuity on the leaf surface in all directions. Search in different directions with O_r as the center, and define the point of which the gradient suddenly changes as D_{ij} (in projection X-Y plane, and $i \in X$, $y \in Y$) on the edge of a fragmental point-cloud, then the adjacent points can be found in the reverse directions. With the corresponding depth of the point defined as d_{ij} , considering that the varying depth values of adjacent points in a region determining the spatial attitude of the leaf point cloud, the depth values in the extended area are filled as

$$d_{j,k} = \text{mean}(d_{m,n} + k_{m,n}\Delta d) \quad (9)$$

where, $\text{mean}(\cdot)$ represents a mean function, Δd represents the sampling distance of the point cloud, $d_{m,n}$ and $k_{m,n}$ represents the depth value and the slope of the point $D_{m,n}$, and

$$D_{m,n} \in U_\delta = \{D_{ij} \mid \|D_{ij} - D_{m,n}\|_2 \leq \delta\}$$

where, U_δ is the neighborhood of D_{ij} . After filling the missing parts of the point cloud with the depth values, a complete and continuous cotton leaf point cloud can be obtained (Fig. 8b).

The values in the data structure are shown in Fig. 9, the area with a depth value of zero in the center is the fragmental part of a leaf point cloud, and the top area with a zero value is the edge. The points in the fragmental area should be extended. The depth values of adjacent points are referenced to fill the depth of the missing area with Eq. 9. After the extension of the point cloud, all points with missing depth values have been repaired, and the depth distribution of any region in the extended point cloud follows the depth distribution law of a cotton leaf. After filtering and extension of the point cloud, the surface of the cotton leaf is smooth and complete, achieving the reconstruction of a cotton leaf with an internal occluded structure repaired.

For aligning RGB images with point-cloud pixels after the reconstruction of point clouds, a three-dimensional Cartesian coordinate is established, and all point clouds of the cotton leaves in the canopy are placed under the same coordinate system, as shown in Fig. 10. Each point cloud of the cotton leaf is overlaid according to its relative position in space, and finally, the whole point cloud of the cotton canopy P is as

$$P = \sum_{i=1}^N P_i + \sum_{j=1}^M P_j \quad (10)$$

where, N represents the number of original point clouds of complete leaves, M represents the number of completed point clouds of fragmental

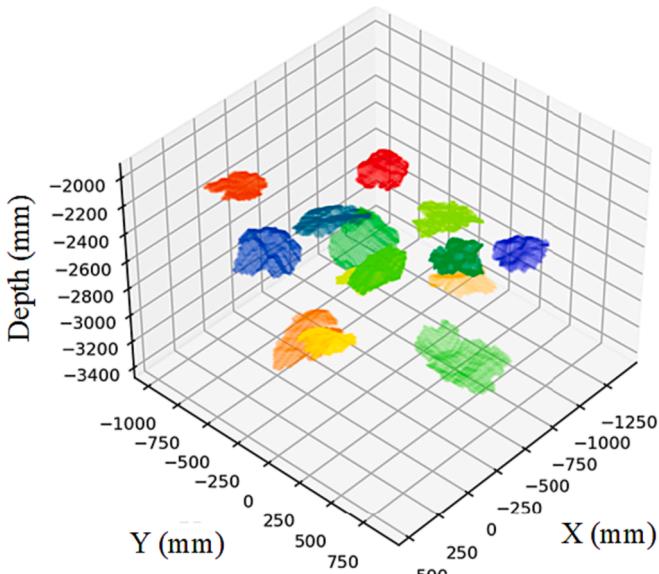


Fig. 10. Reconstructed point cloud of cotton plant.

leaves, P_i represents the original point cloud of the i -th complete leaf, P_j represents the generated point cloud of the j -th fragmental leaf. Then, the internal occluded structures of a canopy have been reconstructed, the point clouds are of better integrity and coherence, and the outliers are eliminated in the data acquisition.

3. Results and discussions

The three-dimensional reconstruction of cotton canopies was based on the PyTorch deep learning framework, and the parameters of the computation hardware are: CPU: Intel (R) Core (TM) i7-12700; GPU: NVIDIA RTX 3060; RAM: 16 GB. The training of CLSCN contains two parts: the front-end instance segmentation network training and the back-end GAN training. The accuracy of leaf segmentation and the precision of leaf completion are analyzed. After the training of networks is finished, the reconstructions of the separated leaf point clouds are carried out to verify the feasibility and reconstruction accuracy of the proposed FLPRA algorithm.

3.1. Test of leaf image completion

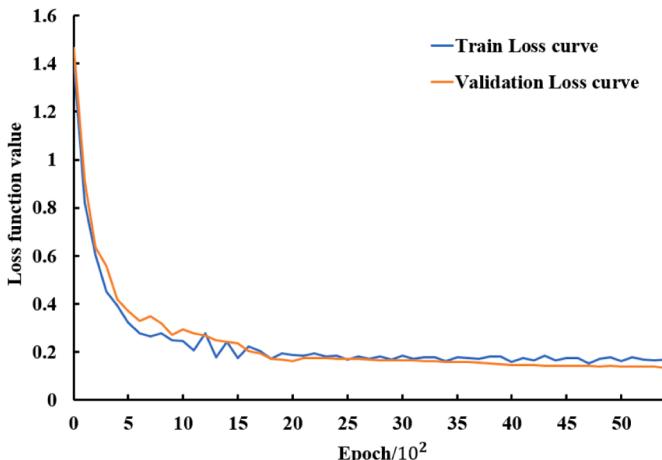
3.1.1. CLSCN training

The image dataset of the cotton canopy for training the front-end instance segmentation network was divided into a training set and a validation set with a ratio of 4:1. The training set contained 2 720 images and were manually semantically annotated using LabelMe software. The classifications, bounding boxes and masks of each leaf image are outputted after the instance segmentation network. A canopy image was segmented into many single-leaf images, and the resolutions were adjusted to 480×480 pixels automatically, resulting in a total of 12 672 independent cotton leaf images, including 9 024 complete leaves and 3 648 fragmental leaves. The complete leaves are used for training the back-end generative network, and the fragmental leaves are used for validating the generative ability of the back-end network. While training, for the front-end network, the batch size was set as 4 and the optimizer used SGD, the initial learning rate was set to 5×10^{-3} , the learning momentum was set to 0.9, the weight decay was set to 5×10^{-4} , and the IoU threshold was 0.7. In back-end GAN, the batch size was set to 8, and the optimizer used Adam, the learning rate of the discriminator was set to 2×10^{-4} and the learning rate of the generator was set to 7×10^{-4} .

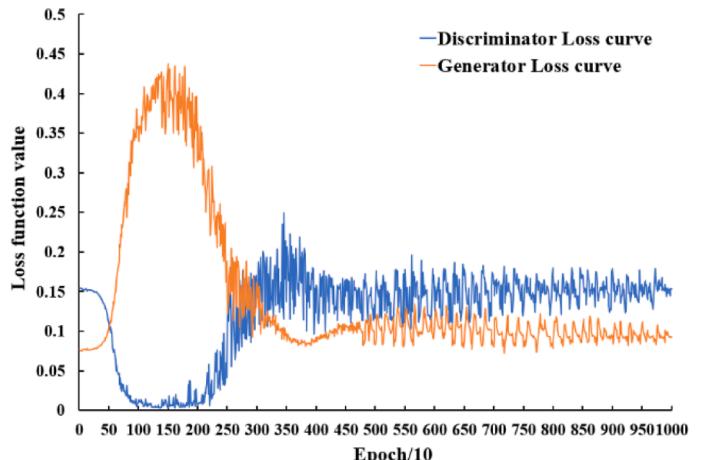
The training curves of front-end instance segmentation network were shown in Fig. 11a, and the total loss showed a decreasing trend during the training process. After about 1 000 epochs, the loss finally converged and reached the termination condition. The training finished after about 6 000 epochs, and the total training loss for both training set and validation set were less than 0.2. The training curves of back-end generative adversarial network were shown in Fig. 11b, the losses of the generator and the discriminator showed a tendency of increasing and decreasing separately, and after more than about 8 000 epochs, the fluctuation of loss tended to stabilize, indicating that the generator could repair fragmental images after the training of the generative network, while the discriminator also distinguished the minor differences between the real and fake images.

3.1.2. Accuracy analysis of instance segmentation

To verify the accuracy of the instance segmentation model, many metrics could be used to evaluate the detection effectiveness, such as *Precision*, *Recall*, and *mean Average Precision*. Generally speaking, instance segmentations are subsequently used for object detection tasks. And in these cases, it is nice to have a metric to compute their accuracy on this subsequent task. Especially in the early days of instance segmentation, there didn't seem to be many datasets with instance segmentation labels, so comparing against bounding box ground truth labels was more essential. More formally, *Precision* is applied to measure prediction results, while *Recall* is adopted to measure the quality of



a.



b.

Fig. 11. Curves of Loss function of CLSCN. a. curves of loss function of leaf instance segmentation network, b. curves of loss function of Generative Adversarial Network.

positive predictions. The *Precision* and *Recall* are defined as.

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}}$$

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}}$$

where, *TP* and *FP* represent the true and false positive number, respectively, and *FN* represents the false negative number. Consequently, the AP value is computed as

$$\text{AP} = \int_0^1 P(R) dR \quad (11)$$

where, *P(R)* presents the PR curve. In this work, one hundred points are sampled on the PR curve for calculation, the threshold of IoU is adjusted from 0.5 to 0.95 with a step of 0.05 for the AP calculation, and the mean value of all the results is taken as the final result. Generally, AP is calculated under a single category, and mAP is the average of AP value under all categories, that is, the mAP is defined as

$$\text{mAP} = \frac{1}{n_c} \sum_i^n \text{AP}_i \quad (12)$$

Specifically, mAP50 and mAP_[0.5:0.95] are denoted as the mAP values for IoU = 0.5 and IoU between 0.5 and 0.95, respectively. In terms of the accuracy, the relevant indexes of the bounding box and the masks are calculated separately. For bounding box, the mAP, mAP50 and mAP75 are 91.25 %, 98.91 % and 98.04 %, and for masks, the mAP, mAP50 and mAP75 are 91.40 %, 99.24 % and 98.17 %.

The Intersection over Union (IoU) metric was employed to assess the level of overlap between the prediction box and the ground truth box, which is defined as the ratio of the overlap area and the union area between the segmentation mask and ground-truth mask. In this work, with the consideration of each pixel given the same importance, the IoU of instance segmentation was performed with 412 images of cotton canopies in the validation set, and the Mean Intersection over Union (mIoU) is used to evaluate the results of the leaf instance segmentation network as the expression in pixel format

$$\text{mIoU} = \frac{1}{k+1} \sum_{i=0}^k \frac{p_{ii}}{\sum_{j=0}^k p_{ij} + \sum_{j=0}^k p_{ji} - p_{ii}} \quad (13)$$

where *k* represents the total class number; *p_{ij}* represents the number of pixels that belong to category *i* but have been misclassified as category *j*; *p_{ji}* represents the number of pixels that belong to category *j* but have been misclassified as category *i*; *p_{ii}* represents the number of pixels belongs to category *i* that have been correctly judged as category *i*.

The test results are shown in Fig. 12. The MIoU of unobstructed leaves was 94.95%, the MIoU of leaves occluded by stems was 86.22%, the MIoU of leaves occluded by other leaves was 85.89%, and in general, the MIoU reached 87.22% for these several cases. Therefore, the leaf instance segmentation network proposed in this work can perform high-quality segmentation of cotton leaves in occlusion after training, which can satisfy the test requirements for cotton canopy leaf segmentation, and provide high-quality mask images for the subsequent training of leaf generative network.

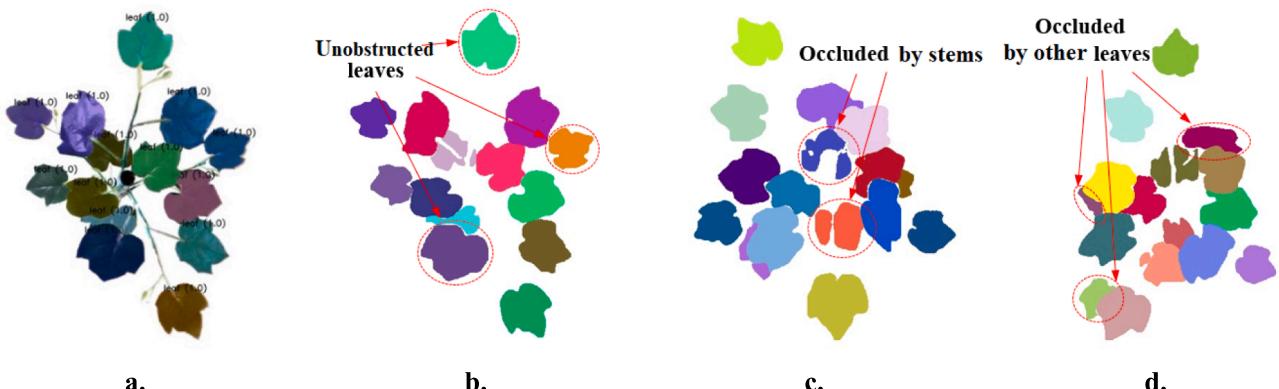


Fig. 12. Instance segmentation of cotton canopy. a. output result of network, b. unobstructed leaves, c. occluded leaves by stems, d. occluded leaves by other leaves.

3.1.3. Accuracy analysis of leaf image completion

Fréchet Inception Distance (FID) is one of the most popular metrics for measuring the feature distance between the real and the generated images. Fréchet Distance is a measure of similarity between curves that takes it into account that the location and ordering of the points along the curves. It can be used to measure the distance between two distributions as well. Mathematically, Fréchet Distance is used to compute the distance between two “multivariate” normal distribution. For a “univariate” normal distribution Fréchet Distance is given as,

$$d(X, Y) = (\mu_X - \mu_Y)^2 + (\delta_X - \delta_Y)^2 \quad (14)$$

where μ and δ are the mean and standard deviation of the normal distributions, X and Y are two normal distributions. The FID for “multivariate” normal distribution is given by,

$$FID = \|\mu_X - \mu_Y\|^2 - Tr(\Sigma_X + \Sigma_Y - 2\sqrt{\Sigma_X \Sigma_Y}) \quad (15)$$

where, X and Y are the real and fake embeddings (activation from the Inception model) assumed to be two multivariate normal distributions, and Tr is the trace of the matrix and Σ_X and Σ_Y are the covariance matrix of the vectors.

Over four hundred cotton leaves within the validation set were randomly selected for the leaf completion test. The occluded leaves were divided into three categories: complete leaves, occluded leaves by stems, and occluded leaves by other leaves. As shown in Table 1, the overlap rates and the FID scores were calculated separately, where all the FID scores were less than 35, with the mean value 21.11. Furthermore, for the fragmental leaves, we used the region overlaps between the after-

completion leaves by the back-end leaf GAN and the before-completion leaves to measure the accuracy performance of leaf completion. Considering that the GAN may generate extra regions beyond a leaf image, we recognized that subtracting the proportion of this extra part is reasonable. By defining the image of the fragmental leaf before the completion as I_o , the completed region after the leaf GAN as I_c , and the ideal completed leaf image as I_t , the accuracy C_d before and after completion is defined as

$$C_d = \frac{(I_o \cup I_c) \cap I_t}{I_t} - \frac{(I_o \cup I_c) \setminus I_t}{I_t} \quad (16)$$

where, the operator ‘\’ means excluding a set from another set, for example, ‘A\B’ means excluding a set B from a set A.

The accuracy values were also shown in Table 1, and the statistical results of leaf image completion are shown in Fig. 13. It can be seen that the completed RGB images output from the back-end GAN can keep their original geometry as much as possible, and the overlap rates were up to 95%~100%. The overlap regions between the repaired cotton leaf image and the original leaf image were all higher than 94%. Thus, it is shown that the Generative Adversarial Network is effective in completing the fragmental leaves of the cotton canopy and is suitable for completing the leaves against various occluded types of the cotton canopy.

3.1.4. Accuracy analysis of FLPRA

Different with RGB images, an appropriate three-dimensional metric is essential to characterize the difference between two point clouds, due to the spatial disorder property of point clouds. There are three

Table 1
Completion results of cotton leaves by back-end GAN.

Leaves	Occlusion rate (%)	Completed leaves	FID score	Accuracy (%)
	0%		2.71	99.98%
	7.65%		22.29	95.28%
	10.66%		19.37	96.24%
	22.16%		28.12	94.36%
	73.94%		33.08	97.26%

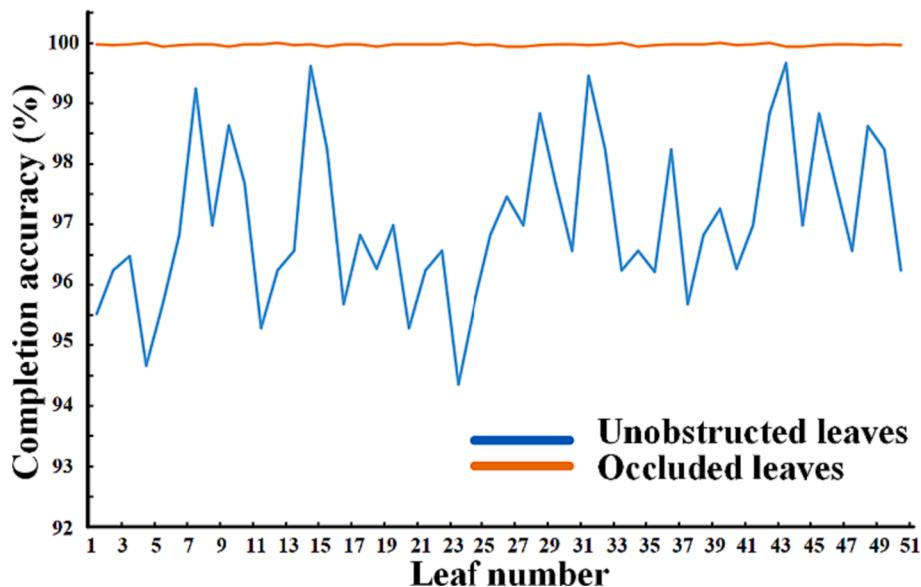


Fig. 13. Completion accuracy of unobstructed leaves and occluded leaves.

commonly-used metrics to compute the difference between two point-clouds, namely, Earth Mover's Distance (EMD), Hausdorff Distance (HD) and Chamfer Distance (CD). The EMD usually involves pairing the original point cloud with the points in the generated point cloud, and calculating their pairing distance (Solomon et al., 2014). The smaller pairing distance, the smaller difference between two point-clouds, mainly reflecting the differences in the density and distribution of the points. The HD measures how far two \mathbb{R}^3 subsets of a metric space are from each other (Chopin et al., 2023). It is the longest distance to travel from a point in one of the two sets to the closest corresponding point in the other set. In other words, it is the greatest of all the distances from a point in one set to the closest point in the other set. The CD is to measure the difference between two point clouds by using the minimum distance (Al-Shurbaji et al., 2022). Different with the EMD and HD, the CD needs not to match each pair points, but needs to calculate the point with the minimum distance from the original point cloud, which mainly reflects the difference between the two point-clouds in the shape and the contour of leaves. Usually, it is necessary to choose the most suitable one to characterize the similarity of two point-clouds. Considering that the

point cloud and the RGB image can be pixel aligned with Realsense L515, and there is no difference in the density of the point cloud distribution, the CD is used to measure the difference between two point-clouds in this work.

Mathematically, the Chamfer Distance between two point-clouds $P_1 = \{x_i \in \mathbb{R}^3\}_{i=1}^n$ and $P_2 = \{x_j \in \mathbb{R}^3\}_{j=1}^m$ is defined as the average distance between pairs of nearest neighbors between P_1 and P_2

$$\text{chamfer}(P_1, P_2) = \frac{1}{2n} \sum_{i=1}^n |x_i - \text{NN}(x_i, P_2)| + \frac{1}{2m} \sum_{j=1}^m |x_j - \text{NN}(x_j, P_1)| \quad (17)$$

where $\text{NN}(x, P) = \operatorname{argmin}_{x' \in P} \|x - x'\|$ is the nearest neighbor function.

A point is marked as high-matched point with the corresponding spatial position if the calculated CD value is less than 10^{-4} . Table 2 shows some fragmental point clouds, reconstructed point clouds, and the calculated CD values refer to the original complete point clouds. The data shown in Table 2 reflect that the CD values of two corresponding

Table 2
Point cloud reconstruction analysis.

Fragmental point cloud	Reconstructed point cloud	Chamfer Distance	Fragmental point cloud	Reconstructed point cloud	Chamfer Distance
		2.3×10^{-4}			1.2×10^{-4}
		3.7×10^{-4}			0.7×10^{-4}

point clouds are very small, verifying the point clouds reconstructed by the proposed CLSCN and FLPRA are closely analogous to the original point clouds. The difference between the reconstructed point cloud and the original point cloud mainly comes from the processing of points and Outlier. The missing part of the fragmental point cloud needs to be reconstructed after filtering, and the reconstructed regions depend on the spatial information of their neighborhoods, so the CD values turn larger when the area of the missing part increases.

3.2. Test of canopy reconstruction

Due to the point cloud reconstruction of leaves are carried out one by one, it is necessary to restore the reconstructed leaves to the corresponding spatial position for constituting an entire cotton canopy. When the leaves are segmented and extracted from the cotton canopy one by one, the spatial location information is recorded including the center point of each cotton leaf. Because the filter and extension of leaf point clouds are based on the center point, based on the space coordinate where one of the leaf point clouds is located, all remaining leaf point clouds relative to their center points can be determined in the original cotton canopy space. All reconstructed cotton leaf point clouds are placed in the same spatial coordinate system to complete an entire canopy reconstruction. Compared with the point cloud of the cotton canopy acquired with RealSense L515, the reconstructed cotton canopy has restored the internal occlusion structure of the canopy, and the point cloud of the canopy is smoother with no outliers, as in Fig. 14. Besides, the regions circled in red were the regions reconstructed for the fragmental parts of cotton leaves. Furthermore, using the Delaunay Triangle mesh generation algorithm based on the Bowyer-Watson method, the reconstructed point clouds are packaged to form a three-dimensional model of the cotton canopy.

The relationship between reconstruction depth and accuracy was furtherly analyzed. The horizontal plane crossing the highest point of the canopy is defined as S, and the vertical distance from a certain point inside the canopy to S is defined as the reconstruction depth. The point cloud reconstruction of cotton plants is performed on several depth levels, which gradually increase in steps of 0.05 m. As shown in Fig. 15, when the depth increases, the reconstruction accuracy of the occluded leaves inside cotton canopies shows a decreasing trend. It reached 82.70 % at 0.05 m depth and decreased to 66.70 % while the reconstruction depth is up to 0.6 m. Compared to only acquire surface information of canopy from a single viewpoint with RGB-D sensor, the proposed methods in this work recover the inner canopy structure with the accuracy maintains above 70 % deeper than 0.4 m depth. The main reason is that as the reconstruction depth increases, the occlusion rates of leaves and branches also increase, the available information for leaves completion becomes less, and it is more difficult to recover the occluded images of cotton leaves, finally resulting in a decreasing overlap rate between the completed leaves by back-end GAN and the ideal completed leaves. In the test, the entire reconstruction process took less than 4.7 s from acquiring RGB-D information from the top view to finishing the corresponding three-dimensional reconstruction. The system has high reconstruction efficiency and can achieve high-throughput three-dimensional reconstruction of cotton canopies.

3.3. Discussion

The efficient three-dimensional reconstruction method proposed in this work mainly contains four steps: data acquisition, instance segmentation of RGB image, point cloud extension, spatial recombination of leaves and branches, etc. It can recover the internal occluded structure, instead of just acquiring the surface information of the canopy

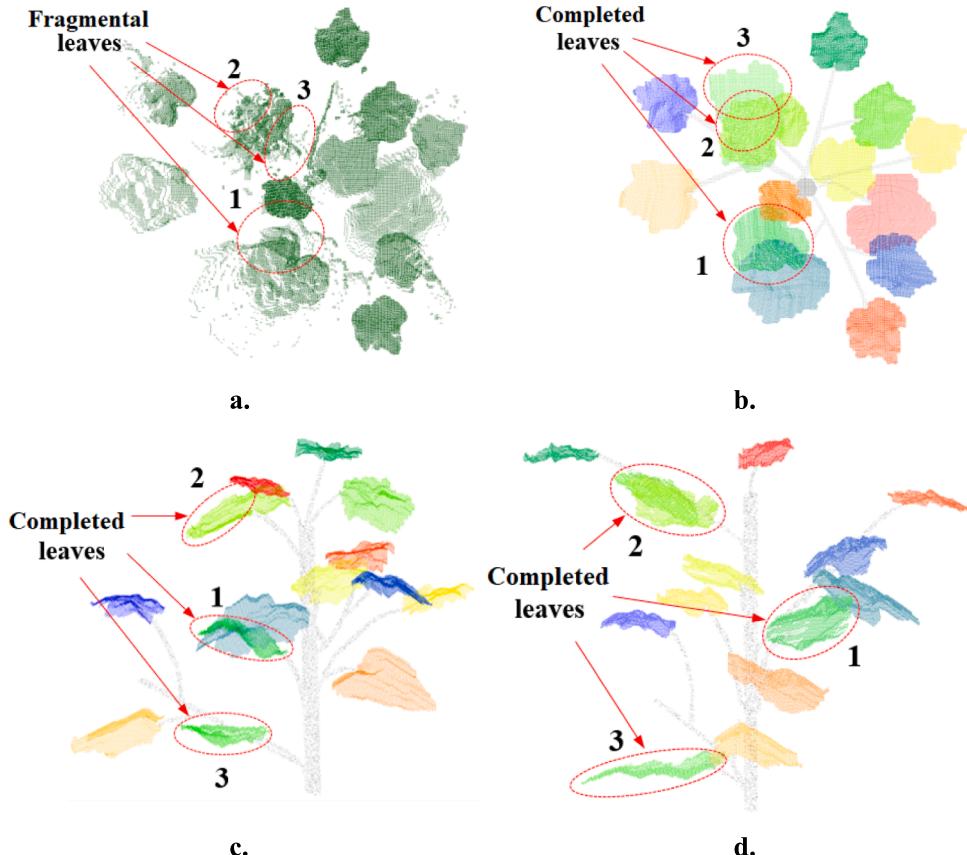


Fig. 14. Reconstruction result of the cotton canopy. a. original point cloud, b. top-view after reconstruction, c. main-view after reconstruction, d. left-view after reconstruction.

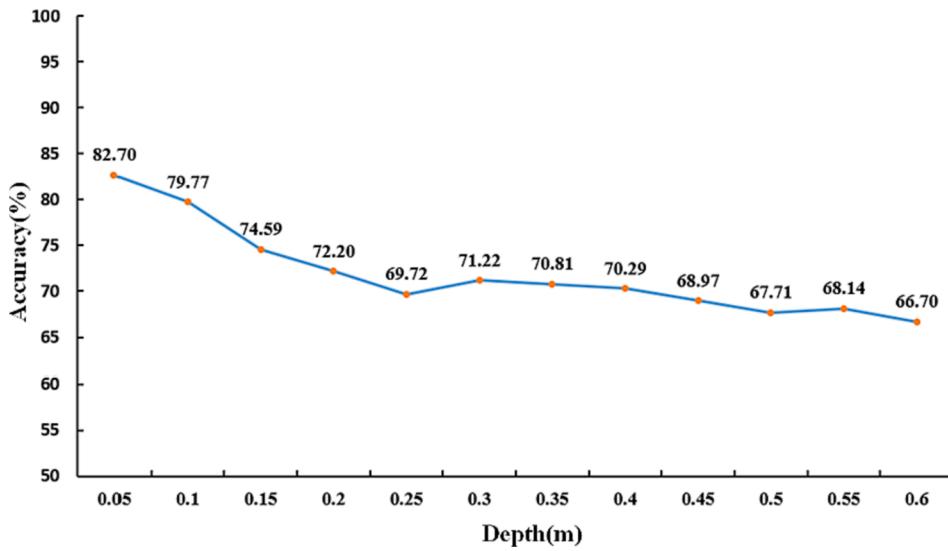


Fig. 15. Relationship between the accuracy and the depth of canopy reconstruction.

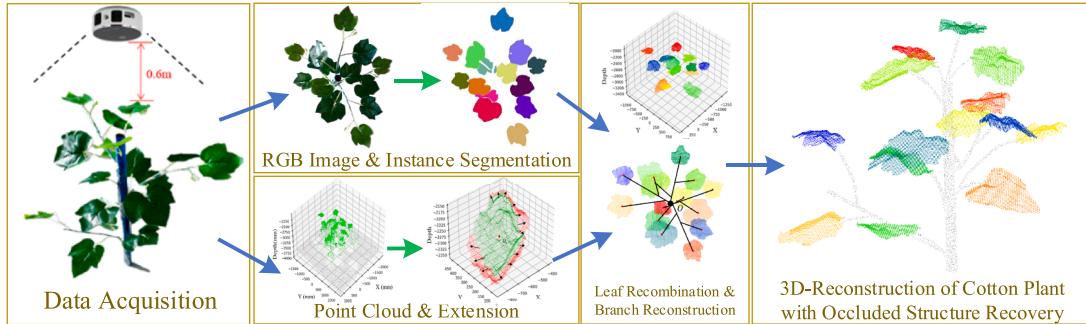


Fig. 16. Reconstruction process of internal structure of cotton canopy.

(Fig. 16). The proposed method has shown a better performance while the qualities of acquired images and point clouds are poor due to many unfavorable factors, such as occlusion among leaves, complex ambient light, the surface reflectivity of canopies and so on. Normally, because of the mutual occlusion of the leaves in the canopy, the information captured by the RGB-D sensor is adequate for learning about the internal canopy structure. The proposed method makes it possible to well reconstruct the occluded structure by an RGB-D sensor in a single view. Compared with the raw point cloud directly acquired by the sensor, the reconstructed point cloud showed the following merits: good consistency with the raw data, complete recovery of the missing points, fewer outliers, smooth spatial continuity.

However, compared to other fine modeling algorithms, the mean reconstruction accuracy of only 71.90% within a depth of 0.6 m is not overly impressive, especially considering the accuracy of only 82.70% for reconstructing the surface of complex canopies. The main reason is that our method focuses on reconstructing the occluded internal parts of a canopy, which enables us to maintain the internal and external integrity of a three-dimensional canopy model from a single viewpoint—a capability that has not been mentioned in previous researches. The tests in this work have verified the feasibility of the proposed method, but the results also indicate that more work needs to be done to improve these networks and algorithms. As mentioned in Section 3.2, as the reconstruction depth increasing, the occluded area increases and the accuracy of reconstructing the internal structure declines rapidly. With an increasing occlusion rate of the internal canopy, less information is available to the leaf GAN, and leaf image recovery becomes more difficult. Therefore, there are positioning errors between the

reconstructed point clouds and the acquired fragmental point clouds, leading to accuracy losses in reconstructing the internal canopy. In addition, environmental stresses, including but not limited to high temperature, low temperature, drought, salinity, heavy metals and so on, have adverse effects on the normal physiological activities of plants. These factors lead to significant differences between real cotton leaves and artificial cotton leaves, limiting the generalizability of the method proposed in this paper.

4. Conclusions and future work

The methods proposed in this work is to deal with the problem of hard construction of crop canopy with occluded inner structures using only external information of the canopy. Compared with traditional plant reconstruction, deep learning network is integrated to construct three-dimensional models, and the test results proved that the proposed CLSCN and FLPRA can greatly improve the integrity of the reconstructed canopy models. The front-end segmentation network of the CLSCN can output high-quality cotton leaf masks, of which the mIoU can be up to 84.65%, and the back-end GAN of CLSCN can complete the occluded leaves with an accuracy of over 94%; the reconstruction accuracy of the final three-dimensional model of the cotton canopy is up to 82.7%. It is an important theoretical and technical support to realize real-time crop status observation and precise field management in agriculture production.

However, due to the large number of parameters of the neural network, it needs much time and work for the reconstruction optimization and accuracy improvement. In the future, the instance

segmentation and image completion operations can be performed using lightweight neural networks to reduce the reconstruction time, and carrying out more tests to optimize reconstruction methods for cotton leaves under environmental stress is also one of the important future works. Meanwhile, the proposed reconstruction methods should be applied to other crops for enhancing the generalization ability of the model.

CRediT authorship contribution statement

Yang Li: Conceptualization, Methodology, Writing – original draft, Funding acquisition. **Shuke Si:** Data curation, Resources, Software, Writing – original draft. **Xinghua Liu:** Validation, Writing – review & editing, Funding acquisition. **Liangliang Zou:** Investigation. **Wenqian Wu:** Formal analysis, Visualization. **Xuemei Liu:** Writing – review & editing. **Li Zhang:** 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.

Acknowledgements

This study was supported by the National Natural Science Foundation of China (32271992), Shandong Province Key R&D Plan Project (Major Science and Technology Innovation) (2022LZGC020), Natural Science Foundation of Shandong Province (ZR2023MC083) and Project of Mechanical Post Expert of Shandong Cotton Industry Innovation Team (SDAIT-03-09).

References

- Al-Shurbaji, T.A., AlKaabneh, K.A., Alhadid, I., Masadeh, R., 2022. An Optimized Scale-Invariant Feature Transform Using Chamfer Distance in Image Matching. *Intell. Automation Soft Comput.* 31, 971–985. 10.32604/iasc.2022.019654.
- Chopin, J., Fasquel, J.-B., Mouchère, H., Dahyot, R., Bloch, I., 2023. Model-based inexact graph matching on top of DNNs for semantic scene understanding. *Comput. Vis. Image Underst.* 235, 103744. <https://doi.org/10.1016/j.cviu.2023.103744>.
- Cuevas-Velasquez, H., Gallego, A.-J., Fisher, R.B., 2020. Segmentation and 3D reconstruction of rose plants from stereoscopic images. *Comput. Electron. Agric.* 171, 105296. <https://doi.org/10.1016/j.compag.2020.105296>.
- Das Choudhury, S., Maturu, S., Samal, A., Stoerger, V., Awada, T., 2020. Leveraging Image Analysis to Compute 3D Plant Phenotypes Based on Voxel-Grid Plant Reconstruction. *Front. Plant Sci.* 11. <https://doi.org/10.3389/fpls.2020.521431>.
- Dugdale, S.J., Malcolm, I.A., Hannah, D.M., 2019. Drone-based Structure-from-Motion provides accurate forest canopy data to assess shading effects in river temperature models. *Sci. Total Environ.* 678, 326–340. <https://doi.org/10.1016/j.scitotenv.2019.04.229>.
- Espejo-Garcia, B., Mylonas, N., Athanasakos, L., Vali, E., Fountas, S., 2021. Combining generative adversarial networks and agricultural transfer learning for weeds identification. *Biosyst. Eng.* 204, 79–89. <https://doi.org/10.1016/j.biosystemseng.2021.01.014>.
- Fang, W., Feng, H., Yang, W., Duan, L., Chen, G., Xiong, L., Liu, Q., 2016. High-throughput volumetric reconstruction for 3D wheat plant architecture studies. *J. Innov. Opt. Health Sci.* 09, 1650037. <https://doi.org/10.1142/S1793545816500371>.
- Gibbs, J.A., Pound, M.P., French, A.P., Wells, D.M., Murchie, E.H., Pridmore, T.P., 2020. Active Vision and Surface Reconstruction for 3D Plant Shoot Modelling. *IEEE/ACM Trans. Comput. Biol. Bioinform.* 17, 1907–1917. <https://doi.org/10.1109/TCBB.2019.2896908>.
- Gigliani, R., Magnanini, E., Nerozzi, F., Muzzi, E., Sinoquet, H., 2005. Canopy probabilistic reconstruction inferred from Monte Carlo point-intercept leaf sampling. *Agric. For. Meteorol.* 128, 17–32. <https://doi.org/10.1016/j.agrformet.2004.09.003>.
- Guo, J., Liu, Y., 2019. Image completion using structure and texture GAN network. *Neurocomputing* 360, 75–84. <https://doi.org/10.1016/j.neucom.2019.06.010>.
- Huang, E., He, Z., Mao, A., Camila Ceballos, M., Parsons, T.D., Liu, K., 2023. A semi-supervised generative adversarial network for amodal instance segmentation of piglets in farrowing pens. *Comput. Electron. Agric.* 209, 107839. <https://doi.org/10.1016/j.compag.2023.107839>.
- Li, Z., Li, Y., Yang, Y., Guo, R., Yang, J., Yue, J., Wang, Y., 2021. A high-precision detection method of hydroponic lettuce seedlings status based on improved Faster RCNN. *Comput. Electron. Agric.* 182, 106054. <https://doi.org/10.1016/j.compag.2021.106054>.
- Liu, X., Liu, X., Li, Y., Yuan, J., Song, L., Li, H., Wu, M., 2020. Estimation model of canopy stratification porosity based on morphological characteristics: A case study of cotton. *Biosyst. Eng.* 193, 174–186. <https://doi.org/10.1016/j.biosystemseng.2020.02.018>.
- Lu, Y., Chen, D., Olaniyi, E., Huang, Y., 2022. Generative adversarial networks (GANs) for image augmentation in agriculture: A systematic review. *Comput. Electron. Agric.* 200, 107208. <https://doi.org/10.1016/j.compag.2022.107208>.
- Madsen, S.L., Dyrmann, M., Jørgensen, R.N., Karstoft, H., 2019. Generating artificial images of plant seedlings using generative adversarial networks. *Biosyst. Eng.* 187, 147–159. <https://doi.org/10.1016/j.biosystemseng.2019.09.005>.
- Nguyen, T.T., Slaughter, D.C., Max, N., Maloof, J.N., Sinha, N., 2015. Structured Light-Based 3D Reconstruction System for Plants. *Sensors* 15, 18587–18612. <https://doi.org/10.3390/s150818587>.
- Okura, F., 2022. 3D modeling and reconstruction of plants and trees: A cross-cutting review across computer graphics, vision, and plant phenotyping. *Breed. Sci.* 72, 31–47. <https://doi.org/10.1270/jbsbs.21074>.
- Pagliari, A., Ammoniaci, M., Sarri, D., Lisci, R., Perria, R., Vieri, M., D'Arcangelo, M.E., Storchi, P., Kartsiotis, S.-P., 2022. Comparison of Aerial and Ground 3D Point Clouds for Canopy Size Assessment in Precision Viticulture. *Remote Sens. (Basel)*. <https://doi.org/10.3390/rs14051145>.
- Rapado-Rincón, D., van Henten, E.J., Kootstra, G., 2023. Development and evaluation of automated localisation and reconstruction of all fruits on tomato plants in a greenhouse based on multi-view perception and 3D multi-object tracking. *Biosyst. Eng.* 231, 78–91. <https://doi.org/10.1016/j.biosystemseng.2023.06.003>.
- Shi, W., van de Zedde, R., Jiang, H., Kootstra, G., 2019. Plant-part segmentation using deep learning and multi-view vision. *Biosyst. Eng.* 187, 81–95. <https://doi.org/10.1016/j.biosystemseng.2019.08.014>.
- Solomon, J., Rustamov, R., Guibas, L., Butscher, A., 2014. Earth Mover's Distances on Discrete Surfaces. *ACM Trans. Graph.* 33. <https://doi.org/10.1145/2601097.2601175>.
- Song, Z., Zhou, Z., Wang, W., Gao, F., Fu, L., Li, R., Cui, Y., 2021. Canopy segmentation and wire reconstruction for kiwifruit robotic harvesting. *Comput. Electron. Agric.* 181, 105933. <https://doi.org/10.1016/j.compag.2020.105933>.
- 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. <https://doi.org/10.3390/agriculture10100462>.
- Wei, B., Ma, X., Guan, H., Yu, M., Yang, C., He, H., Wang, F., Shen, P., 2023. Dynamic simulation of leaf area index for the soybean canopy based on 3D reconstruction. *Eco. Inform.* 75, 102070. <https://doi.org/10.1016/j.ecoinf.2023.102070>.
- Yin, K., Huang, H., Long, P., Gaisinski, A., Gong, M., Sharf, A., 2016. Full 3D Plant Reconstruction via Intrusive Acquisition. *Comput. Graphics Forum* 35, 272–284. <https://doi.org/10.1111/cgf.12724>.
- Zhu, Rongsheng, L.S., Sun, Yongzhe, Cao, Yangyang, Sun, Kai, Guo, Yixin, Jiang, Bo Feng, Wang, Xueying, Li, Yang, Zhang, Zhanguo, Xin, Dawei, Hu, Zhenbang, Chen, Qingshan, 2021. Research Advances and Prospects of Crop 3D Reconstruction Technology. *Smart Agric.* 3, 94–115. 10.12133/j.smartag.2021.3.3.202102-SA002.
- Zhu, T., Ma, X., Guan, H., Wu, X., Wang, F., Yang, C., Jiang, Q., 2023a. A calculation method of phenotypic traits based on three-dimensional reconstruction of tomato canopy. *Comput. Electron. Agric.* 204, 107515. <https://doi.org/10.1016/j.compag.2022.107515>.
- Zhu, Z., Yang, L., Lin, X., Yang, L., Liang, Y., 2023b. GARNet: Global-aware multi-view 3D reconstruction network and the cost-performance tradeoff. *Pattern Recognit.* 142, 109674. <https://doi.org/10.1016/j.patcog.2023.109674>.