

Article

A Method for Identifying Picking Points in Safflower Point Clouds Based on an Improved PointNet++ Network

Baojian Ma ¹, Hao Xia ^{2,*}, Yun Ge ^{2,*}, He Zhang ², Zhenghao Wu ², Min Li ² and Dongyun Wang ³

¹ Department of Mechanical and Electrical Engineering, Xinjiang Institute of Technology, Aksu 843100, China; 11813019@zju.edu.cn

² College of Mechanical and Electrical Engineering, Shihezi University, Shihezi 832003, China; zhangh@stu.shzu.edu.cn (H.Z.); wuzhenghao@stu.shzu.edu.cn (Z.W.); limin@stu.shzu.edu.cn (M.L.)

³ College of Engineering, Zhejiang Normal University, Jinhua 321004, China; zsdwdy@zjnu.edu.cn

* Correspondence: xiahao@stu.shzu.edu.cn (H.X.); gy_mac@shzu.edu.cn (Y.G.); Tel.: +86-188-5398-2140 (H.X.); +86-189-9933-0088 (Y.G.)

Abstract: To address the challenge of precise picking point localization in morphologically diverse safflower plants, this study proposes PointSafNet—a novel three-stage 3D point cloud analysis framework with distinct architectural and methodological innovations. In Stage I, we introduce a multi-view reconstruction pipeline integrating Structure from Motion (SfM) and Multi-View Stereo (MVS) to generate high-fidelity 3D plant point clouds. Stage II develops a dual-branch architecture employing Star modules for multi-scale hierarchical geometric feature extraction at the organ level (filaments and fruit balls), complemented by a Context-Anchored Attention (CAA) mechanism to capture long-range contextual information. This synergistic feature learning approach addresses morphological variations, achieving 86.83% segmentation accuracy (surpassing PointNet++ by 7.37%) and outperforming conventional point cloud models. Stage III proposes an optimized geometric analysis pipeline combining dual-centroid spatial vectorization with Oriented Bounding Box (OBB)-based proximity analysis, resolving picking coordinate localization across diverse plants with 90% positioning accuracy and 68.82% mean IoU (13.71% improvement). The experiments demonstrate that PointSafNet systematically integrates 3D reconstruction, hierarchical feature learning, and geometric reasoning to provide visual guidance for robotic harvesting systems in complex plant canopies. The framework's dual emphasis on architectural innovation and geometric modeling offers a generalizable solution for precision agriculture tasks involving morphologically diverse safflowers.

Keywords: safflower; picking point; 3D point cloud; deep learning; harvesting robots



Academic Editor: Paul Kwan

Received: 29 March 2025

Revised: 27 April 2025

Accepted: 30 April 2025

Published: 2 May 2025

Citation: Ma, B.; Xia, H.; Ge, Y.; Zhang, H.; Wu, Z.; Li, M.; Wang, D. A Method for Identifying Picking Points in Safflower Point Clouds Based on an Improved PointNet++ Network. *Agronomy* **2025**, *15*, 1125. <https://doi.org/10.3390/agronomy15051125>

Copyright: © 2025 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Safflower (*Carthamus tinctorius* L.), a typical terminal-flowering economic crop, possesses remarkable development value. It is widely applied in the pharmaceutical industry, oil production, and feed sectors [1], with significant potential particularly evident in the deep processing and development of the pharmaceutical and healthcare industries. Geographically, China's Xinjiang region has become a globally important safflower industry base due to its unique ecological conditions. Accounting for over 80% of the national planting area [2], it has established a large-scale cultivation pattern. The crop exhibits continuous productivity throughout its growth cycle, enabling three to five successive harvests. Harvest timing significantly influences the stability of its medicinal components: safflower achieves the highest medicinal value three days after blooming. Thereafter, as

the water content gradually decreases, its medicinal value declines. Failure to harvest the filaments from the capitulum in a timely manner will disrupt subsequent flowering, leading to reduced filament yield [3]. Currently, the development of the safflower industry is severely restricted by bottlenecks in harvesting technology. The manual harvesting mode is characterized by extremely low efficiency and exorbitantly high costs. Moreover, existing mechanized harvesting equipment, due to the marked heterogeneity in plant spatial arrangement, is highly prone to positioning errors, leading to a relatively low harvesting success rate and a significantly increased probability of filament mechanical damage [4–6].

In contrast to traditional mechanized systems, robotic harvesting solutions demonstrate superior adaptability to complex field environments and enable precision operation through safflower picking location recognition. Therefore, accurate identification of the spatial coordinates at the filament–fruit ball junction of safflowers constitutes the core determinant for successful safflower harvesting. To achieve precise localization, researchers urgently need to overcome detection interference caused by morphological variations in plants. Zhang et al. enhanced YOLOv3’s robustness in complex environments via backbone network improvement, pooling layer optimization, and attention mechanism integration, attaining 90.89% mean average precision (mAP) [7]. Furthermore, Wang et al. augmented YOLOv7 by integrating Swin Transformer attention and optimizing the loss function, achieving accurate recognition and localization of three safflower targets with 88.5% mAP for multi-category samples [8]. Subsequently, Zhang et al. proposed the improved Faster R-CNN by integrating ResNeSt-101 backbone, achieving 90.49% mAP for safflower filament detection in natural environments [3]. To overcome the limitation of computational latency, Chen et al. developed YOLO-SaFi through YOLOv8n modifications to achieve 93.9% mAP for real-time filament detection in unstructured environments, notwithstanding persistent challenges posed by backlighting and occlusion scenarios [1]. Other terminal-blooming flowers have also attracted focused research attention. For instance, Zhao et al. proposed CR-YOLOv5s by integrating coordinate attention mechanisms into the backbone and replacing conventional convolutions with RepVGG blocks, achieving 93.9% mAP and a 4.5% accuracy improvement over the baseline model for chrysanthemum inflorescence detection [9]. Concurrently, Park et al. developed a lightweight YOLOv4-Tiny model utilizing circular bounding box regression and geometric feature encoding, achieving dynamic classification of chrysanthemum growth cycles and precise differentiation of flowering stages to support staged harvesting decision-making [10]. In their research on jasmine flowers, Zhou et al. implemented a YOLOv7-based framework for real-time jasmine flower detection, achieving 94.8% mAP through five-tier bloom-stage classification that addresses temporal flowering variations in field conditions [11]. In research on woody flowers, intelligent monitoring of rose crops has also attracted the attention of scholars. Shinoda et al. proposed RoseTracker, an automatic rose-growth-monitoring system integrating YOLOv5 object detection, SORT tracking, and regression models to detect rose buds and blooming flowers from top-view videos, addressing missed detection issues in traditional methods and achieving an F1 score of 0.950 in experiments [12]. To address the need for real-time field detection of Damask roses, Fatehi et al. optimized the YOLOv9t model via knowledge distillation (KD), maintaining model lightness (4.43 MB) while improving mAP@0.5 by 0.3% and 0.2%, and enhancing detection speeds by 5.1 FPS and 1.8 FPS, thus effectively balancing accuracy and computational efficiency [13]. Meanwhile, in the field of flower classification, Gupta et al. proposed the Flora-NET model for medicinal flower classification, which uses the DCAFE module and Inv-FR module for feature extraction and optimization, achieving classification accuracies of 91.12% and 91.18% on the Urban Street and Medicinal Blossom datasets, respectively [14]. The above-mentioned literature is

limited to the target recognition of flowers and does not involve research on the accurate positioning method of picking points.

By leveraging flower recognition technology to achieve precise picking position determination, researchers have conducted systematic research on terminal flower picking point localization. The existing methodologies can be predominantly categorized into two types: image-based localization approaches and point cloud-based localization approaches. The image-processing-based localization first segments target crops by leveraging distinct feature differences (e.g., color, contour, and texture) between crops and their backgrounds. Subsequently, the optimal picking position is determined through computational analysis of the spatial geometric configurations. To achieve automated picking of Hangzhou white chrysanthemums, Yang et al. achieved target positioning by segmenting images using a fast FCM algorithm based on the S component of the HSV color space and calculating 3D coordinates via camera calibration [15]. Building upon this foundation, Yang et al. proposed a least squares support-vector-machine-based segmentation algorithm for Hangzhou white chrysanthemums, enabling 3D coordinate calculation and subsequent picking point localization through algorithmic processing [16]. Meanwhile, Xing et al. developed an improved PSO-rotating rectangle fusion model for filament neck recognition and a 3D spatial solution-based localization strategy, achieving 89.75% positioning accuracy [17]. However, traditional image-processing methods are prone to environmental variable interference during picking point identification, resulting in insufficient robustness [18]. In response, deep learning approaches exhibit significant advantages by automatically learning features from deep neural networks and achieve pixel-level segmentation of organs and complete picking point localization by integrating domain-specific prior rules. For instance, Wang et al. proposed a two-stage identification and positioning method for safflower picking points based on the improved Yolov5 algorithm, achieving rapid recognition with a 98% identification success rate [19]. Zhang et al. developed an enhanced YOLOv8s-Seg architecture integrating Principal Component Analysis (PCA)-guided Region of Interest (ROI) extraction and contour fitting algorithms, achieving 92.9% positioning accuracy for safflower picking points [20]. Although RGB image-based picking point identification methods have been widely applied, they are constrained by the two-dimensional perspective, making it difficult to accurately capture object depth information. This leads to potential deviations in the judgment of target object spatial positions and postures. Determining picking points through the fusion of RGB images and depth maps has attracted significant attention from researchers. For example, Chen et al. achieved 83% picking point positioning accuracy for tea buds through YOLOv3-driven segmentation of RGB-D data, leveraging depth-enhanced feature fusion for target localization [21]. Meanwhile, Ge et al. employed Mask R-CNN for instance segmentation of strawberry images to generate segmentation masks, fused the depth maps, and implemented coordinate transformation, density-based clustering, and position optimization algorithms to achieve a harvesting accuracy of 74.1% [22]. Notably, Li et al. developed a 3D localization method tailored to the growth characteristics of tea shoots by fusing RGB images and depth maps, yet positioning deviations were observed in complex environments due to diverse foliage postures [23]. To address this limitation, Li et al. employed lightweight YOLOv3 for tea shoot recognition, fused RGB images depth maps, and achieved harvest point localization by integrating tea shoot growth characteristics [24]. In addition, Zhu et al. utilized an improved YOLOv5 for tea bud detection, fused RGB images depth maps, and computed 3D harvesting coordinates via principal component analysis (PCA) and cuboid fitting, ultimately achieving a 94.4% accuracy rate [25]. While RGB-D images provide depth information, depth maps captured under natural conditions may exhibit data dropout, which not only compromises

the integrity of plant spatial representation but also introduces picking localization errors, introducing new challenges for robotic visual perception systems.

Compared to image-based picking point determination, 3D point cloud analysis enhances localization accuracy by analyzing crop spatial geometric features. Conventional approaches typically employ multi-view 3D reconstruction techniques to acquire plant point clouds, followed by geometric feature analysis for harvesting point localization. For instance, Yoshida et al. proposed a point cloud-based cutting point localization method for tomato-harvesting robots, which voxelizes tomato point clouds and identifies cutting points, achieving 100% accuracy [26]. In another study, Díaz et al. reconstructed grape point clouds via Structure from Motion (SfM), combined 2D sliding window classification with DBSCAN clustering, and localized bud positions using centroid localization, achieving a recall of 0.45 and a localization error of 1.5 cm [27]. While conventional methods have been widely adopted for 3D picking point determination, they inherently suffer from three critical limitations: dependency on manual feature engineering, lack of semantic comprehension capabilities, and inadequate adaptability to unstructured environments or dynamic scenarios. In contrast, deep-learning-driven point cloud segmentation has been widely applied in agricultural robotics. Complex perception challenges are addressed through end-to-end semantic-aware modeling. In inter-plant recognition, three-dimensional point cloud data of plants enable precise differentiation of adjacent plant boundaries based on spatial distribution characteristics and geometric morphology parameters of point clouds, solving the recognition errors caused by plant occlusion in traditional methods and achieving automated segmentation of crop organs [28–30], thus providing accurate data support for rational close planting. In the field of growth monitoring, three-dimensional plant models constructed from point cloud data enable dynamic tracking of key growth indicators such as plant height, crown width, and number of branches. Compared with manual measurement, this approach enables high-frequency, non-contact monitoring, significantly improving data acquisition efficiency and timeliness. In yield prediction, by analyzing parameters such as the number of inflorescences and single-inflorescence volume from point cloud data and establishing yield prediction models using deep learning algorithms, the limitations of traditional empirical prediction can be broken through [31–33]. However, the direct application of deep learning to point cloud data for picking point identification—especially in the specific scenario of safflower harvesting—remains in its infancy.

This paper presents a novel approach for high-precision identification of safflower picking points by integrating deep learning and point cloud analysis. The methodology begins with multi-view RGB image acquisition of safflower plants using a consumer-grade camera, followed by high-fidelity 3D point cloud reconstruction via the SfM-MVS algorithm. The acquired point cloud data are then preprocessed (denoising and downsampling) and semantically annotated using CloudCompare Version 2 to construct a training dataset. Finally, we propose an improved point cloud segmentation model—PointSafNet—and develop a picking point identification framework tailored for safflower harvesting. The primary contributions of this study are as follows:

- (1) A high-fidelity 3D point cloud reconstruction framework for safflower plants in natural scenes is developed, along with a systematic preprocessing pipeline—including denoising, downsampling, and registration—tailored specifically for safflower point cloud data.

- (2) By integrating multi-layer perceptron (MLP) modules with attention mechanisms, we present PointSafNet—an enhanced variant of PointNet++—to improve the semantic segmentation accuracy of safflower point clouds.

- (3) A novel methodology for safflower picking point identification is introduced, combining point cloud processing and geometric feature analysis. First, point cloud

masks of filaments and fruit balls are extracted, with their centroids computed to establish spatial vector relationships. Next, the Oriented Bounding Box (OBB) algorithm performs minimum-volume fitting on fruit ball point clouds, enabling precise determination of 3D picking coordinates through geometric analysis.

2. Materials and Methods

2.1. The Overall Process of Identifying Safflower Picking Points

The overall workflow of the safflower picking point localization method proposed in this study is illustrated in Figure 1, which primarily consists of the following steps: (1) Multi-view images of a single safflower plant in outdoor environments are captured using an RGB camera, followed by 3D reconstruction via SfM-MVS algorithms. Subsequently, the generated point cloud data of the safflower plant are preprocessed. (2) The proposed PointSafNet model is applied to conduct semantic segmentation on the safflower plant point cloud, obtaining point cloud masks for filaments, fruit balls, stems, and leaves, thus enabling precise identification of different plant parts. (3) The centroids of the filament and fruit ball point clouds are computed, and their spatial vector relationship is determined by using the Oriented Bounding Box (OBB) method. Moreover, the 3D picking point location of the safflower is determined through the minimum distance of the OBB. This research introduces an innovative method for the recognition of safflower picking points.

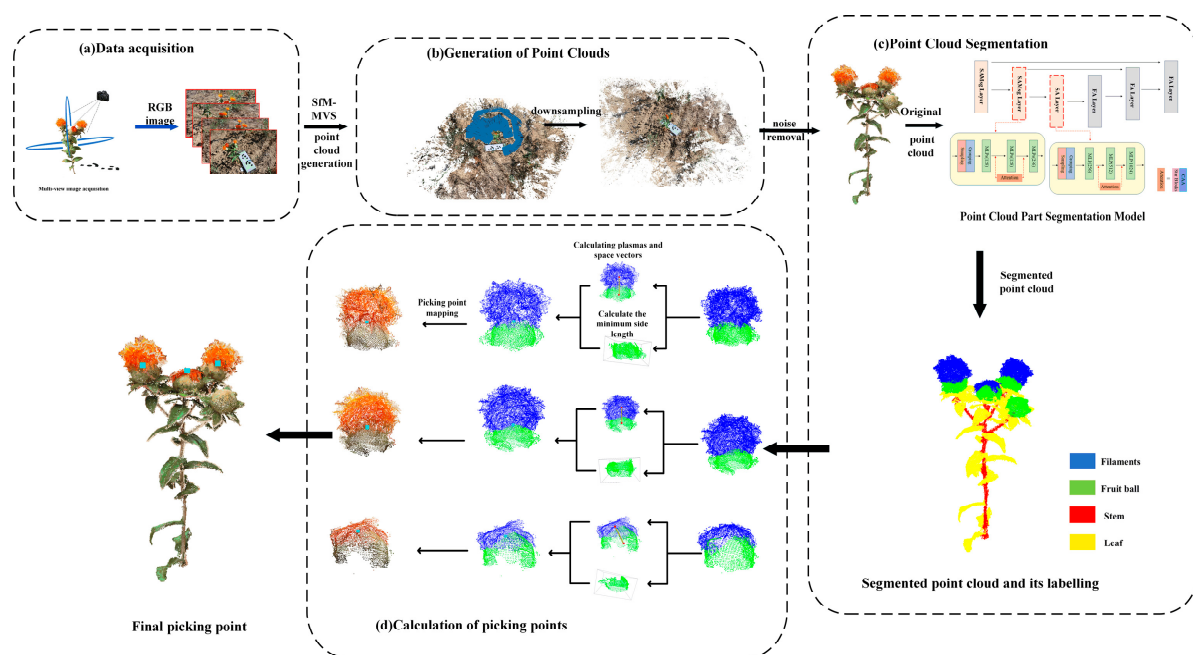


Figure 1. Safflower picking point determination process.

2.2. Three-Dimensional Reconstruction and Preprocessing of Safflower Plants

To systematically acquire observational data covering the full flowering period of safflower, this study conducted field data acquisition in a dryland safflower experimental field located in Yumin County (46.19° N, 82.93° E), Tacheng Prefecture, Xinjiang Uygur Autonomous Region, China. Data collection was primarily performed under overcast conditions or during early-morning low-intensity light periods to mitigate the interference of direct sunlight on image acquisition quality. For image capture, a Canon EOS R50 camera (Canon Inc., Tokyo, Japan) was used to obtain multi-view images of individual safflower plants, with natural lighting maintained throughout the process. Based on the acquired multi-source image dataset (6000 × 4000 pixel resolution), high-precision 3D point clouds

of 150 safflower plants were reconstructed using the Structure-from-Motion (SfM) and Multi-View Stereopsis (MVS) techniques.

Three-dimensional reconstruction of safflower plants was conducted using the Structure from Motion-Multi-View Stereo (SfM-MVS) approach to derive raw point cloud data. As an economical 3D reconstruction technique, SfM-MVS accurately recovers object geometry by acquiring sequential multi-view imagery [34]. The implementation process involves: (1) detecting local keypoints using the algorithm illustrated in Figure 2, followed by Scale-Invariant Feature Transform (SIFT) [35] for image feature extraction; (2) matching detected keypoints between image pairs to establish cross-view correspondences; (3) calculating camera intrinsic/extrinsic parameters via triangulation to determine the 3D spatial coordinates of matched keypoints, thereby achieving a 3D reconstruction of safflower plants. At the technical detail level, the fundamental matrix F is computed from keypoint pairs (x_1, x_2) , and the camera parameters are estimated through $x_1^T F x_2 = 0$ using matched points. Based on camera intrinsics and its position, matched keypoint positions (termed sparse 3D point clouds) are generated via triangulation. Finally, combining camera parameters and computed sparse points, dense point clouds are generated using multi-view stereo matching algorithms [36]. The 3D reconstruction pipeline was implemented using Agisoft Metashape (Agisoft LLC, Saint Petersburg, Russia), a software suite capable of generating high-precision point clouds [37]. During reconstruction, the safflower point cloud exhibited high density, with processing time dependent on the number of images (average duration: approximately 20 min per plant). Each plant's 3D model typically contains 1–2 million points, including safflower plant structures, ground surfaces, calibration targets, and residual noise requiring post-processing removal.

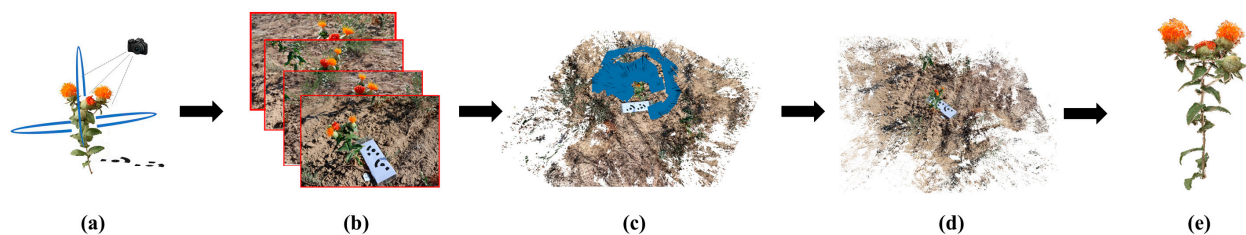


Figure 2. Three-dimensional reconstruction and preprocessing of safflower Plants. (a) Schematic diagram of safflower shooting angles. (b) Multi-view images of safflower. (c) Original point cloud. (d) Point cloud after downsampling. (e) Processed point cloud.

To address the challenge of numerous outliers in safflower point clouds and the associated low efficiency of manual curation, this study employed random downsampling to subsample point clouds [38]. This method reduces data density while preserving spatial distribution, thereby improving computational efficiency. Additionally, random downsampling enhances model robustness, generalization capacity, and performance, facilitating more efficient subsequent analysis. A radius filtering algorithm was then applied for noise removal, which demonstrates superior performance in preserving shape and edge information while maintaining rapid processing speeds [39]. This algorithm effectively handles large quantities of outliers that are difficult to manually remove. These preprocessing steps establish a reliable data foundation for subsequent precise analysis of safflower point clouds. Finally, the point cloud data were exported as an $M \times 6$ array, where M denotes the number of points. The six columns represent X , Y , Z coordinates and corresponding red (R), green (G), and blue (B) color channels, respectively. The data were saved as a TXT file for downstream processing and analysis.

2.3. Annotation of the Point Cloud of Safflower Plants and Model Training

Semantic annotation of safflower point clouds was performed using CloudCompare software (Figure 3). The detailed workflow is as follows: (1) Preprocessed point cloud data were imported into CloudCompare. Using the software's integrated segmentation toolset, precise segmentation of safflower filaments, fruit balls, stems, and leaves was achieved based on morphological characteristics, with concurrent semantic annotation. (2) Segmented and annotated part-specific point clouds were aggregated into complete safflower point cloud datasets. A total of 354 fully annotated safflower point clouds were constructed, following a data format analogous to that of the ShapeNet open-source dataset [40]. Each annotated safflower point cloud contains XYZ coordinates, RGB color values, and semantic class labels. The safflower point cloud dataset consists of 354 point clouds, each representing a single safflower plant. Each plant contains 4 organs, which were independently annotated by 3 personnel. The data were stored in .ply file format. Using stratified random sampling, the dataset was divided into a training set (284 samples), a validation set (35 samples), and a test set (35 samples) at an 8:1:1 ratio.

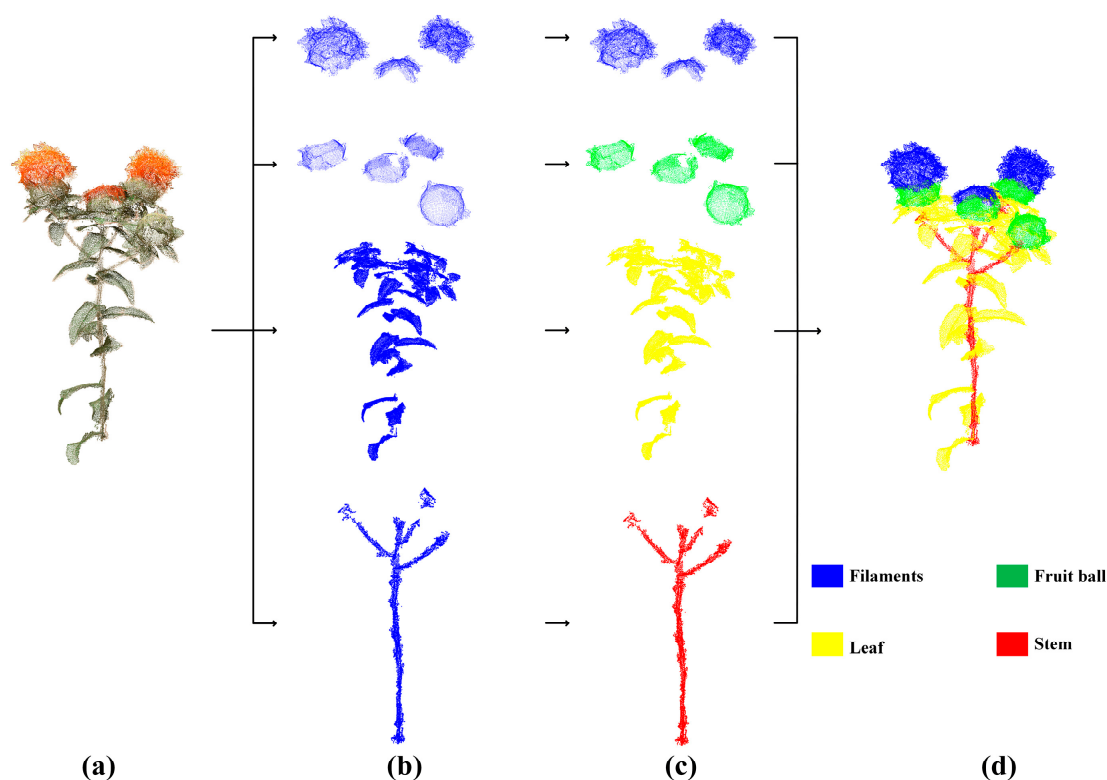


Figure 3. Safflower point cloud. (a) Segmentation point cloud. (b) Segmentation point cloud. (c) Point cloud after annotation. (d) Segmented safflower point cloud.

This study utilized a 64-bit Windows 10 operating system with 32 GB of RAM, an NVIDIA GeForce RTX 3080 graphic card (10 GB GDDR6 VRAM), and a 13th Gen Intel® Core™ i9-13900K processor (24 cores/32 threads, base frequency 3.0 GHz, max turbo frequency 5.8 GHz). The experiments were conducted using the PyTorch 1.13.0 deep learning framework, CUDA version 11.7, and Python version 3.8. The training batch size was set to 8, with an input of 1024 randomly sampled points, with the Adam optimizer employed, an initial learning rate of 0.001, and a weight decay parameter of 0.0001. The cross-entropy loss function was employed. The number of training iterations was 500.

2.4. PointSafNet Model

The PointNet model proposed by Charles R.Q. [41] directly processes point cloud data, primarily addressing inherent properties of point clouds: disorder, and invariance under transformations. However, PointNet's sole focus on individual point characterization and its limited capacity to integrate local contextual information results in significant information loss during classification and segmentation tasks. To address this limitation, the PointNet++ model [42] was developed to enhance local feature learning through hierarchical feature aggregation. By introducing Set Abstraction (SA) and Feature Propagation (FP) layers, PointNet++ achieved multi-scale feature extraction.

To effectively capture safflower point cloud features and enhance semantic segmentation accuracy, this paper proposes the PointSafNet architecture, which is based on PointNet++. The first two Set Abstraction (SA) layers utilize the Multi-Scale Grouping (MSG) strategy to obtain richer point cloud features, which are called SAMsg Layers. The attention mechanism is integrated into the second SAMsg Layer and the third SA layer, while refining the perception layer's architecture. The architectural refinement of the second layer for segmentation enhancement is visually annotated with red dashed lines in Figure 4. Specifically, two key improvements are implemented: (1) integrating Star Blocks from StarNet [43] and Contextual Anchor Attention (CAA) from PKINet [44] to enhance hierarchical feature learning capabilities; (2) replacing traditional convolutions in the Multi-Layer Perceptron (MLP) with depth-wise separable convolution (DSC) to enable comprehensive feature extraction from all points within each neighborhood and facilitate the generation of local features for current keypoints. During segmentation, PointSafNet first predicts semantic categories for each point in the raw point cloud based on ground truth labels. Through iterative sampling, feature combination, and hierarchical extraction, the model progressively acquires multi-scale features. High-dimensional features are then aligned with low-dimensional features using inverse distance interpolation and fused with local contextual information. Finally, the fused features are passed through a fully connected classification head to produce an $N \times M$ matrix (where N denotes the number of points and M the number of segmentation categories), achieving semantic segmentation of the point cloud.

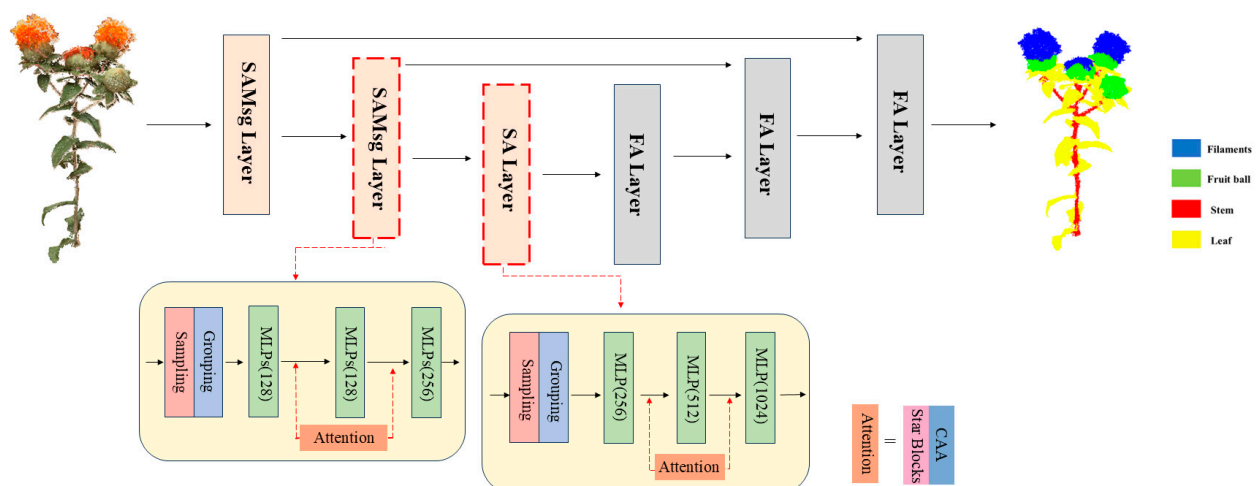


Figure 4. Schematic of PointSafNet model.

The first two SAMsg layers of the PointSafNet model employ a multi-scale grouping strategy for hierarchical feature learning. By selecting distinct radii to query the neighborhoods of keypoints, the model is able to capture fine-grained local features effectively. As shown in Figure 5, the model first uses the Farthest Point Sampling (FPS) method to

select key points based on the spatial distribution of the point cloud. Then, neighborhood selection is performed via ball query, that is, searching spherical regions centered around each key point. In contrast to the single-radius ball query method, multi-scale grouping is capable of effectively capturing local features across multiple spatial scales.

$$U(p_s, R) = \{p_i | p_s - p_i < R, i = 1, 2, 3, \dots, n\} \quad (1)$$

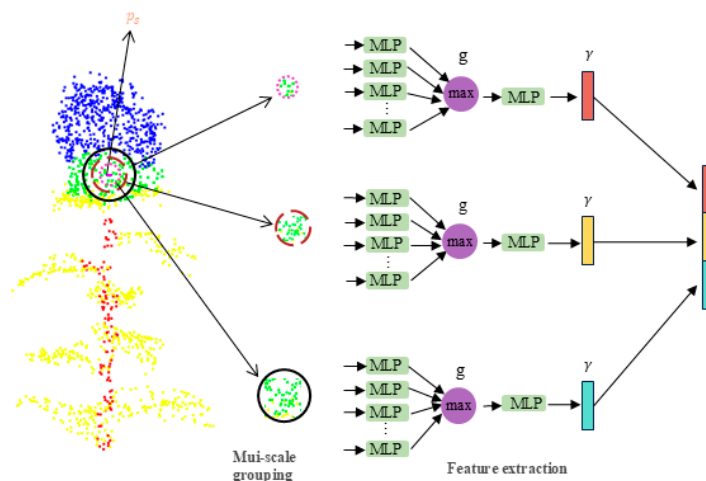


Figure 5. Multi-scale grouping to extract features.

2.4.1. Star Blocks Model

In the PointSafNet architecture, traditional feature extraction methods may fail to fully capture the complexity and diversity of point cloud data. To address this limitation, this study incorporates the Star Blocks model—depicted in Figure 6—to enhance feature representation capabilities. This operation significantly improves feature extraction accuracy without expanding network complexity.

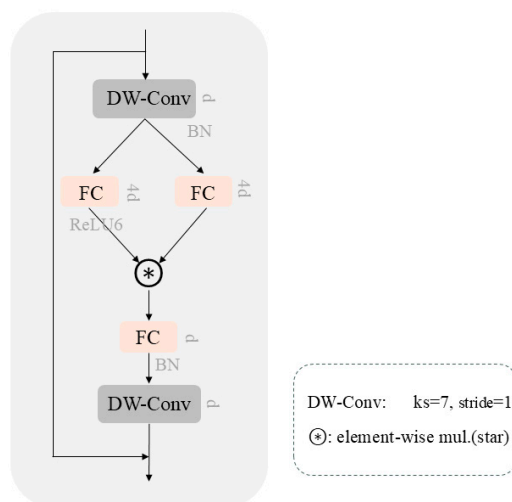


Figure 6. Star Blocks model.

The core operation of Star Blocks, namely the star operation, enhances feature discriminability and captures complex inter-feature relationships through element-wise multiplicative interactions, thereby generating more representative and discriminative feature embeddings. To further augment the network's feature representation capacity, depth-wise separable convolutional layers were inserted both before and after the Star Blocks layer. By setting the channel expansion factor to 4, the network width at each stage was effec-

tively increased. This modification not only improves feature extraction efficiency but also enhances overall network performance.

2.4.2. Contextual Anchor Attention (CAA) Model

The Contextual Anchor Attention (CAA) in the PointSafNet model enhances feature extraction by capturing long-range contextual dependencies in Figure 7. Specifically, the mechanism first extracts critical information from preceding stage features and generates local region features using average pooling and 1×1 convolutions. This enhances the model's ability to learn global point cloud semantics and provides a holistic perspective for later stages. After acquiring global contextual information, CAA computes inter-region correlations through attention mechanisms and adjusts regional representations based on correlation weights. This step focuses the model on discriminative feature regions while suppressing irrelevant background noise. Finally, contextual features are integrated with original features through fusion operations, thereby generating enhanced representative embeddings optimized for downstream point cloud segmentation tasks. As a result, CAA significantly improves model performance and accuracy.

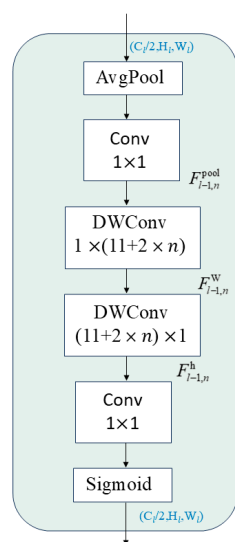


Figure 7. Contextual Anchor Attention model.

2.5. Harvesting Point Recognition Method

Based on the semantic segmentation of safflower plants, a method for identifying the picking points is proposed. The workflow for safflower picking point localization is presented in Figure 8. First, the 3D coordinates of the filament centroid $P_1(x_1, y_1, z_1)$ and fruit ball centroid $P_2(x_2, y_2, z_2)$ are computed independently. The spatial vector $\vec{P}$ between the two centroids is then calculated, and the Euclidean distance $|P|$ is determined. The unit direction vector $\hat{e}$ is subsequently obtained by normalizing $\vec{P}$ with respect to $|P|$. Using the Oriented Bounding Box (OBB) algorithm, the dimensions of the fruit ball's bounding box and edge lengths are determined. First, the convex hull algorithm is used to extract the outer contour of the safflower fruit ball point cloud. Then, PCA is employed to calculate the principal component directions (long, medium, and short axes) of the point cloud. Finally, taking the centroid as the origin and determining the boundaries of each axis of the bounding box along the principal component directions, an OBB that closely fits the point cloud is constructed. Half of the minimum edge length L_{min} is adopted as the spatial

displacement distance for the unit vector, enabling the final picking point coordinates $P_c(x_c, y_c, z_c)$ to be calculated accordingly.

$$x_1 = \frac{\sum_{i=1}^n x_i}{N}, x_2 = \frac{\sum_{i=1}^n x_i}{N} \quad (2)$$

$$y_1 = \frac{\sum_{i=1}^n y_i}{N}, y_2 = \frac{\sum_{i=1}^n y_i}{N} \quad (3)$$

$$z_1 = \frac{\sum_{i=1}^n z_i}{N}, z_2 = \frac{\sum_{i=1}^n z_i}{N} \quad (4)$$

where, in (x_1, y_1, z_1) , x_i, y_i and z_i represent the positional information of the segmented filament points. In (x_2, y_2, z_2) , x_i, y_i, z_i represent the positional information of the segmented fruit ball points. The index i is determined by the number N of segmented point clouds, respectively. n is the point index, and N is the total number of points.

$$\vec{P} = \overrightarrow{P_2 P_1} = P_1 - P_2 = (x_1 - x_2, y_1 - y_2, z_1 - z_2) \quad (5)$$

$$|P| = \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2 + (z_1 - z_2)^2} \quad (6)$$

$$\hat{e} = \frac{\vec{P}}{|P|} \quad (7)$$

$$P_c = P_2 + \frac{1}{2} L_{min} * \hat{e} \quad (8)$$

where $\vec{P}$ represents the spatial vector, $|P|$ denotes the spatial distance between the two centroids, $\hat{e}$ indicates the spatial unit vector, L_{min} represents the shortest edge of the bounding box, and P_c represents the picking point coordinates.

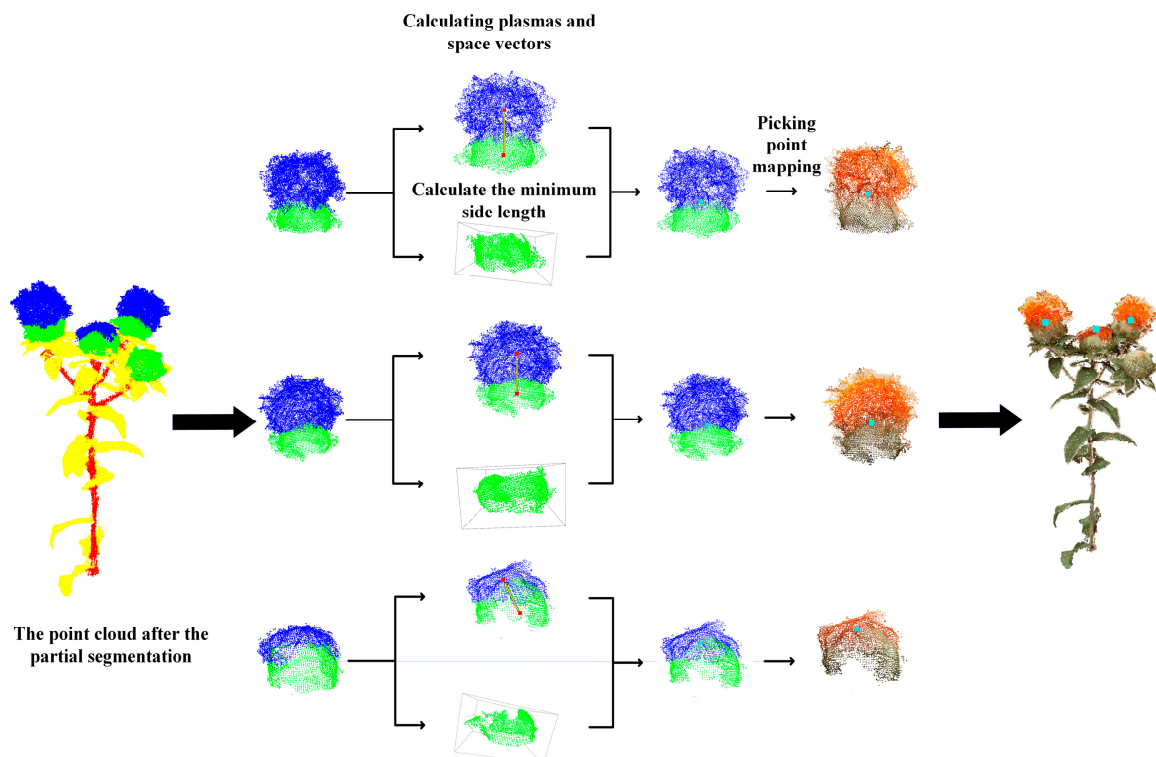


Figure 8. Calculation of picking points.

3. Results

3.1. Experimental Evaluation

The evaluation metrics used in this study include Accuracy, mean Intersection over Union for parts ($mIoU_{part}$), and mean Intersection over Union ($mIoU$). If a point in the safflower plant point cloud is labeled and classified into the correct category, it is considered a true positive (TP); if a point is incorrectly segmented, it is classified a false negative (FN); and if a point label does not exist but is segmented from the point cloud, it is considered a false positive (FP). The specific calculation formulas are as follows:

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad (9)$$

$$mIoU_{part} = \frac{1}{N_{part}} \sum_{i=1}^{N_{part}} IoU_i = \frac{1}{N_{part}} \sum_{i=1}^{N_{part}} \frac{TP_i}{TP_i + FN_i + FP_i} \quad (10)$$

$$mIoU = \frac{1}{N} \sum_{i=1}^N IoU_i = \frac{1}{N} \sum_{i=1}^N \frac{TP_i}{TP_i + FN_i + FP_i} \quad (11)$$

where i represents the categories of point cloud segmentation (filament, fruit ball, stem, and leaf).

3.2. Ablation Experiments

To evaluate the improved PointSafNet models, convolutional layers in different Set Abstraction (SA) layers were replaced with depth-wise separable convolution (DSC). Figure 9 illustrates the experimental configurations for different DSC replacement positions. As shown in Table 1, replacing the convolutional layer in the third SA layer with DSC improved overall network accuracy by 5.23% and $mIoU$ by 11.15%, whereas replacements at other positions did not yield significant performance gains. Ablation experiments were further conducted to investigate the effects of Star Blocks and the CAA attention model at different network locations, as presented in Figure 10. The results demonstrate that PointSafNet significantly enhances network accuracy. The original PointNet++ achieved 79.46% accuracy and 55.11% $mIoU$ in safflower point cloud segmentation. By integrating Star Blocks and the CAA model into the second and third SA layers, respectively, the network's accuracy increased to 86.83% and $mIoU$ improved to 68.82%, representing improvements of 7.37% and 13.71% over the baseline (Table 2).

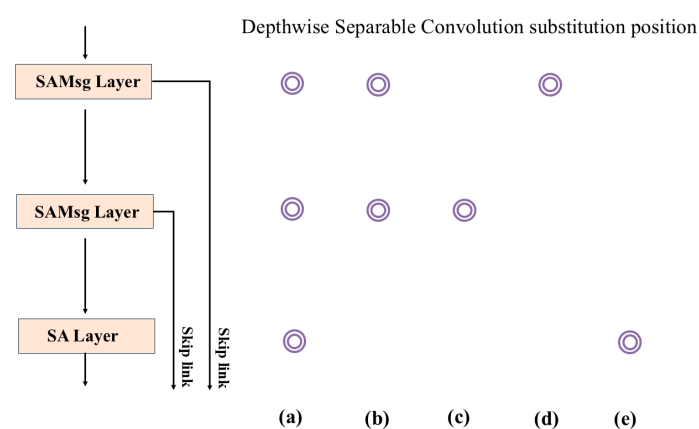
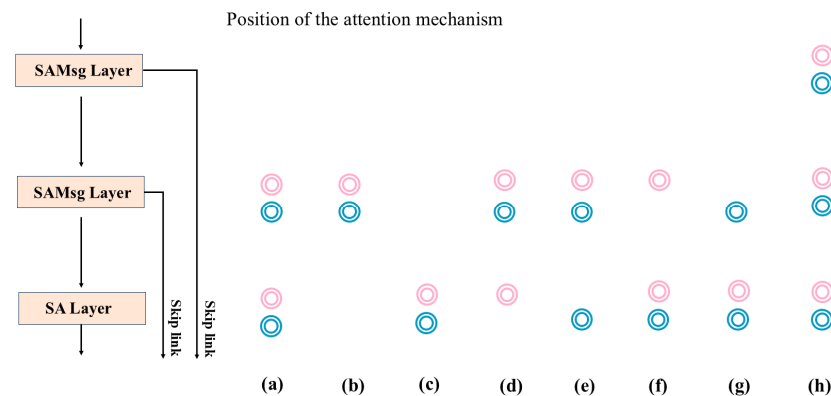


Figure 9. (a–e) Depthwise separable convolution substitution position. The purple circle is where we replaced conv with depth-wise separable convolution.

Table 1. Depth-wise separable convolution ablation experiments.

Model	Accuracy (%)	mIoU _{safflower filaments} (%)	mIoU _{fruit ball} (%)	mIoU _{leaf} (%)	mIoU _{stem} (%)	mIoU (%)
base	79.46	61.74	28.44	75.81	35.46	55.11
(a)	73.98	39.32	10.08	60.49	15.65	43.46
(b)	66.13	18.60	2.30	46.75	0.23	29.48
(c)	81.66	56.94	30.45	77.20	46.19	59.71
(d)	77.30	55.32	25.19	73.45	34.79	51.87
(e)	85.72	72.48	41.29	81.52	52.71	66.26

**Figure 10.** (a–h) Attention models are placed in different locations. Pink circles represent the Star Blocks models, while blue circles represent the CAA models.**Table 2.** Attention models' ablation experiment.

Model	Accuracy (%)	mIoU _{safflower filaments} (%)	mIoU _{fruit ball} (%)	mIoU _{leaf} (%)	mIoU _{stem} (%)	mIoU (%)
base	79.46	61.74	28.44	75.81	35.46	55.11
(a)	86.83	75.52	44.14	82.50	54.65	68.82
(b)	85.67	72.36	41.27	81.95	52.93	66.33
(c)	85.50	72.81	41.64	81.58	52.29	66.40
(d)	85.63	73.19	41.42	81.34	51.79	66.32
(e)	86.29	75.13	42.19	81.68	51.72	67.78
(f)	86.20	75.67	41.96	81.42	51.33	67.79
(g)	86.24	74.21	42.32	81.85	52.24	67.41
(h)	71.74	50.80	22.54	65.10	32.06	48.37

3.3. Performance Comparison of Different Models

To validate the effectiveness of the proposed PointSafNet model, it was compared against several widely used point cloud segmentation networks, including PointNet, PointNet++, and PointNet++_MSG [41,42]. The experimental results are shown in Table 3. The proposed PointSafNet model achieved mIoU_{safflower filaments}, mIoU_{fruit ball}, mIoU_{leaf}, mIoU_{stem}, and mIoU of 75.52%, 44.14%, 82.50%, 54.65%, and 68.82%, respectively. Compared to the PointNet model, these represent improvements of 24.75%, 33.39%, 11.85%, 54.53%, and 33.45% in mIoU_{safflower filaments}, mIoU_{fruit ball}, mIoU_{leaf}, mIoU_{stem}, and mIoU, respectively. Compared to the PointNet++ model, the mIoU_{safflower filaments}, mIoU_{fruit ball}, mIoU_{leaf}, mIoU_{stem}, and mIoU were improved by 13.78%, 15.70%, 6.69%, 19.19%, and 13.71%, respectively. Relative to PointNet++_MSG, the mIoU_{safflower filaments}, mIoU_{fruit ball}, mIoU_{leaf}, mIoU_{stem}, and mIoU were improved by 7.65%, 8.5%, 2.89%, 6.25%, and 7.34%, respectively. Moreover, the PointSafNet model outperforms other models in additional

performance metrics (Accuracy, $mIoU_{part}$, and $mIoU$). Figure 11 provides a visual comparison of the segmentation results for five sample safflower point clouds using PointNet, PointNet++, PointNet++_MSG, and PointSafNet.

Table 3. The performance of different algorithms for semantic segmentation.

Model	Accuracy (%)	$mIoU_{safflower\ filaments}$ (%)	$mIoU_{fruit\ ball}$ (%)	$mIoU_{leaf}$ (%)	$mIoU_{stem}$ (%)	$mIoU$ (%)
PointNet	71.69	50.77	10.75	70.65	0.12	35.37
PointNet++	79.46	61.74	28.44	75.81	35.46	55.11
PointNet++_MSG	82.28	67.87	35.64	79.61	48.40	61.48
PointSafNet	86.83	75.52	44.14	82.50	54.65	68.82

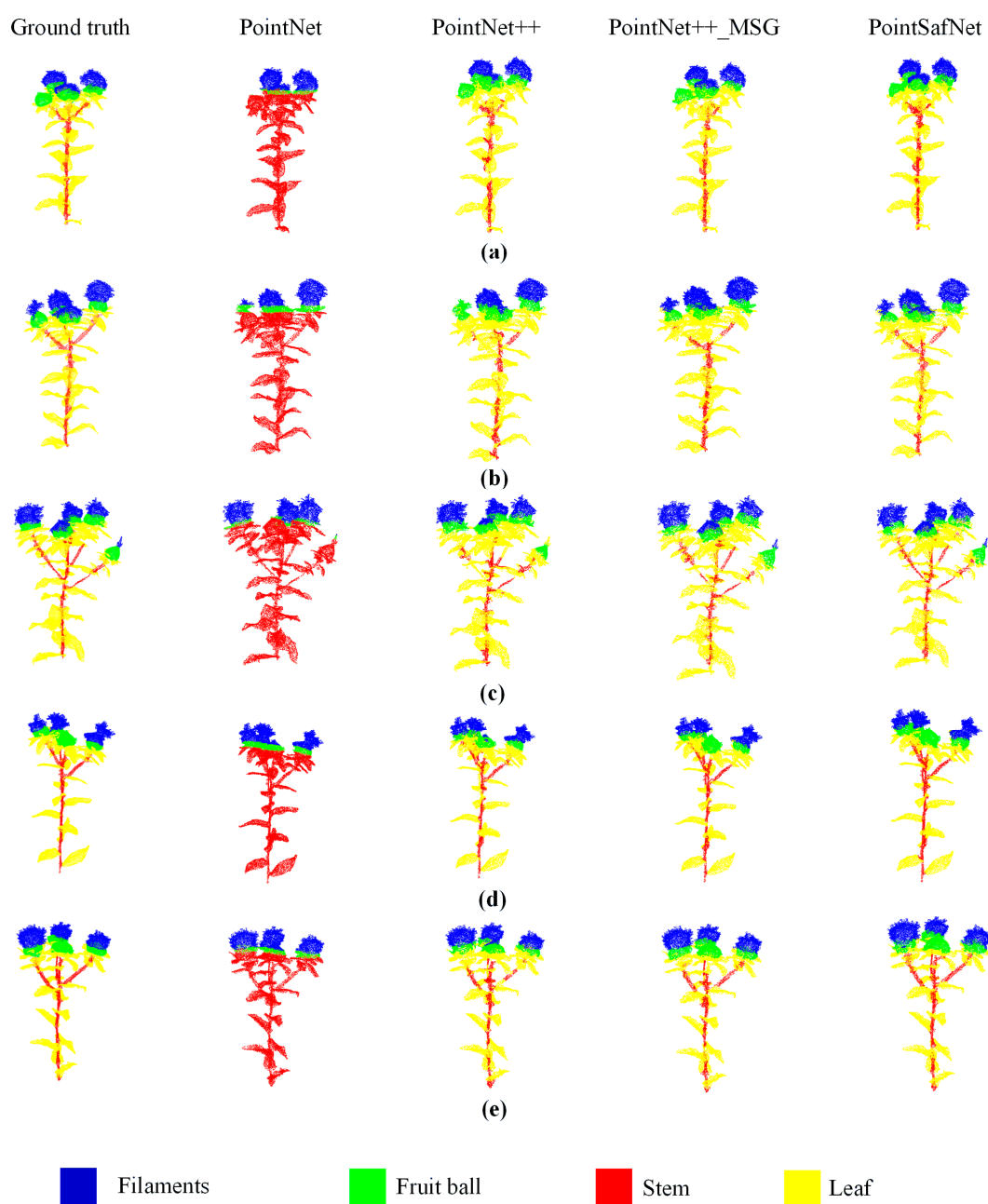


Figure 11. (a–e) Comparison of semantic segmentation results for different models.

In safflower point cloud segmentation, the PointNet model demonstrates suboptimal performance at filament–fruit ball boundaries and exhibits particularly inadequate handling

of stems due to their continuous, elongated geometry. This limitation prevents the effective capture of long-range spatial features, leading to the misclassification of nearly all stem points as leaves (Figure 11d–e). Notably, while PointNet++ achieves enhanced segmentation performance through hierarchical feature aggregation and enables precise classification of complex plant architectures, it struggles to preserve fine-grained structural details essential for high-precision phenotyping applications. To improve safflower segmentation accuracy, PointNet++_MSG incorporates a multi-scale feature fusion strategy, thereby enhancing the classification of detailed structures such as filament tips by capturing local information across multiple scales. However, as illustrated in Figure 11d, this method remains insensitive to subtle geometric differences at stem–leaf boundaries, resulting in incomplete segmentation in these regions. The proposed PointSafNet mitigates this limitation by employing an attention mechanism that dynamically emphasizes local feature weights in critical areas, such as filament–fruit ball transition zones, while simultaneously establishing long-range spatial relationships between stem structures and their surrounding context. The experimental results (Figure 11) reveal that PointSafNet not only significantly reduces stem misclassification rates but also accurately captures subtle variations in filament contours, demonstrating superior feature discriminability in complex deformation regions.

3.4. Picking Point Identification and Error Analysis

This study achieves adaptive picking point localization by integrating spatial features at the safflower filament–fruit ball interface with an OBB minimum distance strategy. The optimal picking range is defined as the spatial region at the filament–fruit ball junction. Tests on 110 safflower plants across different growth stages demonstrate that partially extended filaments in the early flowering stage (Figure 12a) form distinct geometric boundaries with fruit balls, enabling precise segmentation-based localization; fully open filaments and ellipsoidal fruit balls during the full-bloom stage (Figure 12b,c) allow for an accurate 3D coordinate calculation of the junction region using OBB; morphological changes in the post-full-bloom transition (Figure 12d) challenge OBB’s spatial boundary discrimination due to drooping filaments contacting fruit ball surfaces; and complete filament drooping in the decay stage (Figure 12e) causes segmentation misjudgments and localization failures. Through experimental validation, in the picking point identification test, a total of 110 ground-truth picking points were recorded, with 99 model-predicted picking points successfully identified, yielding an identification success rate of 90%. These results validate that the overall recognition accuracy of the proposed method remains consistently at 90%.

Furthermore, the segmentation quality of safflower filaments and fruit balls directly impacts 3D positioning stability. As shown in Figure 13, ineffective segmentation (Figure 13a) or incomplete segmentation (Figure 13b,c) of filaments result in positioning deviations; false positives where filaments are misclassified as fruit balls (Figure 13d,e) lead to abnormal OBB dimensions, preventing accurate determination of fruit ball bounding box edges and thereby causing positioning errors. While PointSafNet mitigates these issues through its attention mechanism, there remains room for improvement in segmenting unopened or completely decayed filament point clouds.

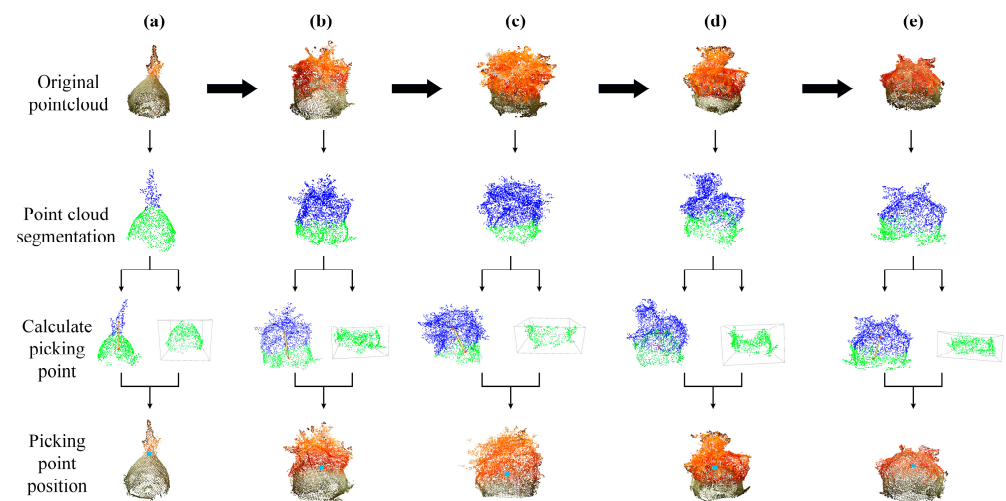


Figure 12. Determining picking points in different flowering states. (a) Filaments begin to open. (b,c) blooming safflower; (d) filaments gradually begin to decay; (e) decaying safflower.

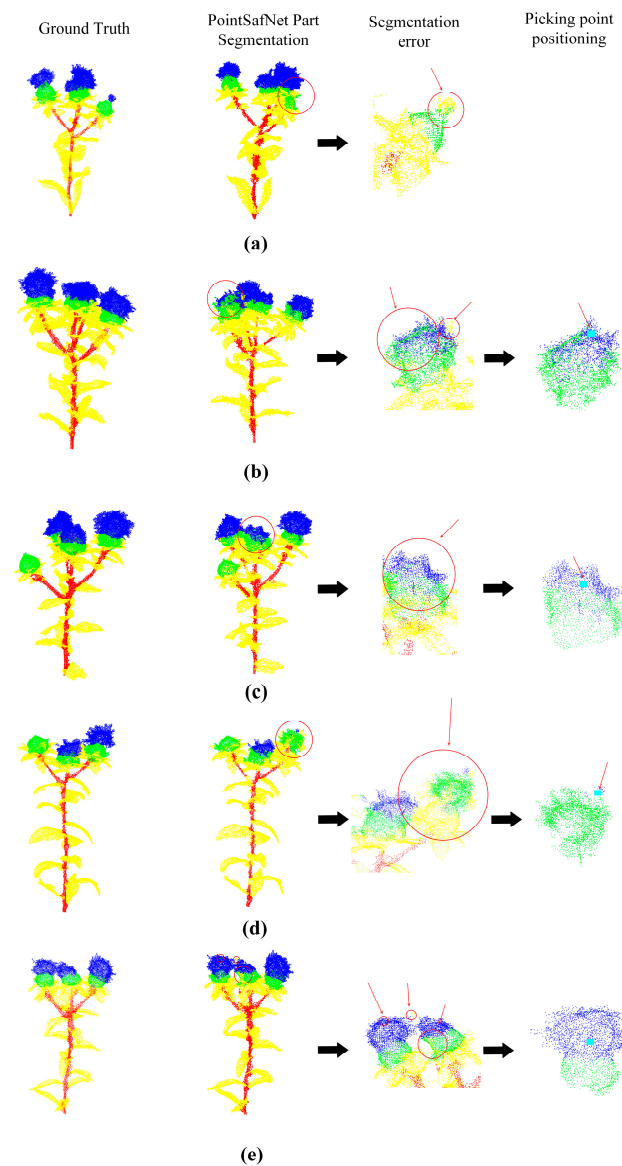


Figure 13. (a–e) Segmentation error and picking point positioning error schema.

4. Discussion

Proposed in this study, a safflower localization method based on 3D point clouds and the PointSafNet network achieves 90% localization accuracy for stamens under over-cast or low-intensity natural light conditions, verifying its effectiveness in multi-stage picking point calculation for single detached plant samples. However, when applied to large-scale complex field scenarios, the following technical bottlenecks and improvement directions remain:

(1) Light sensitivity and complex scene adaptability: Current 3D reconstruction relies on uniform low-light environments, which can effectively suppress specular reflection and shadow interference. However, strong midday light or dynamic lighting conditions may cause texture distortion in RGB images. For example, specular reflection from leaves weakens the feature extraction capability of the SfM-MVS algorithm (Structure from Motion–Multi-View Stereo algorithm) for the boundaries of safflower, thereby leading to reduced point cloud density and data loss. This phenomenon exposes the insufficiency of existing algorithms in physical environmental robustness. Future research could explore the use of shading measures under natural conditions to avoid the impact of lighting on reconstruction results.

(2) Modeling challenges in large-scale cultivation scenarios: In this experiment, a multi-view imaging strategy was adopted for single-plant samples to avoid common occlusion and motion blur issues in densely planted farmland. However, in practical scenarios, leaf and branch occlusion from adjacent plants may cause the key organs of the target safflower to be missing in multi-view images, thereby affecting the integrity of 3D reconstruction. Additionally, plant shaking caused by natural wind introduces registration errors between consecutive images, reducing point cloud alignment accuracy. To address these challenges, it is necessary to develop adaptive 3D reconstruction strategies suitable for large-scale cultivation environments, such as using three RGB-D cameras with fixed configuration for synchronous acquisition of safflower point clouds. This scheme can leverage multi-view depth information to mitigate occlusion issues and reduce registration errors caused by plant movement through synchronous data acquisition.

(3) Algorithmic computational efficiency and deployment feasibility. The current methods rely on high-computing-power platforms for point cloud segmentation and deep learning inference, creating a contradiction with the edge computing resources prevalent in agricultural fields. Although the PointSafNet network optimizes segmentation efficiency via feature aggregation modules, its GPU dependency and real-time performance still struggle to meet on-site field operation requirements. Future research could advance algorithm deployment on embedded devices through model lightweighting (e.g., knowledge distillation, quantization compression) and hardware co-design.

(4) The limited morphological generalizability of fruit ball segmentation: The significantly lower segmentation accuracy of fruit balls (mIoU = 44.14%) compared to filaments primarily stems from their morphological diversity: the geometric features of unexpanded fruit balls in early growth stages and wrinkled structures in senescence phase exhibit high similarity to filaments (Figure 13d,e), while the current training datasets inadequately cover atypical morphological samples. Furthermore, data bias issues further impair the model's ability to resolve continuous developmental stage variations. Proposed improvements include collecting fruit ball data across various growth stages, expanding the training set through 3D rotation and scale augmentation, and introducing 3D geometric constraint losses to guide global shape learning.

(5) Lack of multidimensional agronomic information fusion. Currently, the model only relies on RGB-geometric features for picking point localization, without integrating spectral, thermodynamic, and other biochemical information to assess flower maturity or

pathological status. For example, the absence of multispectral imaging capability results in an inability to calculate vegetation indices such as NDVI (Normalized Difference Vegetation Index), limiting the possibility of simultaneous “localization-quality inspection” integrated operations. Future work could explore multi-sensor fusion architectures to synchronously acquire hyperspectral, 3D point cloud, and thermal imaging data, and establish joint optimization models for picking point decision-making and quality evaluation.

5. Conclusions

To address challenges posed by irregular safflower growth patterns and the lack of spatial information in image-based safflower picking point recognition, this paper presents PointSafNet—a novel 3D point cloud-based method for safflower picking point recognition. By integrating Star Blocks with a Contextual Anchor Attention (CAA) mechanism, this approach significantly improves feature representation capabilities of safflower point clouds, thereby enhancing segmentation accuracy. Building upon this foundation, a new scheme for safflower picking point recognition is further proposed. Key experimental findings are summarized as follows:

(1) In this study, RGB image data of safflowers were acquired from multiple perspectives in natural environments. High-resolution 3D point clouds of safflower plants were reconstructed using the SfM-MVS algorithm, establishing a safflower point cloud dataset that provides a data foundation for subsequent point cloud analysis.

(2) The proposed workflow for safflower picking point recognition integrates three key steps: first, enhancing PointNet++ with Star Blocks and a Contextual Anchor Attention (CAA) mechanism to improve feature representation capabilities; second, segmenting filament and fruit ball point clouds to compute centroids, spatial vectors, and Oriented Bounding Box (OBB) parameters for fruit balls; finally, determining 3D picking coordinates using the minimum edge length of the fruit ball OBB. This feature enhancement strategy significantly improves point cloud segmentation accuracy in complex scenarios, providing a technical foundation for precise positioning.

(3) For the PointSafNet model, the mean intersection over union (mIoU) values for the segmentation of safflower point clouds, $mIoU_{\text{safflower filaments}}$, $mIoU_{\text{fruit ball}}$, $mIoU_{\text{leaf}}$, $mIoU_{\text{stem}}$, and the mIoU are 75.52%, 44.14%, 82.50%, 54.65%, and 68.82%, respectively. Compared with the original model, these indicators increased by 13.78%, 15.70%, 6.69%, 19.19%, and 13.71%, respectively. Moreover, the positioning accuracy of 90% was achieved in the positioning experiment.

The PointSafNet model proposed in this paper exhibits superior performance in safflower picking point localization. Accurate segmentation of safflower filaments and fruit balls is critical to the successful identification of picking points. Segmentation errors directly lead to localization failures or significant deviations. Future research will enhance complex scene robustness, computational efficiency, and segmentation generalizability through the following approaches, and achieve integrated “localization-quality inspection” operations: fusing light shielding measures, designing dynamic light adaptation and occlusion repair algorithms, advancing model lightweighting and cross-modal feature fusion, expanding full-cycle sample data, and constructing multi-sensor joint optimization architectures.

Author Contributions: Conceptualization, B.M., H.X. and D.W.; methodology, B.M. and H.X.; software, B.M. and H.Z.; validation, B.M. and M.L.; formal analysis, B.M., Y.G. and Z.W. investigation, B.M. and H.X.; resources, B.M., H.Z. and Z.W.; data curation, B.M., H.X. and Z.W.; writing—original draft preparation, B.M., H.X. and D.W.; writing—review and editing, B.M., H.X., Y.G. and D.W.; visualization, B.M., H.X. and M.L.; supervision, B.M. and H.X. All authors have read and agreed to the published version of the manuscript.

Funding: This study was supported by the National Natural Science Foundation of China (grant No. 52065057).

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

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

References

- Chen, B.; Ding, F.; Ma, B.; Wang, L.; Ning, S. A Method for Real-Time Recognition of Safflower Filaments in Unstructured Environments Using the YOLO-SaFi Model. *Sensors* **2024**, *24*, 4410. [\[CrossRef\]](#)
- Guo, H.; Luo, D.; Gao, G.; Wu, T.; Diao, H. Design and experiment of a safflower picking robot based on a parallel manipulator. *Eng. Agric.* **2022**, *42*, e20210129. [\[CrossRef\]](#)
- Zhang, Z.; Shi, R.; Xing, Z.; Guo, Q.; Zeng, C. Improved Faster Region-Based Convolutional Neural Networks (R-CNN) Model Based on Split Attention for the Detection of Safflower Filaments in Natural Environments. *Agronomy* **2023**, *13*, 2596. [\[CrossRef\]](#)
- Zhang, Z.; Xing, Z.; Yang, S.; Feng, N.; Liang, R.; Zhao, M. Design and experiments of the circular arc progressive type harvester for the safflower filaments. *Trans. Chin. Soc. Agric. Eng.* **2022**, *38*, 10–21. [\[CrossRef\]](#)
- Zhang, Z.; Zhao, M.; Xing, Z.; Liu, X. Design and test of double-acting opposite direction cutting end effector for safflower harvester. *Trans. Chin. Soc. Agric. Mach.* **2022**, *53*, 160–170. [\[CrossRef\]](#)
- Sun, C.; Ge, Y.; Zhang, H.; Zeng, H.; Zhang, L. Design and experiment of the vertical brush-roller picking device for dry-safflower harvesters. *Trans. Chin. Soc. Agric. Eng.* **2024**, *40*, 203. [\[CrossRef\]](#)
- Zhang, Z.; Xing, Z.; Zhao, M.; Yang, S.; Guo, Q.; Shi, R.; Zeng, C. Detecting safflower filaments using an improved YOLOv3 under complex environments. *Trans. Chin. Soc. Agric. Eng.* **2023**, *39*, 162–170. [\[CrossRef\]](#)
- Wang, X.; Xu, Y.; Zhou, J.; Chen, J. Safflower picking recognition in complex environments based on an improved YOLOv7. *Trans. Chin. Soc. Agric. Eng.* **2023**, *39*, 169–176. [\[CrossRef\]](#)
- Zhao, W.; Wu, D.; Zheng, X. Detection of Chrysanthemums Inflorescence Based on Improved CR-YOLOv5s Algorithm. *Sensors* **2023**, *23*, 4234. [\[CrossRef\]](#)
- Park, H.-M.; Park, J.-H. YOLO Network with a Circular Bounding Box to Classify the Flowering Degree of Chrysanthemum. *AgriEngineering* **2023**, *5*, 1530–1543. [\[CrossRef\]](#)
- Zhou, H.; Luo, J.; Ye, Q.; Leng, W.; Qin, J.; Lin, J.; Xie, X.; Sun, Y.; Huang, S.; Pang, J. Advancing jasmine tea production: YOLOv7-based real-time jasmine flower detection. *J. Sci. Food Agric.* **2024**, *104*, 9297–9311. [\[CrossRef\]](#) [\[PubMed\]](#)
- Shinoda, R.; Motoki, K.; Hara, K.; Kataoka, H.; Nakano, R.; Nakazaki, T.; Noguchi, R. RoseTracker: A system for automated rose growth monitoring. *Smart Agric. Technol.* **2023**, *5*, 100271. [\[CrossRef\]](#)
- Fatehi, F.; Bagherpour, H.; Amiri Parian, J. Enhancing the Performance of YOLOv9t Through a Knowledge Distillation Approach for Real-Time Detection of Bloomed Damask Roses in the Field. *Smart Agric. Technol.* **2025**, *10*, 100794. [\[CrossRef\]](#)
- Gupta, S.; Tripathi, A.K. Flora-NET: Integrating dual coordinate attention with adaptive kernel based convolution network for medicinal flower identification. *Comput. Electron. Agric.* **2025**, *230*, 109834. [\[CrossRef\]](#)
- Yang, Q.; Chang, C.; Bao, G.; Fan, J.; Xun, Y. Recognition and localization system of the robot for harvesting Hangzhou White Chrysanthemums. *Int. J. Agric. Biol. Eng.* **2018**, *11*, 88–95. [\[CrossRef\]](#)
- Yang, Q.; Luo, S.; Chang, C.; Xun, Y.; Bao, G. Segmentation algorithm for Hangzhou white chrysanthemums based on least squares support vector machine. *Int. J. Agric. Biol. Eng.* **2019**, *12*, 127–134. [\[CrossRef\]](#)
- Xing, Z.; Zhang, Z.; Shi, R.; Guo, Q.; Zeng, C. Filament-necking localization method via combining improved PSO with rotated rectangle algorithm for safflower-picking robots. *Comput. Electron. Agric.* **2023**, *215*, 108464. [\[CrossRef\]](#)
- Rong, Q.; Hu, C.; Hu, X.; Xu, M. Picking point recognition for ripe tomatoes using semantic segmentation and morphological processing. *Comput. Electron. Agric.* **2023**, *210*, 107923. [\[CrossRef\]](#)
- Wang, X.; Zhou, J.; Xu, Y.; Cui, C.; Liu, Z.; Chen, J. Location of safflower filaments picking points in complex environment based on improved Yolov5 algorithm. *Comput. Electron. Agric.* **2024**, *227*, 109463. [\[CrossRef\]](#)
- Zhang, H.; Ge, Y.; Xia, H.; Sun, C. Safflower picking points localization method during the full harvest period based on SBP-YOLOv8s-seg network. *Comput. Electron. Agric.* **2024**, *227*, 109646. [\[CrossRef\]](#)
- Chen, C.; Lu, J.; Zhou, M.; Yi, J.; Liao, M.; Gao, Z. A YOLOv3-based computer vision system for identification of tea buds and the picking point. *Comput. Electron. Agric.* **2022**, *198*, 107116. [\[CrossRef\]](#)
- Ge, Y.; Xiong, Y.; Tenorio, G.L.; From, P.J. Fruit Localization and Environment Perception for Strawberry Harvesting Robots. *IEEE Access* **2019**, *7*, 147642–147652. [\[CrossRef\]](#)
- Li, Y.; He, L.; Jia, J.; Lv, J.; Chen, J.; Qiao, X.; Wu, C. In-field tea shoot detection and 3D localization using an RGB-D camera. *Comput. Electron. Agric.* **2021**, *185*, 106149. [\[CrossRef\]](#)

24. Li, Y.; Wu, S.; He, L.; Tong, J.; Zhao, R.; Jia, J.; Chen, J.; Wu, C. Development and field evaluation of a robotic harvesting system for plucking high-quality tea. *Comput. Electron. Agric.* **2023**, *206*, 107659. [\[CrossRef\]](#)
25. Zhu, L.; Zhang, Z.; Lin, G.; Chen, P.; Li, X.; Zhang, S. Detection and Localization of Tea Bud Based on Improved YOLOv5s and 3D Point Cloud Processing. *Agronomy* **2023**, *13*, 2412. [\[CrossRef\]](#)
26. Yoshida, T.; Fukao, T.; Hasegawa, T. A Tomato Recognition Method for Harvesting with Robots using Point Clouds. In Proceedings of the 2019 IEEE/SICE International Symposium on System Integration (SII), Paris, France, 14–16 January 2019; pp. 456–461.
27. Díaz, C.A.; Pérez, D.S.; Miatello, H.; Bromberg, F. Grapevine buds detection and localization in 3D space based on Structure from Motion and 2D image classification. *Comput. Ind.* **2018**, *99*, 303–312. [\[CrossRef\]](#)
28. Qiao, Y.; Liao, Q.; Zhang, M.; Han, B.; Peng, C.; Huang, Z.; Wang, S.; Zhou, G.; Xu, S. Point clouds segmentation of rapeseed siliques based on sparse-dense point clouds mapping. *Front Plant Sci* **2023**, *14*, 1188286. [\[CrossRef\]](#)
29. Luo, J.; Zhang, D.; Luo, L.; Yi, T. PointResNet: A grape bunches point cloud semantic segmentation model based on feature enhancement and improved PointNet++. *Comput. Electron. Agric.* **2024**, *224*, 109132. [\[CrossRef\]](#)
30. Zarei, A.; Li, B.; Schnable, J.C.; Lyons, E.; Pauli, D.; Barnard, K.; Benes, B. PlantSegNet: 3D point cloud instance segmentation of nearby plant organs with identical semantics. *Comput. Electron. Agric.* **2024**, *221*, 108922. [\[CrossRef\]](#)
31. Saeed, F.; Sun, S.; Rodriguez-Sanchez, J.; Snider, J.; Liu, T.; Li, C. Cotton plant part 3D segmentation and architectural trait extraction using point voxel convolutional neural networks. *Plant Methods* **2023**, *19*, 33. [\[CrossRef\]](#)
32. Shen, J.; Wu, T.; Zhao, J.; Wu, Z.; Huang, Y.; Gao, P.; Zhang, L. Organ Segmentation and Phenotypic Trait Extraction of Cotton Seedling Point Clouds Based on a 3D Lightweight Network. *Agronomy* **2024**, *14*, 1083. [\[CrossRef\]](#)
33. Yan, J.; Tan, F.; Li, C.; Jin, S.; Zhang, C.; Gao, P.; Xu, W. Stem–Leaf segmentation and phenotypic trait extraction of individual plant using a precise and efficient point cloud segmentation network. *Comput. Electron. Agric.* **2024**, *220*, 108839. [\[CrossRef\]](#)
34. Bayati, H.; Najafi, A.; Vahidi, J.; Gholamali Jalali, S. 3D reconstruction of uneven-aged forest in single tree scale using digital camera and SfM-MVS technique. *Scand. J. For. Res.* **2021**, *36*, 210–220. [\[CrossRef\]](#)
35. Lowe, D.G. Object recognition from local scale-invariant features. In Proceedings of the Seventh IEEE International Conference on Computer Vision, Corfu, Greece, 20–27 September 1999; Volume 1152, pp. 1150–1157.
36. Seitz, S.M.; Curless, B.; Diebel, J.; Scharstein, D.; Szeliski, R. A Comparison and Evaluation of Multi-View Stereo Reconstruction Algorithms. In Proceedings of the 2006 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'06), New York, NY, USA, 17–22 June 2006; pp. 519–528.
37. Jarahizadeh, S.; Salehi, B. A Comparative Analysis of UAV Photogrammetric Software Performance for Forest 3D Modeling: A Case Study Using AgiSoft Photoscan, PIX4DMapper, and DJI Terra. *Sensors* **2024**, *24*, 286. [\[CrossRef\]](#) [\[PubMed\]](#)
38. Liu, J.-P.; Wu, M.-H.; Tsang, P.W.M. 3D display by binary computer-generated holograms with localized random down-sampling and adaptive intensity accumulation. *Opt. Express* **2020**, *28*, 24526–24537. [\[CrossRef\]](#)
39. Tsirikolias, K. Low level image processing and analysis using radius filters. *Digit. Signal Process.* **2016**, *50*, 72–83. [\[CrossRef\]](#)
40. Chang, A.X.; Funkhouser, T.; Guibas, L.; Hanrahan, P.; Huang, Q.; Li, Z.; Savarese, S.; Savva, M.; Song, S.; Su, H.; et al. ShapeNet: An Information-Rich 3D Model Repository. *arXiv* **2015**, arXiv:1512.03012. [\[CrossRef\]](#)
41. Qi, C.R.; Su, H.; Mo, K.; Guibas, L.J. PointNet: Deep Learning on Point Sets for 3D Classification and Segmentation. *arXiv* **2016**, arXiv:1612.00593. [\[CrossRef\]](#)
42. Qi, C.; Yi, L.; Su, H.; Guibas, L.J. PointNet++: Deep Hierarchical Feature Learning on Point Sets in a Metric Space. In Proceedings of the Conference on Neural Information Processing Systems, Long Beach, CA, USA, 4–9 December 2017.
43. Ma, X.; Dai, X.; Bai, Y.; Wang, Y.; Fu, Y. Rewrite the Stars. *arXiv* **2024**, arXiv:2403.19967. [\[CrossRef\]](#)
44. Cai, X.; Lai, Q.; Wang, Y.; Wang, W.; Sun, Z.; Yao, Y. Poly Kernel Inception Network for Remote Sensing Detection. In Proceedings of the 2024 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), Seattle, WA, USA, 16–22 June 2024; pp. 27706–27716.

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.