Computers and Electronics in Agriculture 212 (2023) 108116

Contents lists available at ScienceDirect

Computers and Electronics in Agriculture

journal homepage: www.elsevier.com/locate/compag

Human-centered approach for an efficient cucumber harvesting robot system: Harvest ordering, visual servoing, and end-effector

Yonghyun Park [a,b], Jaehwi Seol [a,b], Jeonghyeon Pak [a,b], Yuseung Jo [a,b], Changjo Kim [a], Hyoung Il Son [a,b,c,*]

[a] Department of Convergence Biosystems Engineering, Chonnam National University, Yongbong-ro 77, Gwangju 61186, Republic of Korea
[b] Interdisciplinary Program in IT-Bio Convergence System, Chonnam National University, Yongbong-ro 77, Gwangju 61186, Republic of Korea
[c] Research Center for Biological Cybernetics, Chonnam National University, Yongbong-ro 77, Gwangju 61186, Republic of Korea

ARTICLE INFO

Keywords:
Cucumber harvesting robot
End-effector
Harvest ordering
Harvesting robot system
Visual servoing

ABSTRACT

This paper presents a human-centered approach for an efficient cucumber-harvesting robot system. Specifically, harvest ordering, visual servoing, and end-effector-based manipulation functionalities were integrated to realize efficient and stable harvesting. The proposed approach involved determining the optimal harvest ordering, guiding the end-effector to the cucumber pedicel through visual servoing, and designing an end-effector to effectively harvest long cucumbers. The performance of the system was evaluated through preliminary and field experiments. The results of the preliminary experiments showed that harvest ordering decreased the harvesting time and travel length and increased the battery efficiency. The visual servoing was robust, and pedicels could be rapidly detected at a speed of 16–23 FPS through computer vision technologies. The pedicel could be accurately positioned within the cutting area of the end-effector. Furthermore, the proposed end-effector could effectively cut thin cucumber pedicels (3–6 mm), with a 100.0% success rate. Field experiments were conducted at three sites in Korea: Green Monsters, Sangju smart-valley, and Fresh-farm. The harvest success rate at the three sites ranged from 50.9% to 60.0%, with an overall value of 56.6%. The overall average harvest time of 56.0 s. The positional accuracy of the system was within the optimal range of 0–30 mm. Furthermore, the primary causes of harvest failure were analyzed, and future research directions to improve the performance of harvest robots were discussed.

1. Introduction

Cucumbers are among the most commonly consumed vegetables worldwide. Cucumbers have secured a pivotal position in the global vegetable consumption landscape due to their fast growth and maturation cycles that are highly adaptable to various climates, from temperate to tropical (Mao et al., 2020a). To provide a steady supply of this essential vegetable throughout the year, cucumber smart farms have been established worldwide. Despite this advancement, the agricultural sector encounters multiple challenges. One notable issue is the shrinking labor force attributable to the aging population, with this problem being particularly prominent in rural communities. Additionally, high operational costs, such as those of energy, water, agricultural chemicals, and labor, threaten the sustainability of these farms. Furthermore, the labor-intensive nature of farming practices aggravates the already complex issue of agricultural labor scarcity. To address these problems, especially that of the large workforce required in harvesting processes, the development of harvesting robots has garnered considerable interest.

Several robots have been developed to harvest fruits and vegetables such as apples, peppers, tomatoes, kiwis, and strawberries (Li et al., 2019; Yu et al., 2019; Song et al., 2021). Harvest robot systems typically involve autonomous robots that can continuously perform harvesting tasks such as fruit detection, reaching, cutting and catching, and loading (Xiong et al., 2020). For example, an existing apple harvest robot incorporates a manipulator, vacuum-based end-effector, and mobile vehicle (Hu et al., 2022). Field tests have demonstrated that the rotational pull pattern is more effective in picking apples than a simple pull pattern. Similarly, an existing red pepper harvesting robot performs fruit detection, location, holding posture estimation, and motion control (Arad et al., 2020). Experimental findings show

* Correspondence to: Yongbong-ro 77, Gwangju 61186, Republic of Korea.
E-mail addresses: dk03378@jnu.ac.kr (Y. Park), seol0810@jnu.ac.kr (J. Seol), poooodg@jnu.ac.kr (J. Pak), chossbb68@jnu.ac.kr (Y. Jo), ckddnckdwh12@jnu.ac.kr (C. Kim), hison@jnu.ac.kr (H.I. Son).
URL: https://www.hralab.com (H.I. Son).

https://doi.org/10.1016/j.compag.2023.108116
Received 1 March 2023; Received in revised form 1 July 2023; Accepted 25 July 2023
Available online 10 August 2023

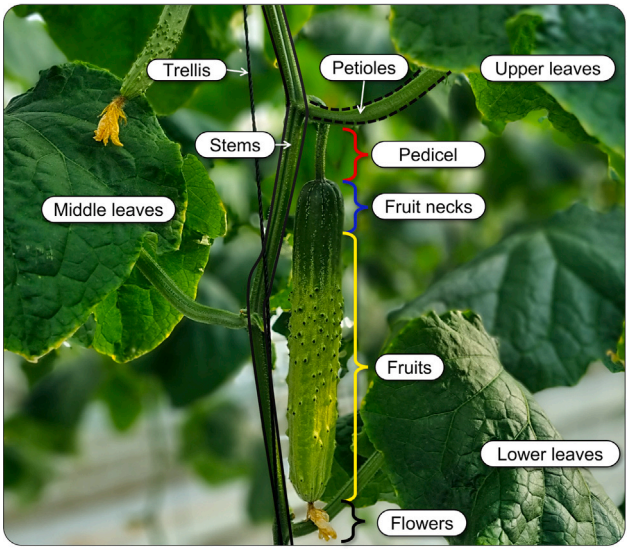


Fig. 1. Structure of a cucumber plant.

that leaves and neighboring fruits can affect the detection accuracy and interfere with the movement of the end-effector. The robot performance can potentially be enhanced by addressing these problems. The success rate and efficiency of tomato harvesting robots have been enhanced by accurately locating and estimating the optimal position for holding the fruit (Rong et al., 2022). In general, the tomato positions must be continuously tracked because the end-effector tends to shake during harvest. Kiwi harvesting multi-arm robots have been noted to successfully harvest 51.0% of all kiwis in an orchard (Williams et al., 2019). By reducing the drop rate and knock-off rate by 6.7%, the success rate can be further improved. Efforts are being made to minimize the kiwi harvest time by distributing uniform work among multiple arms (Barnett et al., 2020). A strawberry harvest robot has been developed to incorporate an autonomous robot that can continuously pick strawberries in multiple tunnels (Xiong et al., 2019). Overall, harvesting robots are being integrated into various agricultural systems to perform tasks continuously and autonomously. However, it is necessary to consider a system holistically rather than focusing on individual elements. Additionally, field experiments have reported improvements by exploring failure cases related to problems such as inaccurate fruit positioning, incorrect route planning, and fruit damage caused by the end-effector. Thus, research on cucumber harvest robots must take into account these challenges in the design phase.

Despite these considerations, most of the existing research on cucumber-harvesting robots has been focused on detecting cucumbers (Mao et al., 2020b; Bai et al., 2022; Kim et al., 2023). This is because the surrounding stems and berries are similar in color to the cucumber, which renders its detection challenging.(Fig. 1). Some attempts have been made to develop integrated autonomous harvesting robots (Van Henten et al., 2002, 2003, 2009). In these studies, field tests were conducted in highly variable environmental conditions, failure causes were analyzed through discussion, and suggestions for improvements were provided. However, research targeting the autonomous harvesting of cucumbers is limited, and none of the existing systems have been successfully commercialized. While the individual systems has been extensively studied, it is necessary to conduct empirical research that integrates the overall system and addresses problems in irregular field conditions. Therefore, research on cucumber harvesting must be guided by integrated and autonomous methodologies.

As shown in Fig. 2(a), previous laboratory studies have developed, combining the principles of three-dimensional (3D) perception, manipulation, and end-effector-based manipulation (Jun et al., 2021). However, problems related to fruit localization, path planning, shaking, and fruit damage while cutting are encountered in atypical environments, which necessitate human intervention To resolve these potential problems, it is necessary to design robots that can ultimately perform tasks autonomously by emulating human-like processes. As shown in Fig. 2(b), recent laboratory studies have designed and conducted preliminary experiments with a human-centered robot system for autonomous cucumber harvesting (Park et al., 2022b). This approach was aimed at emulating human harvesting methods by integrating harvest ordering, visual servoing, and end-effector-based manipulation. The system was noted to be effective in decreasing the travel length and operating time, and the end-effector could realize stable grasping and cutting. This framework highlights the necessity of effectively linking the sensory information acquired during the harvesting process with robot motion control, mirroring the approach adopted by humans. As shown in Fig. 3, the human approach for harvesting cucumbers involves three main stages :

1. Harvest ordering: Humans visually detect the cucumbers to be harvested and determine the optimal harvesting order. This process involves identifying cucumbers and considering factors such as the position of the cucumber in the plant and the overall layout of the field.
2. Approaching: Upon deciding the harvest order, the human harvester approaches the cucumber and begins hand manipulations. The harvester carefully positions their hand and fingers to securely grasp the cucumber, while being careful to avoid any damage.
3. Grasping and cutting: After securely holding the cucumber, the harvester uses their other hand to cut the pedicel using pruning shears. This process requires precise cutting to cleanly sever the pedicel without damaging the cucumber.

Considering these aspects, this paper proposes a human-centered approach for an efficient cucumber-harvesting robot system. The proposed system integrates harvest ordering, visual servoing, and an end-effector. The objectives of this study are as follows: (a) determining the harvest order that allows the robot to perform harvesting along the optimal path, (b) guiding the end-effector to the pedicel by visually servoing the robot in real-time, (c) designing an end-effector that can efficiently harvest long cucumbers, and (d) conducting field tests to assess the performance of the proposed cucumber-harvesting robot (Fig. 4). It is anticipated that the proposed cucumber-harvesting robot system can effectively address the unique challenges of cucumber harvesting to realize efficient and stable harvesting.

1.1. Related work

1.1.1. Harvest ordering

Most of the existing research on harvest ordering has been focused on enhancing perception accuracy, given the complexities associated with fruit detection in unstructured and dynamic operational environments. In other words, the literature on harvesting robots involves only a few reports on harvest planning and ordering. Notably, the robot operation may be inefficient in the absence of a path plan in the dense and unstructured environments typically encountered in agricultural applications. Thus, the optimized harvest order must be derived based on detected information, similar to the sequence planning problem commonly encountered in mobile robot applications (Pak et al., 2022).

In the initial studies on harvest ordering, the minimum path was determined based on traveling salesman problems (TSP) to plan the harvesting sequence (Edan et al., 1991). However, there remained considerable scope for further optimization. Recently, research attention has shifted toward task scheduling and planning (Kurtser and Edan, 2020). For example, given that harvest sequence planning relies on perception, three sensing methods were defined and simulated. The travel

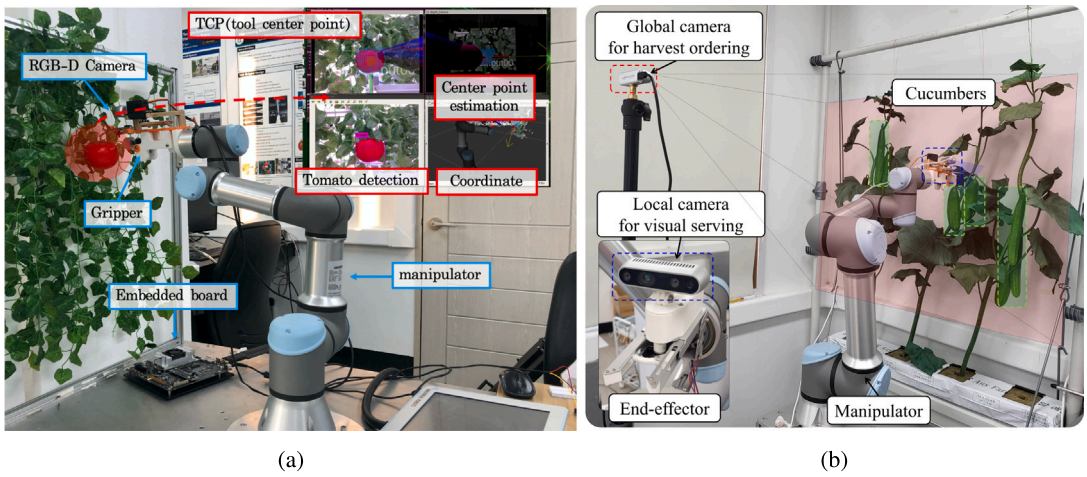


Fig. 2. Previous harvesting robot systems: (a) tomato Jun et al. (2021), (b) cucumber (Park et al., 2022b).

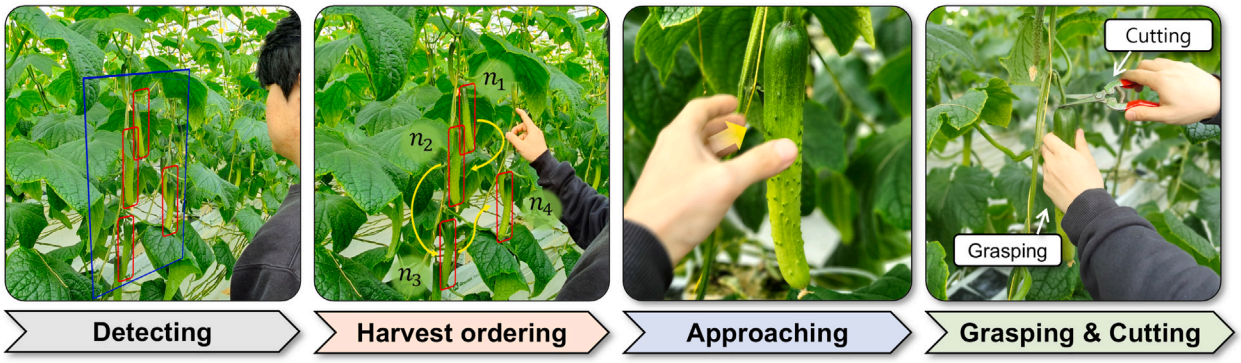


Fig. 3. Human approach for harvesting cucumbers.

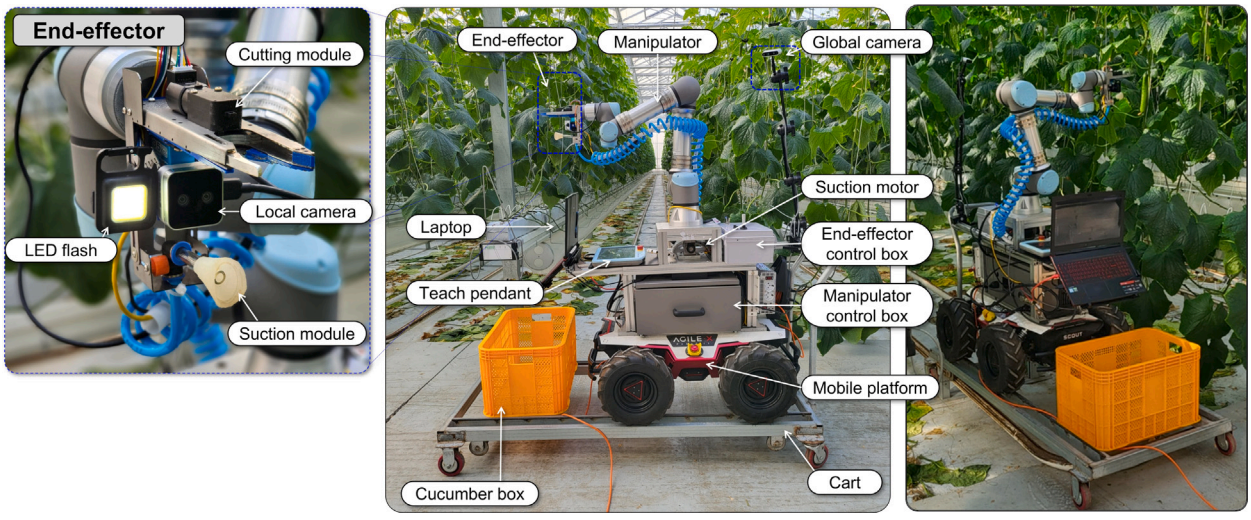


Fig. 4. Proposed cucumber-harvesting robot system.

cost was calculated based on the distance between the detected crops, thus formulating a TSP problem with the objective of minimizing the cost. Additionally, the harvesting cycle time was reduced by devising motion plans aimed at minimizing the travel distance

Ning et al. (2022) developed a method for planning an anti-collision picking sequence after accurate identification in a dense and complex environment. A heuristic, experience-dependent method was chosen to classify picking clusters and plan intra-group picking sequences. This method resulted in enhanced harvesting efficiency because collision-free path planning reduced damage and prevented shaking. However, the cycle time was not reduced.

1.1.2. Visual servoing

In recent years, visual servoing has been increasingly used in agricultural applications. As a robotic control system, visual servoing frameworks use visual feedback from a camera to guide the movements

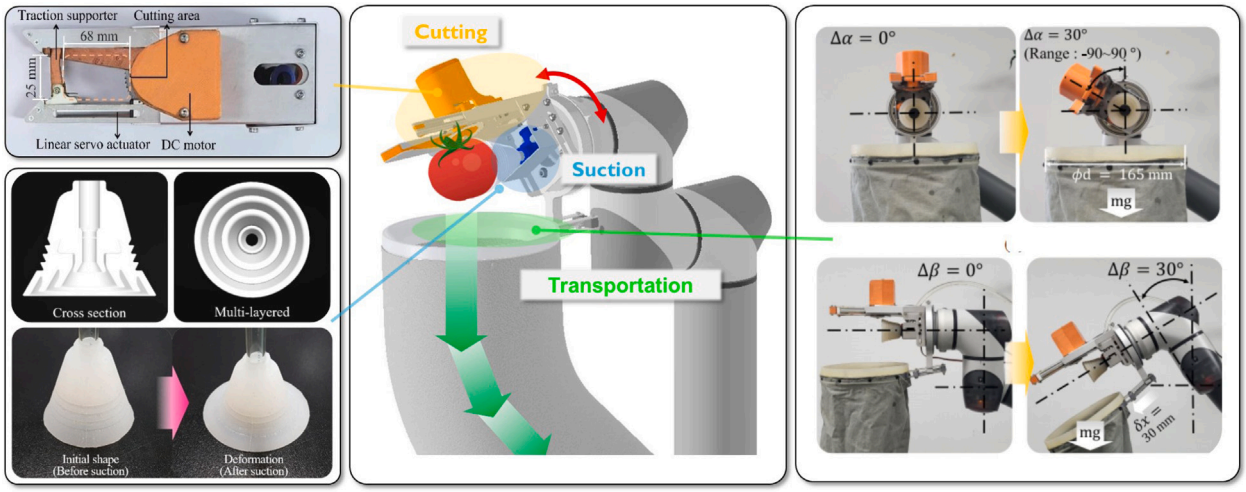


Fig. 5. End-effector for a fruit and vegetable harvesting robot (Park et al., 2022a).

of the end-effector, such as a gripper or hand (Dewi et al., 2018). This capability allows the robot to accurately track and approach harvest targets, such as fruits and vegetables (Mehta et al., 2016). However, the highly dynamic and unstructured nature of the operating environment poses a significant challenge for the use of visual servoing in the agricultural domain (Bai et al., 2023). Fields often feature uneven terrains, and the position and orientation of individual plants can drastically vary, which makes it difficult for the robot to accurately track and approach targets. These aspects often result in decreased performance and higher failure rates.

Various researchers have endeavored to address these challenges. For example, Lee et al. (2012) developed algorithms for real-time object detection and tracking, which enable the robot to accurately pinpoint and track its target. Other researchers have deployed machine learning techniques to enhance the adaptability of a robot to environmental changes and unexpected situations. Mehta and Burks (2014) applied a computationally efficient method for 3D fruit position estimation using a monocular camera and introduced a rotation and hybrid translation controller for manipulator control. Furthermore, Lehnert et al. (2019) used a 3D camera array and a robotic manipulator to determine the optimal view of an object in a highly occluded and unstructured environment. Notably, most of these studies primarily focused on the challenges associated with crop detection. In general, visual servoing and detection issues are intertwined, and the incorporation of visual servoing into agricultural robot systems necessitates substantial research to surmount these obstacles (Fig. 5).

1.1.3. End-effector

An end-effector must exhibit excellent grasping and cutting abilities to be effective in unstructured and unpredictable agricultural environments (Huang et al., 2021). The grasping ability is necessary for handling the external force generated by the contact between the robot and the environment (e.g., other fruits, stems, or hanging gutters), which is a long-standing challenge for harvesting robots. To solve this problem, finger-type and suction-type grasping approaches have been developed.

Finger-type approaches enable the realization of human-like grasping control (Gao et al., 2022b). Although optimized grasping control for harvesting can be achieved, the end-effector is large (Roshanianfard and Noguchi, 2020), which increases the possibility of collision in agricultural environments and may result in undesired movement of the target fruits. Such undesired movement can considerably decrease the performance of harvesting robots (e.g., in terms of the harvest success rate and harvest time)

Suction-type approaches are characterized by simple grasping control and low collision possibilities owing to the compact size of the end-effectors (Hayashi et al., 2010; Huang et al., 2021). However, the control strategy must be designed considering the texture, sphericity, and stiffness of the target fruits and vegetables (Jun et al., 2021).

Previously, the authors of this work introduced an innovative end-effector capable of harvesting a variety of fruits and vegetables without requiring additional, complex controls (Park et al., 2022a). For efficient harvesting, an end-effector was designed, integrating modules such as a kirigami-based suction module for grasping, a circular-saw-based cutting module, and a transportation module for directing the harvested produce via free fall, aided by gravity, into the storage containers. The effectiveness of each module in minimizing harvesting time and enhancing productivity was validated through both laboratory and field experiments.

1.2. Contribution

The contributions and novelty of this study can be summarized as follows:

1. Development of a human-centered approach to cucumber harvesting that combines three key techniques for efficient and stable harvesting.
2. Evaluation of the performance of the proposed harvesting robot system through preliminary and field experiments.
3. Establishment of the foundation for future research to promote the development of more efficient harvesting robot systems by analyzing and discussing key issues through field experiments conducted in three sites.

2. Harvesting robot system

2.1. Robot system setup

The proposed robot system consists of a manipulator, an end-effector, and a mobile platform, as shown in Fig. 4. A manipulator is attached to the mobile platform. An end-effector is attached to the tool center point of the manipulator. The end-effector is equipped with a local camera and cutting and suction modules. In addition, an LED flash is installed that is always switched on to maintain an appropriate amount of ambient light to facilitate detection. The hardware and software setup is shown in Fig. 6. The controllers, which drive the manipulator, end-effector, and mobile platform, are connected to and controlled by the laptop. The laptop specifications are as follows:

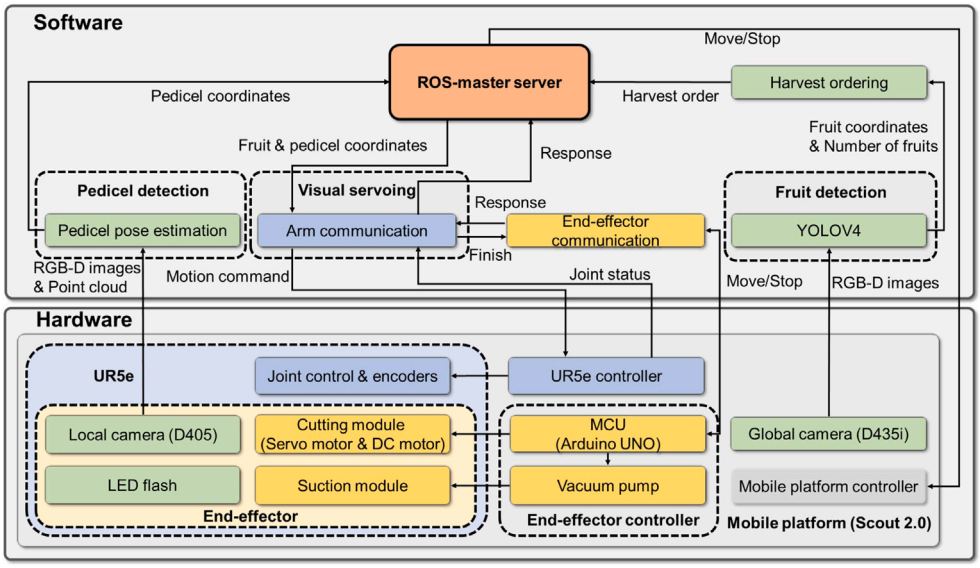


Fig. 6. Hardware and software setup of the proposed cucumber-harvesting robot system.

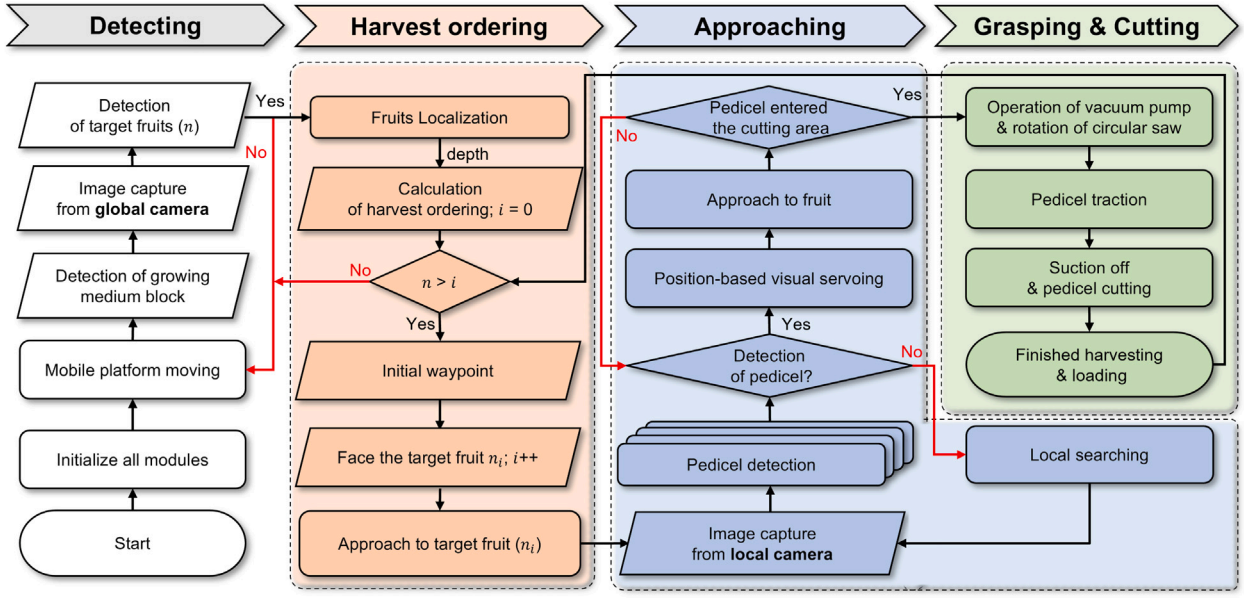


Fig. 7. Process flow of the proposed cucumber-harvesting robot system.

Intel (R) Core (TM) i7-9750H, 16 GB RAM, Geforce RTX 2060, with Ubuntu 18.04. The laptop transmits and receives data to and from each controller through robot operating system topics. The camera is operated by local cameras attached to the end-effector and global cameras attached to the manipulator.

The process flow of the proposed cucumber-harvesting robot system is shown in Fig. 3. As shown in Fig. 7, the process is designed considering the human approach. Similar to humans, the robot detects the cucumber, determines the harvest order, approaches the target, grasps it, and cuts it. The mobile platform locates a growing medium and detects cucumbers using the global camera. Harvest ordering is performed to determine the harvest order of n detected cucumbers and execute the first approach to the vicinity of the selected cucumber. Next, a second approach is performed, in which pedicels are detected and positioned using a local camera. Finally, the pedicel is cut after grasping the cucumber with the end-effector. The robot then returns to the initial pose and prepares for the subsequent harvest.

2.2. Harvest ordering

2.2.1. Perception

Compared with fruits such as tomatoes, apples, and oranges, the detection of cucumbers is challenging as their color is similar to those of the surrounding leaves and pedicels. Therefore, to effectively detect cucumbers using machine vision in real-world scenarios, robust data regarding cucumbers are required.

In this study, image data were collected from a cucumber farm in Jeollabuk-do and Chungcheongnam-do, Republic of Korea. The data were acquired using an RGB-D camera with a resolution of 1920 × 1080. Approximately 5000 images and 30 videos were collected using three cameras, and 20% and 10% of these images were randomly selected and used as the test and validation sets, respectively.

The YOLOV4 model, based on the YOLOV4-tiny architecture consisting of 29 convolutional layers, was used. The YOLOV4-tiny is a lightweight version of the YOLOV4 that is designed for fast and efficient

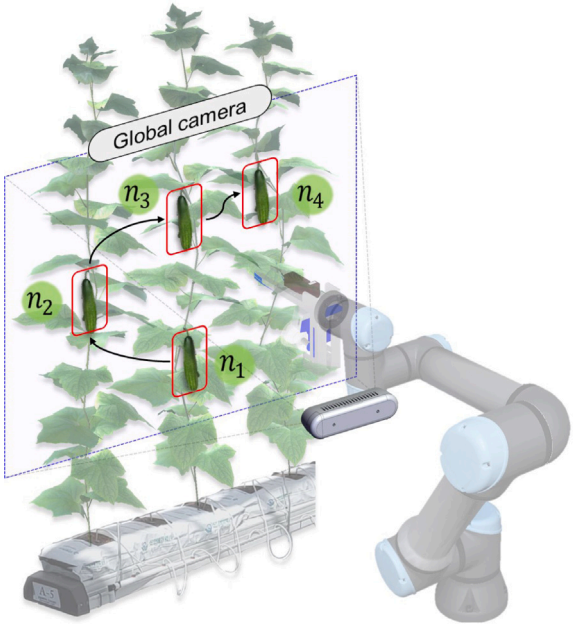


Fig. 8. Harvest ordering.

object detection on mobile and embedded devices. Compared with other deep learning-based detection methods, such as Faster RCNN, SSD, RefineDet, and EfficientDet the YOLOV4-tiny can maintain a high detection accuracy at a high speed (Cui et al., 2023). The cucumbers detected with a confidence score of more than 0.95 were identified by a bounding box. In other words, the cucumbers detected with high confidence were accurately localized within the image. The results of the cucumber detection system were evaluated based on the accuracy of the bounding box in localizing the cucumbers in the images.

2.2.2. Genetic-algorithm-based harvest ordering

As shown in Fig. 3, the method adopted by humans to determine the harvesting order was analyzed. According to the behavior of most agricultural workers, the first cucumber to be harvested is typically the one that is most visible and closest in proximity, followed by cucumbers that are adjacent to the first one. In other words, once the first cucumber in a group is determined, the harvesting order is based on the distance from this cucumber.

Fig. 8 demonstrates that a smart farm environment can be discretized into a set of fruits, $N = \{(x_1, y_1, z_1), (x_1, y_2, z_2), \dots, (x_n, y_n, z_n)\}$. The goal is to minimize the total harvesting time by optimizing the path of the manipulator for each fruit in N. The optimal order of harvest depends on the robot kinematics and plant structure and can be determined by applying the TSP to the fruit coordinates. A genetic-algorithm-based approach was used to optimize the order of fruit harvesting.

Generally, the solution to the TSP problem based on a genetic algorithm is aimed at finding the shortest distance without considering the environmental conditions. However, the fruit-growing environment is unstructured and often presents several obstacles that must be considered. Therefore, the proposed method introduces a new harvesting cost and models fruits in a manner that takes into account the environmental characteristics. The objective is to derive an optimal harvesting sequence that goes beyond simply minimizing the distance.

The fruits can be modeled as an undirected weighted graph with vertices, $G(V, E)$. The paths are the edges of the graph, and the distances of the paths are the weights of the edges. The optimal harvesting sequence problem is aimed at determining if the graph has a TSP solution with a maximum cost of K. If no path exists between multiple crops, an adequately long edge is introduced to complete the graph without affecting the optimal travel. After fruit localization, the fruits identified in the scene are prioritized for optimizing the harvesting sequence. It is assumed that the path of the manipulator to and from each fruit remains linear. The harvesting cost is defined as the duration of motion of the manipulator between two coordinates, i.e., the time required by the robotic arm for performing point-to-point motion. In this context, the harvesting cost is defined as the total path distance.

The path distance is calculated from the global camera to the position of the fruits. The total distance is calculated as Eq. (1).

$$d_{i,j} = \sqrt{(x[i] - x[j])^2 + (y[i] - y[j])^2 + (z[i] - z[j])^2}. \quad (1)$$

Note that the harvesting order is calculated considering the harvesting difficulty in addition to the distance. Considering the harvest ordering in the Z direction offers the advantage of robustness against obstacle perception. Therefore, the weights are defined considering the harvesting difficulty (i.e., the distance of z-direction), and the optimal harvesting sequence is $min(\sum w_{i,j})$.

$$w_{i,j} = d_{i,j}\sqrt{(z[i] - z[j])^2}. \quad (2)$$

2.2.3. Preliminary experiment

The efficiencies of two methods, i.e., random selection and the proposed method, for determining the harvesting order of four cucumbers were evaluated. The end-effectors approached the cucumbers in the order determined by each method. The following parameters were measured during the experiment: travel length (TL), which is the distance traveled by the end-effector, and operation time (OT), which is the duration required for harvesting. The results are summarized in Table 1.

Table 1
Results of preliminary experiment of harvesting ordering.

	TL (mm)	OT (s)
Random	9674.33	68
Proposed	5722.96	54

The results show that OT and TL are reduced through harvest ordering. Notably, this experiment by detecting four cucumbers. Therefore, the improvements are expected to be more notable when harvesting a larger number of cucumbers. The gap between the TL and OT, which served as a general quantitative indicator, is significant. This gap serves to reduce the unnecessary movement of the robot, thereby conserving battery time. Battery time is analogous to the physical stamina of humans and cannot be ignored. Overall, the proposed harvest ordering system can extend the battery of the robot.

2.3. Visual servoing

2.3.1. Pedicel detection for visual servoing

Fig. 3 shows that humans perform constant visual monitoring of the position of the cucumbers to accurately access the pruning shears. Next, the pedicel is placed in the cutting area of the pruning shears to the desired pedicel cutting point. In other words, the humans update the position of the pedicel in real-time and accurately position the pruning shears at the desired point. Harvesting robots must exhibit similar precision in their movements. To this end, a rapid pedicel detection method and shear access system must be developed.

However, the achievement of precise access in field conditions is challenging because of certain variables. For example, fruit motion caused by external factors such as interference from the end-effector during harvesting may result in manipulation errors in visual servoing. To address this problem, the altered pose of the fruit must be rapidly detected and corrected. Recent detection studies have extensively utilized deep-learning-based detection methods. According

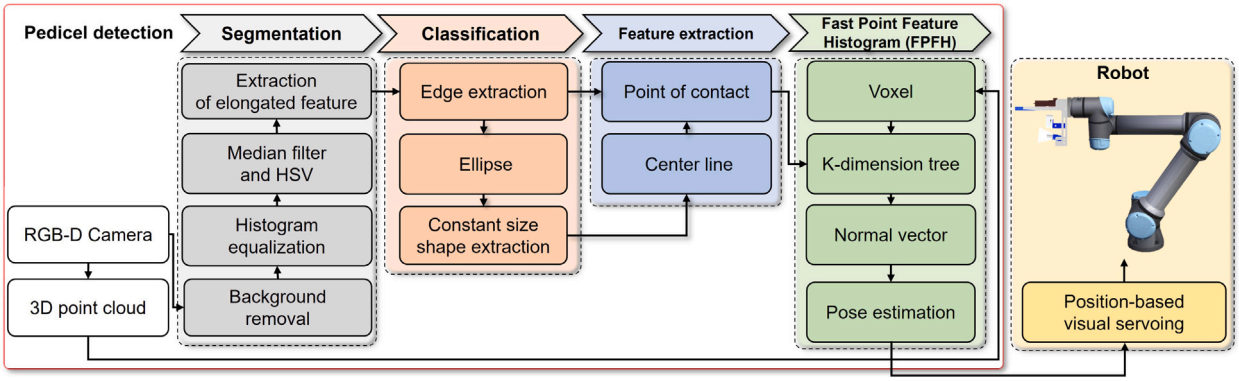


Fig. 9. Process flow of cucumber pedicel detection.

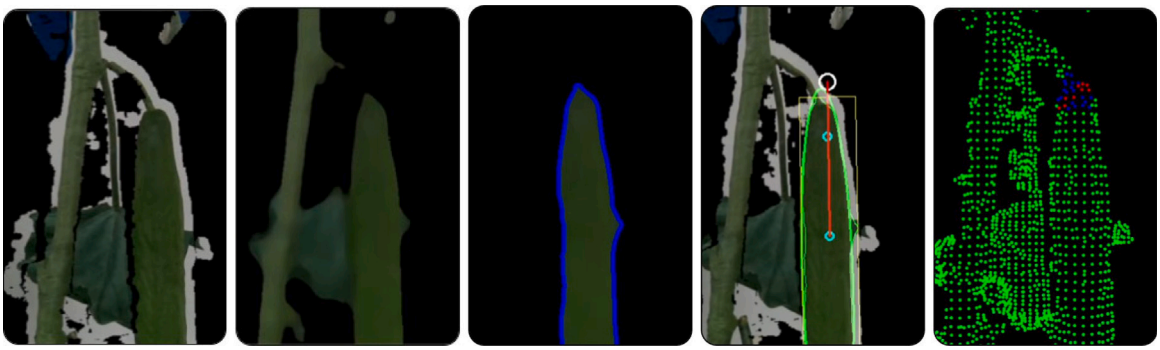


Fig. 10. Pedicel detection process: (a) background removal, (b) noise removal using a median filter, (c) classification of only the cucumber shape with an aspect ratio of 1:3–1:6, (d) extraction of the approximate position of the pedicel, (e) pedicel feature extraction using fast point feature histogram (FPFH) (red points).

to Gao et al. (2022a), a detection method with an average precision of 99.35% for fruit and stem detection can be considered adequately reliable. This method can be executed on a CPU with a frame rate of 2–5 frames per second (FPS). Efforts have been made to enhance the frame rate by applying a lightweight fruit detection algorithm, achieving a high frame rate ranging between 5.0 and 24.8 FPS (Zhang et al., 2021). Notably, the existing studies focused solely on detection, whereas when incorporating visual servoing for harvesting, it is important to estimate the appropriate grasping end-effector posture for each fruit pedicel before performing the manipulation (Wang et al., 2022).

Kim et al. (2022) introduced a deep learning network pipeline named Deep-ToMaToS, which can simultaneously estimate six-dimensional (6D) poses (3D transformation and 3D rotation) with a high average accuracy of 96%. However, this 6D pose estimation process necessitates a large inference time, which decreases the frame rate from 45.81 FPS to 7.26 FPS. Thus, when utilizing deep learning for pose estimation, the frame drop must be considered. In addition, the learning process requires a considerable amount of data. Overcoming the challenge of real-time pose updates remains critical to enhance the performance of visual servoing.

In this study, a computer-vision-based detection method was used for fast pedicel pose estimation. This method aims to enhance visual servoing performance by accurately estimating poses despite the shaking of fruits, as mentioned previously. Pedicel detection and pose estimation are realized in a simple manner. Moreover, the proposed approach is versatile and can be easily applied to other fruits without the need for learning processes In the proposed framework, pedicel detection is realized using a local camera attached to the end-effector, as shown in Figs. 4 and 9. Pedicel detection based on this process flow is shown in Fig. 10.

The background is removed using depth data from the RGB-D image to remove any obstacles or distractions that may interfere with the shape detection of the cucumber. Histogram equalization is applied to ensure reliable stability under various lighting conditions in complex and dynamic outdoor environments. This process allows the hue-saturation-value (HSV) color space to be more uniformly distributed, leading to increased stability in the thresholding process (Fig. 10(a)). To reduce shape detection failures due to noise, a median filter is applied, and remaining obstacles (i.e., the end-effector, trellis rope, and clamp) that are not cucumber-like (i.e., not green) are removed from the HSV color space (Fig. 10(b)). Therefore, considering the characteristics of the green series of cucumbers, the thresholds are set as ($H \in [60, 100], S \in [75, 255], V \in [80, 255]$).

Next, the characteristics of long cucumbers are exploited to classify the objects in the scene. Research suggests that cucumbers exhibit a length and diameter of approximately 250 mm and 40 mm, respectively, resulting in a ratio of 1:6. Therefore, the system is configured to detect only the cucumber shapes that fit this ratio among all the detected shapes. Additionally, only half of the cucumber shape is detected as the region of interest (ROI) of the local camera during the approach of the end-effector. Thus, the system is designed to detect shapes within a ratio range of 1:3 to 1:6. The edges of the detected cucumber shapes are extracted and fitted with an ellipse, and the height and width data of the ellipse are adjusted to fit the shape of the cucumber (Fig. 10(c)).

Finally, the intersection of the extracted edges and centerline of the ellipse is used to draw a circle representing the average pedicel pose (Fig. 10(d)). This pedicel pose is converted from pixel data on the two-dimensional (2D) plane to feature points on the 3D point cloud using camera parameters and depth data. The adjacent point clouds are then clustered to form an ROI. The fast point feature histogram (FPFH) algorithm is used to rapidly extract the feature points of the pedicel. The FPFH algorithm represents the geometric feature of the surface at

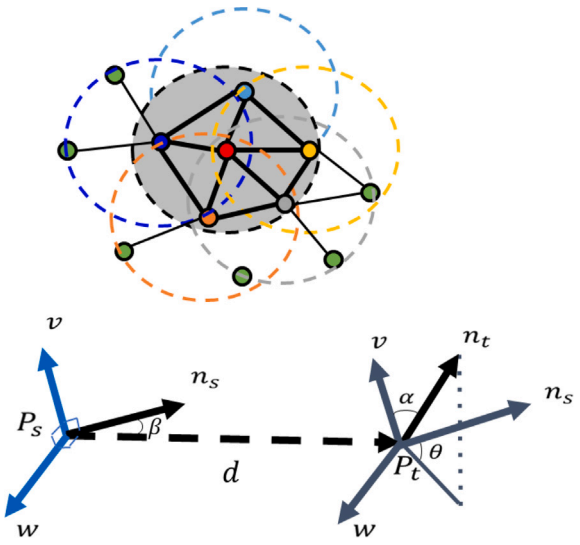


Fig. 11. Fast point feature histogram (FPFH) algorithm.

a given sample point and its neighboring points. This information is used to define a local coordinate system and extract feature points using curvature and direction differences between the cucumber and pedicel surface normal vectors (Fig. 10(e)). As shown in Fig. 11, given a set of points P_s and its neighboring points P_t in the 3D point cloud, the FPFH algorithm first defines a local coordinate system (u, v, w) as follows:

$$u = n_s, \tag{3}$$

$$v = \frac{(p_t - p_s)}{|p_t - p_s|} \times u, \tag{4}$$

$$w = v \times v, \tag{5}$$

where n_s is the surface normal vector of P_s. Next, a set of tuples (α, ϕ, θ) between each point and its neighboring points is computed using the local coordinate system:

$$\alpha = v \cdot n_t, \tag{6}$$

$$\phi = u \cdot \frac{(p_t - p_s)}{d}, \tag{7}$$

$$\theta = \arctan\left(w \cdot n_t, u \cdot n_t\right), \tag{8}$$

where $d = |p_t - p_s|$ is the Euclidean distance between P_s and P_t, and n_t is the surface normal vector of P_t. Finally, the FPFH histogram of P_s is computed by weighting the simplified point feature histogram (SPFH) values ω_i of its neighboring points:

$$FPFH\left(p_s\right) = SPFH\left(p_s\right) + \frac{1}{k}\sum_{i=1}^{k}\frac{1}{\omega_i} \cdot SPFH\left(p_i\right). \tag{9}$$

The FPFH descriptor is calculated for each point in the point cloud, and it represents the distribution of the local surface normal orientations around the point. The calculation of the FPFH descriptor involves considering a set of k neighboring points around each point and computing the histogram of their surface normal orientations. The weight of each neighboring point can be calculated based on a weighting function, such as the inverse distance or ω_i. The histograms of the cucumber fruit section and pedicel section are different, and this difference is exploited to extract the point cloud of the pedicel portion.

The optimization problem can be formulated as follows: Given two point clouds P and Q, the objective is to find the transformation matrix $\mathbf{T}$ such that the transformed source point cloud P' best matches the target point cloud Q in a least-squares sense. Here, $\mathbf{T}$ can be estimated by minimizing the sum of the squared distances between the closest points in the two clouds:

$$\mathbf{T} = \arg\min_{\mathbf{T}} \sum_{i=1}^{n} \|\mathbf{T}p_i - q_{j(i)}\|^2. \tag{10}$$

where p_i is the ith point in the source point cloud, P, $q_{j(i)}$ is the closest point in Q to p_i, and n is the number of points in P. The ICP algorithm is used to solve this optimization problem by iteratively refining an initial estimate of $\mathbf{T}$. In each iteration, the algorithm computes the closest points between the source and target point clouds, updates $\mathbf{T}$ based on the closest points, and repeats the process until convergence or until the maximum number of iterations is reached.

The attitude and orientation of $\mathbf{T}$ can be obtained using the orientation representation in the rotation matrix. In particular, a rotation matrix is a square matrix that represents a rotation in 3D space and can be used to calculate the attitude and orientation of $\mathbf{T}$.

2.3.2. Position-based visual servoing (PBVS)

PBVS is applied for robot-based cucumber harvesting. PBVS involves calculating the desired pose of the end-effector based on the coordinates of a target feature point, such as the cutting point of the cucumber pedicel. The desired pose of the end-effector is calculated using $\mathbf{T}$.

Specifically, $\mathbf{T}$ is calculated based on the pose of the target feature point in the 3D space and pose of the end-effector in the image plane. As shown in Fig. 12(a), the desired pose of the end-effector, F_{c*}, in the image plane is represented as a homogeneous image coordinate. The current pose of the end-effector, F_c, in the image plane can be represented as a homogeneous image coordinate. The error e between the desired and current end-effector poses can be calculated as follows:

$$e = F_{c*} - F_c. \tag{11}$$

e is used in the control law to generate the joint velocity command, $\boldsymbol{\omega}$, for the robot arm. The control law can be represented as follows:

$$\boldsymbol{\omega} = -\lambda \mathbf{J}^{+}(F_{c*} - F_c). \tag{12}$$

where $\boldsymbol{\omega}$ is the joint velocity command, $\mathbf{J}$ is the Jacobian matrix, and λ is a gain parameter that indicates the tracking speed and stability. A higher λ results in faster tracking but may lead to oscillations and instability. In contrast, a lower λ results in slower tracking but improved stability. The value of λ must thus be appropriately tuned to optimize the system performance.

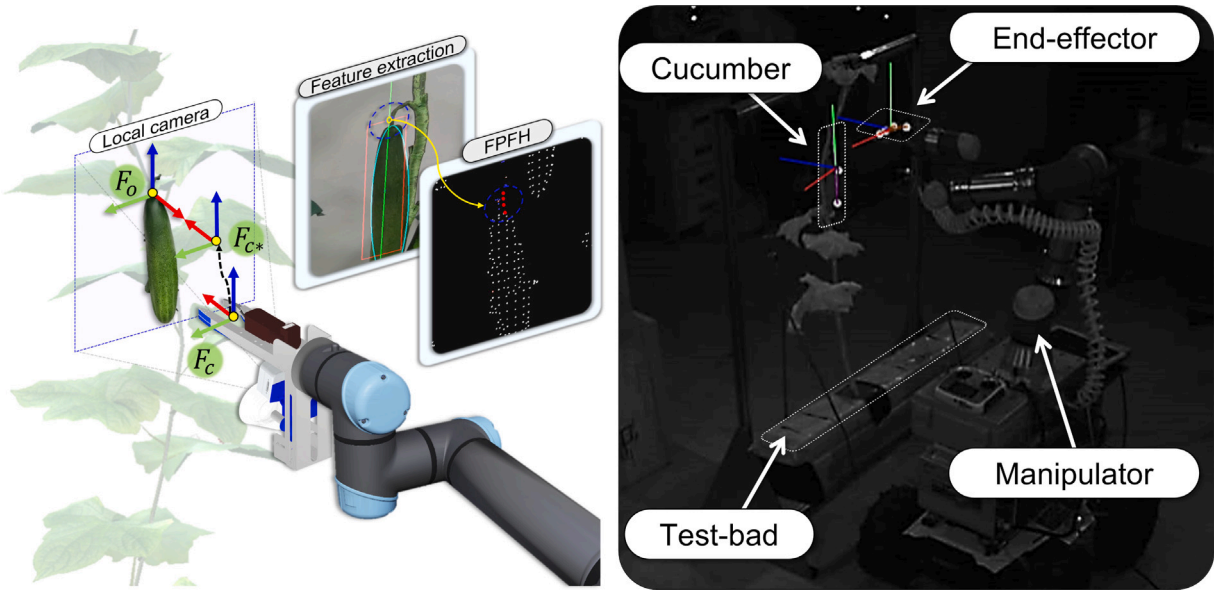
By implementing PBVS to track the cutting point of the cucumber pedicel, the robot can maintain the correct cutting pose, even if the cucumber shakes or moves due to interference from the robot movement. This framework thus improves the accuracy and reliability of cucumber harvesting by robots. Notably, $\mathbf{T}$ plays a crucial role in PBVS by providing the necessary information to calculate the desired pose of the end-effector based on the target feature point.

2.3.3. Preliminary experiment

A preliminary experiment was performed to validate the proposed visual servoing system for robot-based cucumber harvesting. The experiment was performed in a laboratory environment that mimicked a cucumber smart farm greenhouse. Eight motion capture cameras were used to measure the pose of the cucumber and end-effector during the experiment.

Three markers were attached to the model cucumber and end-effector, as shown in Fig. 12(b). The target pose was the pose in which the cucumber pedicel entered the cutting area of the end-effector. The experiment was repeated 100 times, in which the end-effector was moved from an arbitrary pose to the target pose. The objective was to manipulate F_c in the image plane to shift to F_{c*}. F_c measured by the motion capture system was set as F_a. F_{c*} was set to ensure that the pedicel can be positioned in the end-effector cutting area. The pose error (PE) was calculated as follows:

$$PE = F_{c*} - F_a. \tag{13}$$



(a) (b)

Fig. 12. Position-based visual servoing (PBVS): (a) visual servoing process, (b) motion capture camera view.

Table 2
Results of the preliminary experiment for visual servoing.

	Position [mm]			Orientation [deg]		
	x	y	z	x	y	z
PE average	1.44	−2.45	−0.29	−0.25	0.63	0.91
RMSE	4.90	4.73	1.95	1.71	3.38	2.74

The root mean square error (RMSE) was used to calculate the effective value:

$$\mathrm{RMSE} = \sqrt{\frac{\sum_{i=1}^{n} PE^2}{n}}. \tag{14}$$

The results of the preliminary experiment were used to evaluate the performance of the proposed visual servoing system. The accuracy and reliability of the system were determined by comparing F_c to F_{c*}. The experimental results are summarized in Table 2. The results show that the pedicel can be located within the cut area of the proposed end-effector, as shown in Fig. 13. Because the pedicel can be positioned within the cutting area, the end-effector can be efficiently manipulated to the pedicel in real-time through the proposed pedicel detection and visual servoing strategy. Furthermore, the FPS achieved by the proposed method (16–23 for images with 640 × 480 pixels) is higher than that of deep learning methods. This high FPS ensures real-time performance during the implementation of visual servoing.

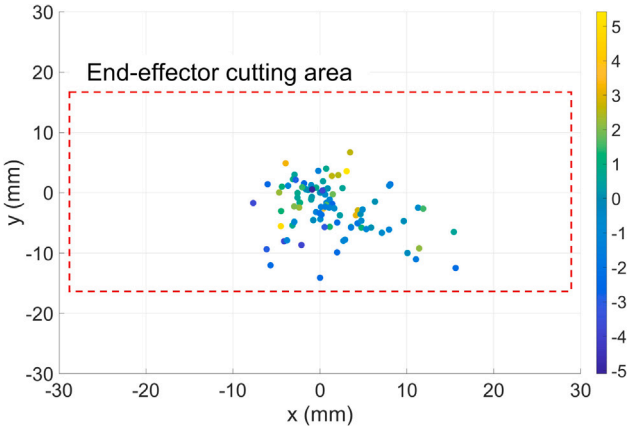


Fig. 13. Position error of the end-effector.

2.4. End-effector

The end-effector functions similar to a human hand in the human-centered harvesting approach. As shown in Fig. 3, humans perform harvesting by isolating individual fruits by separating the pedicels connected to the stem. The harvesting method involves plucking, twisting, and cutting processes. However, the end-effector of harvesting robots is typically composed of a rigid body, and thus, plucking and twisting may damage the fruit surface. To address this problem, the concept of manual harvesting, in which fruits are treated as soft materials and pedicels are used to harvest fruits, has been adopted.

Specifically, following this mechanism, an end-effector for tomato harvesting, a spherical fruit, was developed, and its performance was evaluated in field conditions. Unlike tomatoes, cucumbers are cylindrical and have a bumpy surface. In general, rod-shaped fruits are more susceptible to shaking than sphere-shaped fruits because of the large distance between the pedicel–fruit contact position and center of mass. Therefore, the design of an end-effector for rod-shaped fruits, such as cucumbers, requires more rigid considerations. In this study, the previously developed end-effector was modified to suit cucumber-harvesting requirements. Particularly, the proposed end-effector was designed to imitate manual harvesting mechanisms with cutting and grasping modules. The following sections describe the end-effector and highlight its novelty compared with the existing end-effectors.

2.4.1. Cutting module

In this study, a circular saw-type cutting module was used for the harvesting process. Cutting, a critical step in manual harvesting, is typically executed using scissors. Previously, Jun et al. (2021) introduced a scissor-type end-effector for tomato harvesting. However, the large size of scissor-type end-effectors often leads to conflicts in dense and irregular agricultural environments. These conflicts with environmental objects can introduce errors in robot perception and control, resulting

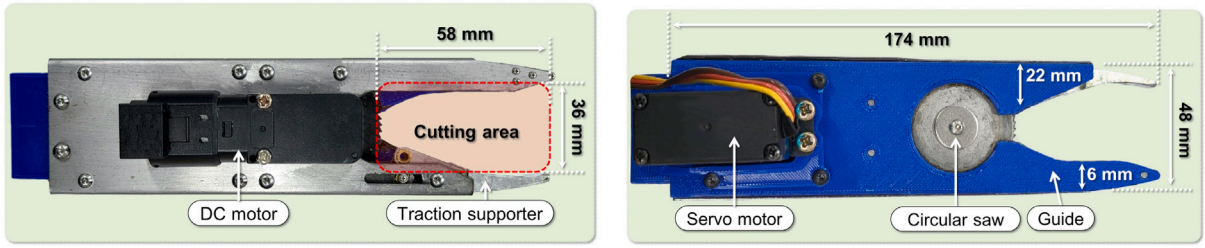


Fig. 14. Cutting module of the proposed end-effector: (a) top view, (b) bottom view.

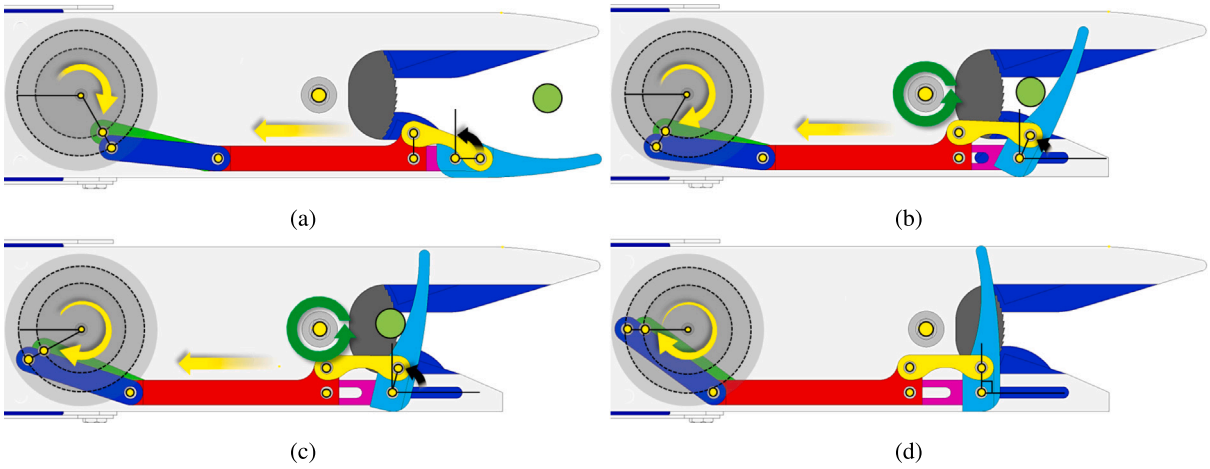


Fig. 15. Mechanism of the proposed cutting module: (a) the pedicel enters the cutting area, (b) the servo motor rotates to pull the traction supporter to tow the pedicel, and the dc motor rotates the circular saw, (c) the pedicel is fully towed to the circular saw, (b) the pedicel is cut, and the servo motor is reversed to return the traction supporter to its initial position.

in reduced overall performance in terms of the harvesting time and success rate.

To address this issue, a compact circular saw-type end-effector was designed in this study to minimize undesirable contact with the environment (Park et al., 2022a). Generally, a circular saw tends to recoil off the pedicel due to its rotation. To overcome this problem, a traction supporter that imitates manual harvesting mechanisms, specifically dual-arm mechanisms, was introduced. In manual harvesting operations, one hand is used for cutting, while the other is engaged in holding and pulling the cucumber and drawing the pedicel to the cutting blade or scissors. The traction support feature of the circular saw-type end-effector emulates this second hand.

The traction supporter holds the cucumber and guides the pedicel into the circular saw of the cutting module. This holding and pulling action enables a reliable cut regardless of the pedicel thickness. The performance results of the circular saw-type cutting module demonstrated a 93.8% success rate when applied to pedicels with a thickness of 4– 11 mm, given an end-effector angle of 0°. Field observations have indicated that the thickness of cucumber pedicels typically falls within the 3–6 mm range, which suggests that the circular saw-type cutting mechanism can be efficiently utilized under such conditions. The focus of this study was to develop a three-point linkage cutting module to enhance the performance of the traction supporter and ensure a consistent pedicel length. A limitation of circular saw-type end-effectors was identified: Such end-effectors fail to transmit sufficient force to pull a thick tomato pedicel due to constraints in the traction mechanism. Consequently, modifications were made to the traction mechanism, as shown in Fig. 16, to augment the force conveyed to the pedicel via the traction supporter.

The traction mechanism within the three-point linkage cutting module was divided into traction and open-close operations. The traction

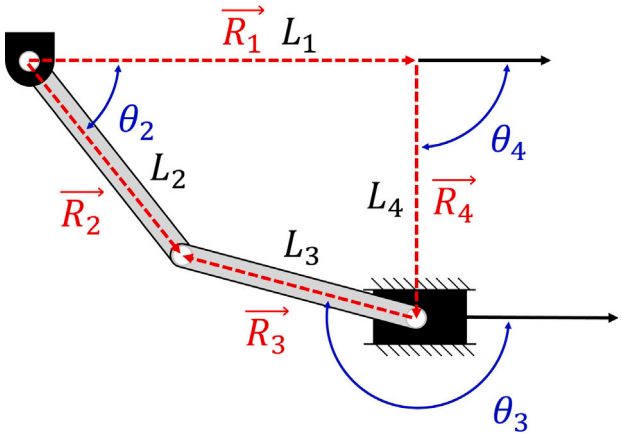


Fig. 16. Simplified offset slider-crank linkage of the cutting module.

was powered by a slider-crank mechanism linked to the servo motor, as illustrated in Fig. 15. The function of the slider-crank mechanism was to pull the pedicel into the circular saw, making it an essential component of the traction operation.

For the open-close operation, two links were interconnected based on the crank radius, with the discrepancy in the radii leading to a difference in the operational range between the two links. Each slider was connected to these two links, and the sliders, in turn, were connected to disparate positions on the traction supporter.

By exploiting the differences in the connection positions and operational range, a three-point linkage structure was generated. This

Table 3
Design parameters for the offset slider-crank.

Parameter	Symbol	Short radius linkage	Long radius linkage
Angle [deg]	θ_1	0	0
	θ_2	60	60
	θ_3	165	175
	θ_4	90	90
Length [mm]	L_1	32	38
	L_2	12	17
	L_3	30	30
	L_4	17	17

structure enabled the traction supporter to execute an open–close operation, providing access to the cutting area of the end-effector in an idle state. Even in the presence of perception and control errors during the approach process, successful entry could be achieved due to the presence of an approach margin. The following expression pertains to the vector equation of the offset slider-crank in the proposed cutting module:

$$-\overrightarrow{R_1} + \overrightarrow{R_2} - \overrightarrow{R_3} - \overrightarrow{R_4} = 0. \tag{15}$$

As depicted in Fig. 16, the vector equation of slider-crank can be represented as follows:

$$-L_1 e^{j\theta_1} + L_2 e^{j\theta_2} - L_3 e^{j\theta_3} - L_4 e^{j\theta_4} = 0. \tag{16}$$

The design parameters for the offset slider-crank of the proposed cutting module are presented in Table 3.

The operating range of the link connected to the crank can be expressed based on Euler's formula and the design parameters. The operating range of the slider can be calculated as

$$\begin{aligned} e^{j\theta} &= cos\theta + jsin\theta, \\ L_1 &= L_2 cos\theta_2 - L_3 cos\theta_3. \end{aligned} \tag{17}$$

The cutting module was equipped with a local camera for pedicel detection and a grasping module at the bottom. Details of the grasping module are presented in the following section.

2.4.2. Grasping module

The grasping module in the proposed cucumber-harvesting system was significantly modified compared with the existing end-effector. This module plays a crucial role in mitigating the reaction force resulting from the rotation of the circular saw in the cutting module. In the previous framework, a suction-cup-based approach was used for the grasping module due to its compact size and simple control mechanism. The circular-saw-type end-effector in the previous research utilized conical-shaped general suction cups to grip tomatoes. This type of a grasping module is generally suitable for spherical fruits, such as citrus fruits or apples.

However, cucumbers present a different challenge due to their rod-like shape and low sphericity. Applying the existing conical suction cup to cucumbers may not ensure a secure grip, as the attachment success may vary with the approach position. Restricting the attachment point is not beneficial in reducing the reaction force, given that defining a generalized approach position is challenging in non-standard agricultural environments. Therefore, the conical suction cup of the existing end-effector is not well-suited for cucumber attachment scenarios.

To solve this problem, as depicted in Fig. 4, a grasping module designed for grasping cylindrical fruits such as cucumbers was established. Similar to the existing framework, a suction-cup-based grasping module was used to ensure compactness. However, the grasping power was increased by changing the circular grasping area of the suction cup to an elliptical shape. The dimensions of the ellipse were determined based on the minimum values obtained by examining 10 commercially available cucumbers. The grasping area of the suction cup, which is related to the grasping power, can be expressed as follows:

$$b^2x^2 + a^2y^2 = a^2b^2. \tag{18}$$

In this equation, the values of a and b are 14 and 10, respectively. By changing the grasping area to an ellipse, the grasping area can be kept constant regardless of the grasping position. Using the calculated grasping area, the grasping power can be determined as follows:

$$F_{z,max} = \pi r^2 \Delta V, \tag{19}$$

where ΔV represents the pressure difference between the suction cup and atmosphere. The grasping area considerably affects the grasping force, as indicated in the equation. If the grasping area is not sufficient, the grasping force may not be adequate, and grasping may fail depending on the grasping position.

2.4.3. Preliminary experiment

The performance of the end-effector for tomatoes has already been validated in the field. Thus, this end-effector can be considered suitable for cutting cucumber pedicels, which are typically thinner (3–6 mm) than tomato pedicels (4–11 mm). However, a key difference between this study and the previous study pertains to the evaluation of the cucumber grasping. Specifically, in this study, the feasibility of grasping cucumbers was evaluated in test beds resembling cucumber farms. The performance metric was the grasping success rate, which is defined as maintenance of the grip even when the robot returns to its initial position after adsorption. Evaluating the success rate of cucumber grasping is important for addressing the issue of grasping ability being dependent on the grasping position in cucumber harvesting scenarios. The grasping module of the proposed end-effector was compared with the conical suction cup used in the existing framework. For the preliminary experiment, a test bed was set up to allow cucumbers to hang like they would in real fields. The results revealed that the proposed end-effector has a grasping success rate of 100%, whereas the success rate of their previous end-effector was 0%.

3. Field experiments

3.1. Experimental design

The experimental setup is shown in Fig. 4. Field tests were conducted by integrating each verified system in the overall framework. Experiments were conducted in actual cucumber greenhouses to verify the effectiveness of the proposed cucumber-harvesting robot system. As shown in Fig. 17, the experiment was conducted in three greenhouse environments to examine the effect of the environmental conditions on the system performance. The study areas were Boryeong Green monsters (GM), Chungju Fresh-farm (FF), and Sangju smart-valley (SV), in which *Cucumis sativus L* are grown. Experiments were conducted in different sites because the pedicel thickness, length, and space between the stems and pedicels differ across the field. The harvesting sequence was as shown in Fig. 18.

The average parameters in the three sites are shown in Fig. 19. The local camera used in the system has a maximum detection depth between 70 and 500 mm. The global camera used in the system has a maximum recognition depth between 28 and 3000 mm. The effective depth is appropriate because the depth required for detection with the global camera is about 1000 mm and the depth required for detection with the Local camera is about 300 mm. In addition, the intertidal distance between cucumber stems is about 400 mm, and 3–4 stems are detected within the ROI of the global camera (the average number of cucumbers is 4–6). The distance between the mobile chassis and the planting row depends on the size and shape of the mobile platform and greenhouse layout. A reasonable distance is typically set between 600 and 700 mm to allow adequate space for maneuvering and avoid collisions.

To verify classification performance, measure Precision, Recall, Accuracy, and F1-score based on classification evaluation metrics. True positive (TP) refers to the detection of cucumbers that are present in the field and harvested correctly by the end-effector. False positive (FP)

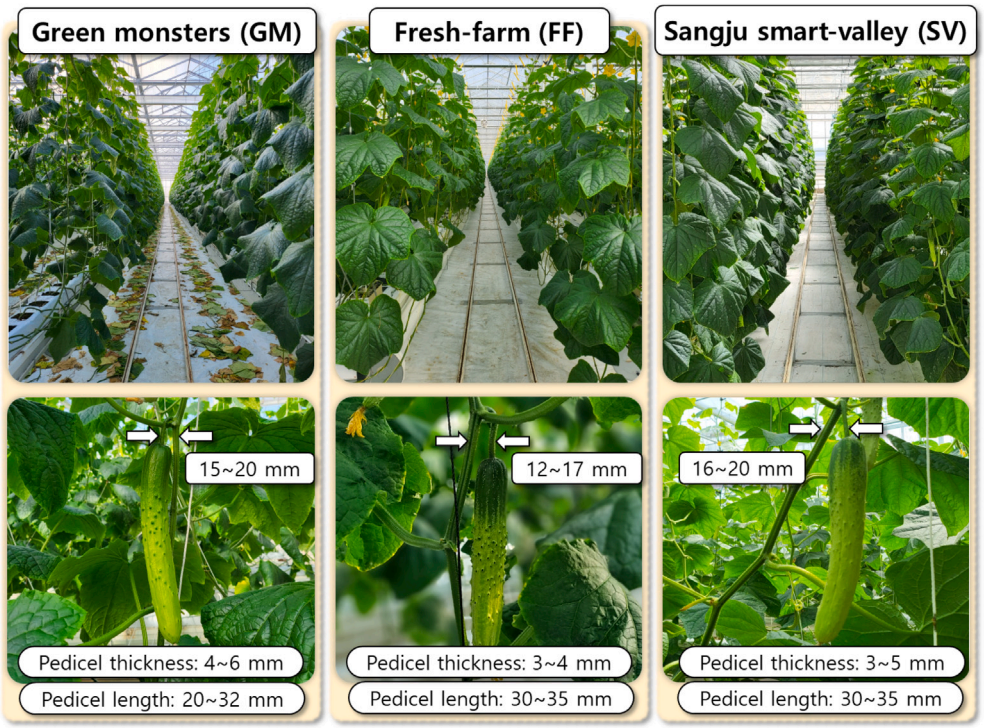


Fig. 17. Experimental environment and cucumber characteristics in three sites.

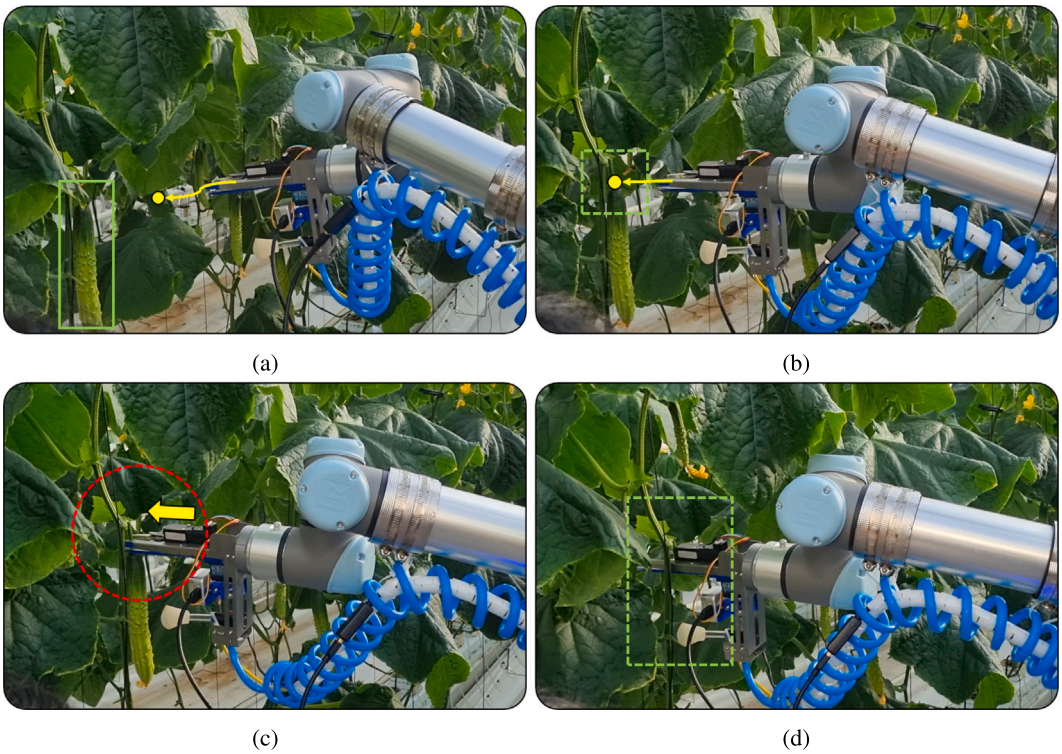


Fig. 18. Harvest in progress according to the harvest sequence of the proposed cucumber-harvesting system: (a) The global camera detects cucumbers and approaches the nearest cucumber, (b) the local camera detects the cucumber pedicel and approaches the end-effector, (c) the pedicel is input to the end-effector cutting area, and (d) the end-effector cuts the pedicel to complete the harvest and returns to its initial position.

refers to the detection of cucumbers that are not present in the field but are still harvested by the end-effector. False negative (FN) refers to the failure of the end-effector in recognizing cucumbers that are present in the field, resulting in these cucumbers not being harvested. True negative (TN) refers to the correct detection of cucumbers that are not present in the field or not harvested by the end-effector.

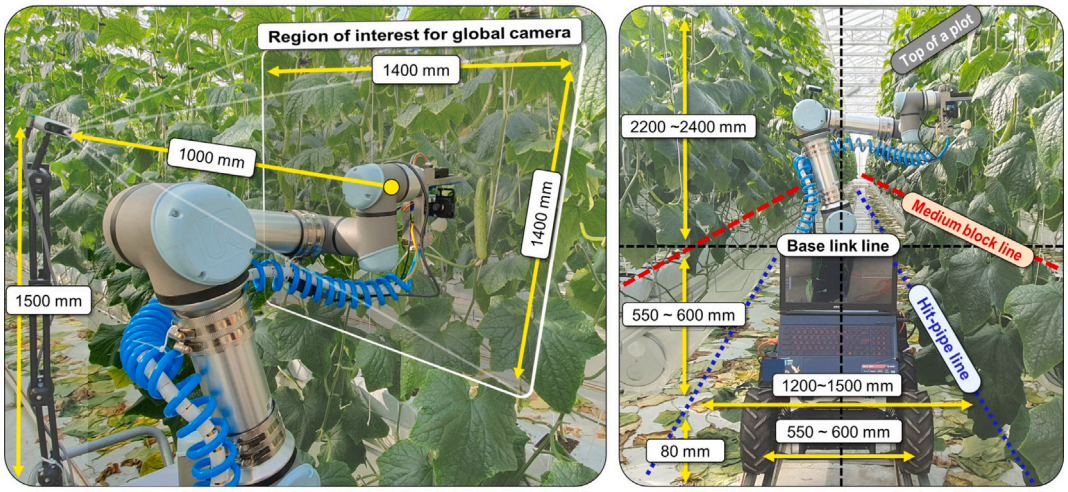


Fig. 19. Parameters in the cucumber greenhouse environment.

The precision is calculated as

$$Precision = \frac{TP}{TP + FP}, \tag{20}$$

The recall can be calculated as

$$Recall = \frac{TP}{TP + FN}, \tag{21}$$

The accuracy is calculated as

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN}, \tag{22}$$

The F1-score is calculated as

$$F1 - score = 2\frac{(Precision \times Recall)}{(Precision + Recall)}. \tag{23}$$

In general, the F1-score is a valuable measure of the overall accuracy of a classification system when the data are imbalanced, i.e., there is a significant disparity between the number of positive and negative cases. In such cases, the use of only precision or recall as a single metric may be misleading. Because the F1-score considers both the precision and recall, it can effectively indicate the overall performance of the system.

3.2. Results and evaluation

The success rate of the proposed cucumber-harvesting robot system was calculated by dividing the yield into detection, entry, and cutting. A total of 265 cucumbers were harvested from the three sites (GM, SV, and FF), with 150 successful harvests. The system performance was evaluated in terms of the success rate, harvest time, precision, and recall rate. As shown in Table 4, the success rate for the GM site was 56.3%, with 98 successful detections out of 112, 82 successful entries out of 112, and 63 successful cuts out of 112. The SV site had a success rate of 50.9%, with 35 successful detections out of 53, 29 successful entries out of 53, and 27 successful cuts out of 53. The FF site had a success rate of 60.0%, with 89 successful detections out of 100, 63 successful entries out of 100, and 60 successful cuts out of 100. The results showed that the success rate of the proposed system ranged from 50.9% to 60.0% at different sites, with an overall success rate of 56.6%. The average damage rate for the harvested cucumbers was 4.7%, with seven damaged cucumbers. The FF site exhibited the highest success rate of 60.0%. These results indicated that the system performed differently in different environments, and further research is needed to optimize the system performance in different conditions.

The harvesting time is shown in Fig. 20. As shown in Fig. 18, first, the global camera detected the cucumber, and then, the local camera detected the cucumber pedicel. Next, the cucumber was cut, and the system returned to its initial pose. The average harvesting time for the first and second approaches was 13.2 s and 30.7 s, with a range of 11–15 s and 21–42 s, respectively. The average cut-off time was 5 s, with a range of 2–7 s. The return times were mostly constant, with an average of 7.3 s and a range of 7–8 s.

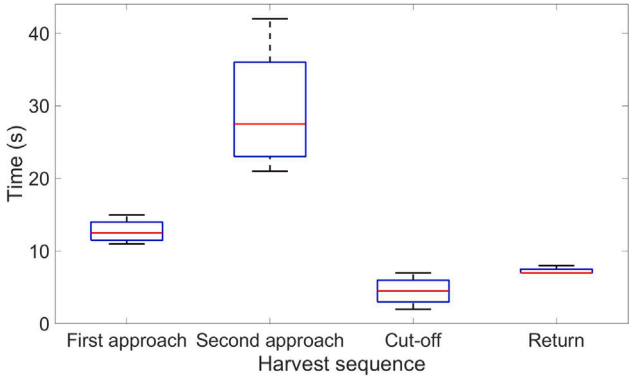


Fig. 20. Harvesting time by sequence.

Table 5 summarizes the results of the performance evaluation of the cucumber-harvesting robot system at the three sites. The system performance was evaluated based on four metrics: precision, recall, accuracy, and F1-score. The precision represents the fraction of correctly harvested cucumbers to the total number of harvested cucumbers. The recall represents the fraction of correctly harvested cucumbers to the total cucumbers present in the field. The accuracy represents the ratio of correct predictions to the total number of predictions. The F1-score is the harmonic mean of precision and recall and provides a balanced picture of the overall performance of the system.

For the GM site, the precision, recall, accuracy, and F1-score were 95.1%, 91.6%, 87.5%, and 93.3%, respectively. For the SV site, the precision, recall, accuracy, and F1-score were 81.4%, 77.7%, 66.0%, and 79.5%, respectively. For the FF site, the precision, recall, accuracy, and F1-score were 95.7%, 92.7%, 89.0%, and 94.2%, respectively. Although the proposed system achieved satisfactory precision and accuracy, there is scope for further improving the recall.

In general, it is challenging to set up a motion capture system in the field, as illustrated in Fig. 12(b), which makes it difficult to accurately measure locations and determine the accuracy indicators in the field. To overcome this challenge, the positional accuracy was evaluated by measuring the length of the cut pedicel. The assumption was that the

Table 4
Field experiment results of the proposed cucumber-harvesting robot system.

Sites	Evaluation	Harvest sequence			Total	
		Detection	Entry	Cutting	Harvesting success	Damage
Green monsters (GM)	Success rate (%)	87.5	83.7	76.8	56.3	4.8
	Number of successes (pcs)	98	82	63	63/112	3
Sangju Smart-valley (SV)	Success rate (%)	81.4	82.9	93.1	50.9	7.4
	Number of successes (pcs)	35	29	27	27/53	2
Fresh-farm (FF)	Success rate (%)	89.0	70.8	95.2	60.0	3.3
	Number of successes (pcs)	89	63	60	60/100	2
Total	Success rate (%)	83.8	78.4	86.2	56.6	4.7
	Number of successes (pcs)	222	174	150	150/265	7

Table 5
Pedicel detection precision, recall, accuracy, and F1-score for different sites.

Site	Precision	Recall	Accuracy	F1-score
Green monsters (GM)	0.951	0.916	0.875	0.933
Sangju Smart-valley (SV)	0.814	0.777	0.660	0.795
Fresh-Farm (FF)	0.957	0.927	0.890	0.942

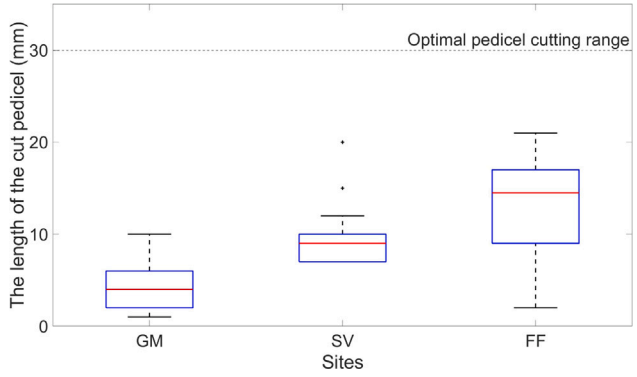


Fig. 21. Length of the cut pedicel.

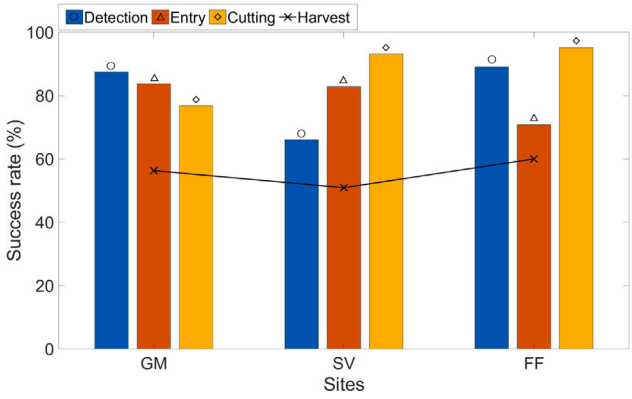


Fig. 22. Analysis of field test results by harvest sequence.

accuracy of the approach could be determined if the pedicel was cut in the correct pose. The experiment results are shown in Fig. 21. The optimal cut length for the pedicel, as preferred by different farmhouses, was between 0 and 30 mm, and all the cuts were made within this range. This finding highlighted that the success rate of the harvest can be increased by improving the accuracy of cucumber detection and reducing interference from the end-effector. Furthermore, the results of the preliminary experiment demonstrated that the proposed system can achieve a high positional accuracy.

4. Discussion

4.1. Failure cases

The results of extensive experiments showed that the proposed cucumber-harvesting robot system performed well in the detection and cutting sequences (success rates of 83.8% and 86.2%, respectively) but had lower success rates in the entry sequence (78.4%). Furthermore, the total harvesting success rate was only 56.6%. Several causes of harvest failure were identified and broadly categorized as follows:

- Pedicel detection failed due to obstacles (stems, leaves, etc.).
- The stem and pedicel both entered the cutting area owing to the width of the end-effector pedicel entrance.
- Interference occurred between the stem and pedicel during entry because of the size of the end-effector.

The primary cause for harvest failure was the presence of obstacles (stems, leaves, etc.), which made it challenging to detect cucumbers and cucumber pedicels. The experimental results presented in Table 4 were analyzed to identify the reason for the problems. As shown in Fig. 22, the detection success rate varied across the study sites. Table 5 shows that the detection success rate at SV (precision of 0.814, recall of 0.777, accuracy of 0.660, and F1-score of 0.795) was lower than those at the other two sites. The detection success rate was affected by the presence of defoliation in the farmhouses: Specifically, GM and FF had a more organized cultivation environment due to defoliation, resulting in fewer leaves and higher detection success rates. To address the challenge of detecting cucumbers and pedicels in the presence of obstacles, it is necessary to develop a system that can accurately detect and access cucumbers even in scenarios involving environmental disruptions.

The second main reason for harvest failure was the width of the end-effector pedicel entrance, which allowed both the stem and pedicel to enter the cutting area. As shown in Fig. 23(a), the space between the stem and pedicel is narrow. Fig. 22 shows that among the three sites, FF corresponded to the lowest entry success rates (70.8%, compared with 83.7% and 82.9% at GM and SV, respectively). This phenomenon of both the stem and pedicels entering the cutting area occurred because of the distinct characteristics of the cucumbers at each site. Fig. 17 shows that the space between the stem and pedicel is the smallest at FF. As shown in Fig. 14, the width of the end-effector pedicel entrance is 36 mm, which can accommodate both the stem and pedicel. Twelve of the 26 failed entries after detection, i.e., 46.2%, occurred at a high rate (6.1% for GM and 6.9% for SV). The end-effector pedicel entrance width must be decreased during the design phase to avoid such problems. Assuming that the end-effector is miniaturized, the entry success rate at FF is expected to be increased to 84.3%, with the values for GM and SV improving to 88.8% and 88.6%, respectively.

In particular, the end-effector size considerably affects the success of harvesting cucumbers as it affects the ability of the end-effector entering the space between the stem and pedicel. As shown in Fig. 22, the entry success rate was low at GM owing to the short and thick pedicel of the cucumbers at this site, resulting in interference at the

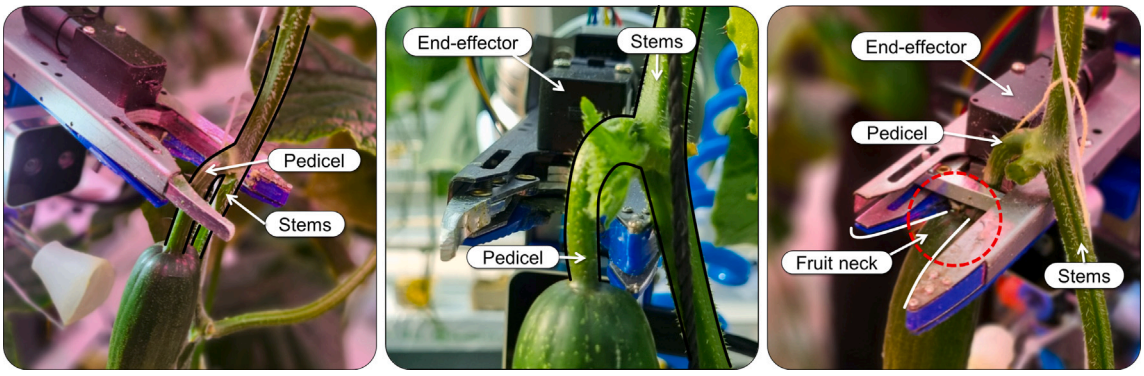


Fig. 23. Problems encountered by the end-effector during harvesting: (a) The stem and pedicel both entered the cutting area owing to the width of the end-effector pedicel entrance, (b) the end-effector size leads to interference between the stem and pedicel during entry, (c) the end-effector guide size leads to interference between the stem and pedicel during entry.

neck of the fruit, as shown in Fig. 23(b). The same phenomenon frequently occurred when the pedicel was towed with a towing supporter, as shown in Fig. 23(c). The guide to the cutting module in the system helped guide the pedicel to a point on the circular saw to ensure fast and accurate cutting. However, the guide may disrupt the cutting process if the pedicel is short and thick, as in the GM environment. To overcome this issue, the end-effector must be miniaturized. The end-effector cutting module of the proposed system has a novel linkage structure, as shown in Fig. 15, which is simple and can facilitate the miniaturization of the device. Specifically, miniaturization can be achieved by simplifying the linkage structure and frame of the end-effector.

4.2. Future work

The proposed robot can be further improved by focusing on the following aspects: (1) miniaturization of end-effectors, (2) development of robust visual servoing through accurate pedicel detection, (3) development of lift platforms for a wide range of harvests, and (4) extension of the proposed method to harvest a variety of fruits.

As a follow-up to this study, ongoing efforts are being made to miniaturize the end-effector. Previous studies have shown improvements by transitioning from a scissors-type end-effector structure (Jun et al., 2021) to a circular saw-type end-effector (Park et al., 2022a). In particular, these modifications have resulted in a high cutting rate with reduced fruit damage, thereby enabling efficient pedicel cutting. However, the mechanism used by such end-effectors requires two types of motors: a linear servo actuator and a DC motor. Servo motors exhibit reduced torque with decreasing size, and the cost increases as the required torque increases. Furthermore, the length of the end-effector must be increased according to the stroke of the linear servo actuator. To address these challenges, the end-effector has been enhanced in this study by adopting a linkage structure, as shown in Fig. 15. This linkage structure converts circular motion into linear motion and allows a typical servo motor to generate linear motion similar to a linear servo actuator. This modification can promote motor miniaturization and yield a high cutting success rate, as indicate in Table 4. Although the proposed end-effector is still not adequately compact, the miniaturization of the linkage structure can help obtain a more compact design. In future work, the proposed end-effector will be miniaturized by using a more compact design.

Second, this study aims to realize robust visual servoing through efficient pedicel detection. The current pedicel detection system uses the FPFH algorithm to search for pedicel cut points, which results in a higher frame rate than that achieved using deep learning. However, the experimental results summarized in Table 4 indicate that detection based on geometry has a low success rate due to obstacles such as leaves and stems. This aspect was noted to be particularly challenging in this study as the target fruit, cucumbers, had a color almost identical to that of the stems and pedicels. To address this issue, a system is being designed that combines the proposed pedicel detection technology with deep learning to achieve fast detection. The proposed pedicel detection technology can extract feature points using FPFH and extract the pedicels by exploiting the curvature differences when the approximate location of the pedicel is known. By implementing deep-learning-based 2D location estimation, the orientation can be estimated, resulting in improved performance during frame drops. In addition, video stabilization techniques are being explored to ensure robustness against shaking by stabilizing motion using the extracted point cloud as a key point. The objective is to minimize shaking by predicting the key point movement through filters. In summary, the combination of the proposed pedicel detection technology, deep learning, and video stabilization is expected to improve the detection accuracy and enable robust visual servoing.

Third, the range of operation of the harvest robot is limited, mainly due to environmental factors. For instance, although the fruits located at higher positions can be successfully detected by the vision system, they remain unreachable due to the workspace constraints. Consequently, a significant portion of potential harvest cannot be accessed by the robot. These workspace limitations predominantly stem from the physical structure of the robotic arm and its range of motion, which fails to extend sufficiently high to reach fruits positioned higher on the plants. This configuration hinders efficient and comprehensive harvesting. A potential solution to this problem is to incorporate a lift mechanism into the robotic system. This adjustment can extend the vertical reach of the robot, enabling access to higher fruits on the plant. By incorporating this lift mechanism, the range of harvest able areas can be significantly expanded, thereby enhancing not only the efficiency of the robot but also its overall yield. Such a robot would be able to execute a more comprehensive and effective harvesting process that aligns better with the actual requirements of the field conditions.

Lastly, this study aims at expanding the applicability of the harvesting robot by enabling the identification of fruits and vegetables such as tomatoes, Korean melons, and paprika. Enhancing the efficiency of the proposed framework will facilitate its use for harvesting other fruits and vegetables. In other words, by enhancing the performance of each component of the human-centered approach, an efficient cucumber-harvesting robot system that can harvest other fruits and vegetables can be developed. These three improvements will be implemented as soon as possible, and their versatility will be confirmed through field experiments involving fruits and vegetables other than cucumbers.

5. Conclusions

This study was aimed at establishing a human-centered approach for an efficient cucumber-harvesting robot system. Harvest ordering, visual servoing, and end-effector-based manipulation were integrated to achieve efficient and stable harvesting. The proposed approach involved the determination of the optimal harvest order, guiding of the end-effector to the cucumber pedicel through visual servoing, and design of an end-effector that can effectively harvest long cucumbers. The performance of the harvesting robot was evaluated through preliminary and field experiments.

The implementation of harvest ordering reduced the harvest time and TL and increased the battery efficiency. Robust visual servoing was achieved through fast pedicel detection at a frame rate of 16–23 FPS using computer vision techniques. The pedicel positioning error was calculated, and the results confirmed that the pedicel could be positioned within the cutting area of the end-effector. The proposed end-effector effectively cut a thin cucumber pedicel (3–6 mm) with a 100% success rate, whereas previous grasping modules failed at grasping cucumbers.

Field experiments were conducted at three sites: GM, SV, and FF. The results showed that the success rates ranged from 50.9% to 60.0% at the three sites, with an overall success rate of 56.6%. The average damage rate for the harvested cucumbers was 4.7%, with seven damaged cucumbers. The precision, recall, accuracy, and F1-score for GM were 0.951, 0.916, 0.875, and 0.933, respectively. The corresponding values for SV were 0.814, 0.777, 0.660, and 0.795, and those for FF were 0.957, 0.927, 0.890, and 0.942. The average harvesting time was 13.2 s and 30.7 s for the first and second approaches, respectively, with an average harvest time of 56.0 s. The positional accuracy of the system was determined to be within the optimal range of 0–30 mm.

Furthermore, failure cases were thoroughly investigated to identify and address the leading causes of cucumber harvesting failures. Through targeted research in the identified directions, the harvesting robot performance can be enhanced, and these efforts are anticipated to promote the development of more efficient harvesting robot systems.

CRediT authorship contribution statement

Yonghyun Park: Conceptualization, Methodology, Writing – original draft, Visualization, Investigation. **Jaehwi Seol:** Conceptualization, Methodology, Writing – original draft, Investigation. **Jeonghyeon Pak:** Conceptualization, Methodology, Writing – original draft, Investigation. **Yuseung Jo:** Conceptualization, Methodology, Writing – original draft, Investigation. **Changjo Kim:** Conceptualization, Methodology, Investigation. **Hyoung Il Son:** Conceptualization, Writing – review & editing, Supervision, Project administration, Funding acquisition.

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

No data was used for the research described in the article.

Acknowledgments

This research was supported by the Korea Institute of Planning and Evaluation for Technology in Food, Agriculture and Forestry (IPET) through the Smart Farm Innovation Technology Development Program funded by the Ministry of Agriculture, Food and Rural Affairs (MAFRA), South Korea (421031-04), and, in part, by Regional Innovation Strategy (RIS) through the National Research Foundation of Korea (NRF) funded by the Ministry of Education (MOE), South Korea (2021RIS-002).

References

Arad, B., Balendonck, J., Barth, R., Ben-Shahar, O., Edan, Y., Hellström, T., Hemming, J., Kurtser, P., Ringdahl, O., Tielen, T., et al., 2020. Development of a sweet pepper harvesting robot. J. Field Robotics 37 (6), 1027–1039. http://dx.doi.org/10.1002/rob.21937.

Bai, Y., Guo, Y., Zhang, Q., Cao, B., Zhang, B., 2022. Multi-network fusion algorithm with transfer learning for green cucumber segmentation and recognition under complex natural environment. Comput. Electron. Agric. 194, 106789. http://dx.doi.org/10.1016/j.compag.2022.106789.

Bai, Y., Zhang, B., Xu, N., Zhou, J., Shi, J., Diao, Z., 2023. Vision-based navigation and guidance for agricultural autonomous vehicles and robots: A review. Comput. Electron. Agric. 205, 107584. http://dx.doi.org/10.1016/j.compag.2022.107584.

Barnett, J., Duke, M., Au, C.K., Lim, S.H., 2020. Work distribution of multiple Cartesian robot arms for kiwifruit harvesting. Comput. Electron. Agric. 169, 105202. http://dx.doi.org/10.1016/j.compag.2019.105202.

Cui, M., Lou, Y., Ge, Y., Wang, K., 2023. LES-YOLO: A lightweight pinecone detection algorithm based on improved YOLOv4-tiny network. Comput. Electron. Agric. 205, 107613. http://dx.doi.org/10.1016/j.compag.2023.107613.

Dewi, T., Risma, P., Oktarina, Y., Muslimin, S., 2018. Visual servoing design and control for agriculture robot; a review. In: 2018 International Conference on Electrical Engineering and Computer Science. ICECOS, IEEE, pp. 57–62. http://dx.doi.org/10.1109/ICECOS.2018.8605209.

Edan, Y., Flash, T., Peiper, U.M., Shmulevich, I., Sarig, Y., 1991. Near-minimum-time task planning for fruit-picking robots. IEEE Trans. Robot. Autom. 7 (1), 48–56. http://dx.doi.org/10.1109/70.68069.

Gao, F., Fang, W., Sun, X., Wu, Z., Zhao, G., Li, G., Li, R., Fu, L., Zhang, Q., 2022a. A novel apple fruit detection and counting methodology based on deep learning and trunk tracking in modern orchard. Comput. Electron. Agric. 197, 107000. http://dx.doi.org/10.1016/j.compag.2022.107000.

Gao, J., Zhang, F., Zhang, J., Yuan, T., Yin, J., Guo, H., Yang, C., 2022b. Development and evaluation of a pneumatic finger-like end-effector for cherry tomato harvesting robot in greenhouse. Comput. Electron. Agric. 197, 106879. http://dx.doi.org/10.1016/j.compag.2022.106879.

Hayashi, S., Shigematsu, K., Yamamoto, S., Kobayashi, K., Kohno, Y., Kamata, J., Kurita, M., 2010. Evaluation of a strawberry-harvesting robot in a field test. Biosyst. Eng. 105 (2), 160–171. http://dx.doi.org/10.1016/j.biosystemseng.2009.09.011.

Hu, G., Chen, C., Chen, J., Sun, L., Sugirbay, A., Chen, Y., Jin, H., Zhang, S., Bu, L., 2022. Simplified 4-DOF manipulator for rapid robotic apple harvesting. Comput. Electron. Agric. 199, 107177. http://dx.doi.org/10.1016/j.compag.2022.107177.

Huang, M., He, L., Choi, D., Pecchia, J., Li, Y., 2021. Picking dynamic analysis for robotic harvesting of agaricus bisporus mushrooms. Comput. Electron. Agric. 185, 106145. http://dx.doi.org/10.1016/j.compag.2021.106145.

Jun, J., Kim, J., Seol, J., Kim, J., Son, H.I., 2021. Towards an efficient tomato harvesting robot: 3D perception, manipulation, and end-effector. IEEE Access 9, 17631–17640. http://dx.doi.org/10.1109/ACCESS.2021.3052240.

Kim, S., Hong, S.J., Ryu, J., Kim, E., Lee, C.H., Kim, G., 2023. Application of amodal segmentation on cucumber segmentation and occlusion recovery. Comput. Electron. Agric. 210, 107847. http://dx.doi.org/10.1016/j.compag.2023.107847.

Kim, J., Pyo, H., Jang, I., Kang, J., Ju, B., Ko, K., 2022. Tomato harvesting robotic system based on deep-ToMaToS: Deep learning network using transformation loss for 6D pose estimation of maturity classified tomatoes with side-stem. Comput. Electron. Agric. 201, 107300. http://dx.doi.org/10.1016/j.compag.2022.107300.

Kurtser, P., Edan, Y., 2020. Planning the sequence of tasks for harvesting robots. Robot. Auton. Syst. 131, 103591. http://dx.doi.org/10.1016/j.robot.2020.103591.

Lee, D., Kim, G., Kim, D., Myung, H., Choi, H.T., 2012. Vision-based object detection and tracking for autonomous navigation of underwater robots. Ocean Eng. 48, 59–68. http://dx.doi.org/10.1016/j.oceaneng.2012.04.006.

Lehnert, C., Tsai, D., Eriksson, A., McCool, C., 2019. 3D move to see: Multi-perspective visual servoing towards the next best view within unstructured and occluded environments. In: 2019 IEEE/RSJ International Conference on Intelligent Robots and Systems. IROS, IEEE, pp. 3890–3897. http://dx.doi.org/10.1109/IROS40897.2019.8967918.

Li, Z., Miao, F., Yang, Z., Wang, H., 2019. An anthropometric study for the anthropomorphic design of tomato-harvesting robots. Comput. Electron. Agric. 163, 104881. http://dx.doi.org/10.1016/j.compag.2019.104881.

Mao, S., Li, Y., Ma, Y., Zhang, B., Zhou, J., Kai Wang, 2020a. Automatic cucumber recognition algorithm for harvesting robots in the natural environment using deep learning and multi-feature fusion. Comput. Electron. Agric. 170, 105254. http://dx.doi.org/10.1016/j.compag.2020.105254.

Mao, S., Li, Y., Ma, Y., Zhang, B., Zhou, J., Wang, K., 2020b. Automatic cucumber recognition algorithm for harvesting robots in the natural environment using deep learning and multi-feature fusion. Comput. Electron. Agric. 170, 105254. http://dx.doi.org/10.1016/j.compag.2020.105254.

Mehta, S., Burks, T., 2014. Vision-based control of robotic manipulator for citrus harvesting. Comput. Electron. Agric. 102, 146–158. http://dx.doi.org/10.1016/j.compag.2014.01.003.

Mehta, S.S., MacKunis, W., Burks, T.F., 2016. Robust visual servo control in the presence of fruit motion for robotic citrus harvesting. Comput. Electron. Agric. 123, 362–375. http://dx.doi.org/10.1016/j.compag.2016.03.007.

Ning, Z., Luo, L., Ding, X., Dong, Z., Yang, B., Cai, J., Chen, W., Lu, Q., 2022. Recognition of sweet peppers and planning the robotic picking sequence in high-density orchards. Comput. Electron. Agric. 196, 106878. http://dx.doi.org/10.1016/j.compag.2022.106878.

Pak, J., Kim, J., Park, Y., Son, H.I., 2022. Field evaluation of path-planning algorithms for autonomous mobile robot in smart farms. IEEE Access 10, 60253–60266. http://dx.doi.org/10.1109/ACCESS.2022.3181131.

Park, Y., Seol, J., Pak, J., Jo, Y., Jun, J., Son, H.I., 2022a. A novel end-effector for a fruit and vegetable harvesting robot: mechanism and field experiment. Precis. Agric. 1–23. http://dx.doi.org/10.1007/s11119-022-09981-5.

Park, Y., Seol, J., Pak, J., Jo, Y., Kim, C., Son, H.I., 2022b. Human-centered robot system for autonomous cucumber harvesting: Preliminary result. In: International Conference on Control, Automation and Systems. ICCAS, IEEE, pp. 1689–1690.

Rong, J., Wang, P., Wang, T., Hu, L., Yuan, T., 2022. Fruit pose recognition and directional orderly grasping strategies for tomato harvesting robots. Comput. Electron. Agric. 202, 107430. http://dx.doi.org/10.1016/j.compag.2022.107430.

Roshanianfard, A., Noguchi, N., 2020. Pumpkin harvesting robotic end-effector. Comput. Electron. Agric. 174, 105503. http://dx.doi.org/10.1016/j.compag.2020.105503.

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. http://dx.doi.org/10.1016/j.compag.2020.105933.

Van Henten, E.J., Hemming, J., Van Tuijl, B., Kornet, J., Meuleman, J., Bontsema, J., Van Os, E., 2002. An autonomous robot for harvesting cucumbers in greenhouses. Auton. Robots 13 (3), 241–258. http://dx.doi.org/10.1023/A:1020568125418.

Van Henten, E., Van Tuijl, B.v., Hemming, J., Kornet, J., Bontsema, J., Van Os, E., 2003. Field test of an autonomous cucumber picking robot. Biosyst. Eng. 86 (3), 305–313. http://dx.doi.org/10.1016/j.biosystemseng.2003.08.002.

Van Henten, E., Van't Slot, D., Hol, C., Van Willigenburg, L., 2009. Optimal manipulator design for a cucumber harvesting robot. Comput. Electron. Agric. 65 (2), 247–257. http://dx.doi.org/10.1016/j.compag.2008.11.004.

Wang, X., Kang, H., Zhou, H., Au, W., Chen, C., 2022. Geometry-aware fruit grasping estimation for robotic harvesting in apple orchards. Comput. Electron. Agric. 193, 106716. http://dx.doi.org/10.1016/j.compag.2022.106716.

Williams, H.A., Jones, M.H., Nejati, M., Seabright, M.J., Bell, J., Penhall, N.D., Barnett, J.J., Duke, M.D., Scarfe, A.J., Ahn, H.S., et al., 2019. Robotic kiwifruit harvesting using machine vision, convolutional neural networks, and robotic arms. Biosyst. Eng. 181, 140–156. http://dx.doi.org/10.1016/j.biosystemseng.2019.03.007.

Xiong, Y., Ge, Y., Grimstad, L., From, P.J., 2020. An autonomous strawberry-harvesting robot: Design, development, integration, and field evaluation. J. Field Robotics 37 (2), 202–224. http://dx.doi.org/10.1002/rob.21889.

Xiong, Y., Peng, C., Grimstad, L., From, P.J., Isler, V., 2019. Development and field evaluation of a strawberry harvesting robot with a cable-driven gripper. Comput. Electron. Agric. 157, 392–402. http://dx.doi.org/10.1016/j.compag.2019.01.009.

Yu, Y., Zhang, K., Yang, L., Zhang, D., 2019. Fruit detection for strawberry harvesting robot in non-structural environment based on mask-RCNN. Comput. Electron. Agric. 163, 104846. http://dx.doi.org/10.1016/j.compag.2019.06.001.

Zhang, W., Liu, Y., Chen, K., Li, H., Duan, Y., Wu, W., Shi, Y., Guo, W., 2021. Lightweight fruit-detection algorithm for edge computing applications. Front. Plant Sci. 12, 740936. http://dx.doi.org/10.3389/fpls.2021.740936.