IAES International Journal of Robotics and Automation (IJRA)
Vol. 15, No. 2, June 2026, pp. 353~364
ISSN: 2722-2586, DOI: 10.11591/ijra.v15i2.pp353-364

Design of beefsteak tomato harvesting robot system in greenhouse

Thien An Dinh, So Nam Phung, Tri Cong Phung
Faculty of Mechanical Engineering, Ho Chi Minh University of Technology (HCMUT), Ho Chi Minh City, Vietnam
Vietnam National University Ho Chi Minh City, Linh Xuan Ward, Ho Chi Minh City, Vietnam

Article Info

Article history:

Received Aug 24, 2025
Revised Feb 25, 2026
Accepted May 16, 2026

Keywords:

Beefsteak tomato
Greenhouse
Harvesting robot
Motion simulation
Robot kinematics

ABSTRACT

One challenge for tomato harvesting robots is that some of the tomato stems were not detectable because they were hidden behind the leaves or other obstacles. The primary objective of this research is to design, simulate, and experiment with a tomato harvesting robot and propose an improved detection algorithm to overcome the above problem. The suggested detection algorithm is designed to first detect the tomato fruit itself, and if the stem is not visible, the system will automatically adjust the camera's viewing angle to provide a better perspective and uncover the hidden stem. Simulation and experimental tests were carried out in a real tomato greenhouse to evaluate the cutting and holding mechanism, as well as the camera-based detection algorithm. These experimental results confirmed the effectiveness of the gripper and detection system and revealed several challenges in the harvesting algorithm. By integrating advanced algorithms for tomato detection and harvesting, this robot will reduce damage to the tomatoes, ensuring higher quality and yield.

Corresponding Author:

Tri Cong Phung
Faculty of Mechanical Engineering, Ho Chi Minh University of Technology (HCMUT)
268 Ly Thuong Kiet Street, Dien Hong Ward, Ho Chi Minh City, Vietnam
Email: ptcong@hcmut.edu.vn

1. INTRODUCTION

Beefsteak tomatoes are among the largest varieties of cultivated tomatoes in Vietnam. They typically weigh between 250 to 300 grams, and their diameter is around 10 to 15 centimeters. In greenhouse hydroponic systems, where plants may be suspended or trained along hanging systems, the fruits may form around 50 centimeters to 1 meter above the ground. To ensure the fruits have enough space to grow, each cluster is limited from 2 to 3 fruits and is neatly pruned. The manual harvesting method is the traditional method of harvesting. Tomatoes are hand-picked by cutting the stem or pulling the fruit off the plant. However, this method is time-consuming, labor-intensive, and costly. The modern method is using tomatoes harvesting robots. Robots are equipped with sensors and cutting mechanisms to detect and pick tomatoes autonomously. Robots also can be programmed to move through the field, identify ripe tomatoes, and harvest them. This method has advantages of increasing automation, saving labor costs, reducing errors, and less damage to the tomato plant. However, there are many challenges for automation of robot harvesting.

The end-effector, an important component of harvesting robots, is designed depending on fruit detachment methods. Theres are various available methods such as gripping and pulling, gripping and spinning, gripping and cutting, and detaching using vacuums [1]. The gripping and pulling methods are the simple method for detaching the tomato from the stem. However, the above method can damage the fruit because of the force of gripping [2]. To overcome this disadvantage, a perceptual softened effector is designed to reduce the effect of gripping force to the tomato [3]. Another potential solution of tomato detachment method to reduce the applied force is gripping and spinning method. The end-effector was designed to separate the fruit from the stalk based on the linear motion of the constraint part and the rotating gripper [4]. The design of end-effectors with a sucking function can reduce the damage of the fruit compared with method using grasping force. An end-effector will absorb fruits directly by using a vacuum suction cup to generate negative pressures [5]. Compared to mechanical arms, the picking efficiency of such a robot using vacuum technology is higher. Gao *et al.* [6] developed a pneumatic finger-like end-effector for cherry tomato harvesting robot. The end-effector had the capability of picking cherry tomatoes continuously and steadily using pneumatic power.

These above methods have weak points, which is impacting the force directly on the tomato and can damage the fruit. Thus, the gripping and cutting method is suggested because this method uses the tool to cut the peduncle of the fruit, do not affect the force directly to the fruit [7], [8]. Park *et al.* [9] proposed a novel end-effector by integrating the cutting module, the suction module, and the transporting module into one tool. This end-effector can harvest a variety of fruits and vegetables, reduce harvesting time, and improve productivity. Yeshmukhametov [10] designed a semi-spherical gripper tool with cutting blades on the edges of the cup for grasping and detaching tomato. A consecutive procedure was built for harvesting tomatoes. This paper concentrates on designing the tool for gripping and cutting method.

The second problem that needs to be considered in the harvesting robot system is tomato detection algorithms using vision sensors. Machine learning and deep learning are methods often used in object detection. Yuan [11] proposed an improved ResNet34 model (ESA-ResNet34) for crop pest and disease detection. The model employs ResNet34 as its backbone and introduces an efficient spatial attention mechanism (effective spatial attention, ESA) to focus on key regions of the images. Wu *et al.* [12] proposed a decomposed uncertainty guided matting algorithm, which explores the explicitly decomposed uncertainties to efficiently and effectively improve the results. Zhang *et al.* [13] proposed a progressive sample selection framework with contrastive loss for noisy labels. This framework operates in two stages, using robust and contrastive losses to augment the robustness of the model. Wang [14] designed an improved Faster R-CNN model named MatDet for tomato maturity detection. This model was built based on three small models including ResNet-50, RoIAlign, and Path Aggregation Network to improve the robustness and accuracy of the proposed method. Afonso [15] suggested a MaskRCNN algorithm to detect tomatoes and the pixels corresponding to this object based on images taken in a greenhouse.

One advanced method for detecting tomatoes is using YOLO model because of fast detection ability. A tomato detection model RC-YOLOv4 is constructed based on a new backbone network R-CSPDarknet53 by fusing the residual neural network to establish the jump connection between the front and back layers to improve the detection accuracy of tomato in natural environment [16]. Ge *et al.* [17] proposed a visual object tracking network based on the YOLOv5s model to identify and count tomatoes in different growth periods for prediction of tomato cultivation. A lightweight real-time tomato detection and point-picking integrated network model based on YOLOv5 (TDPPL-Net) was proposed [18]. Zheng [19] suggested method including depth filtering and an inter-frame prediction algorithm based on YOLOv8 model to address key challenges such as background interference, occlusion, and double counting. For harvesting tomatoes, the robot system needs to recognize ripe fruit and unripe fruit. Qi [20] proposed an enhanced algorithm to simultaneously estimate the count and ripeness levels for cherry tomato bunches. Ripe values were determined by the proportional relationship between the number of ripe and unripe fruits in a tomato bunch.

Stem cutting offers a viable alternative that minimizes the risk of damaging the harvested produce. However, it presents a more significant challenge. Identifying and cutting a small branch can be more difficult in real-world field settings due to dense foliage and obstructions than capturing the entire fruit. Rong *et al.* [21] proposed for peduncle cutting point localization and pose estimation. Images captured in real time at a fixed long-distance were detected using the YOLOv4-Tiny detector with a precision of 92.7% and a detection speed of 0.0091 second per frame. However, due to the complex natural environment and tracking stability, there are still considerable challenges for automated yield estimation to be deployed in practice. Therefore, Rong *et al.* [22] presented an improved tomato cluster counting method that combines object detection, multiple object tracking, and specific tracking region counting. To reduce background tomato misidentification, the YOLOv5-4D was proposed that fuses RGB images and depth images as input. One of the challenges when harvesting tomato using robot systems is that it needs to estimate the ability of harvesting. Fujinaga *et al.* [23] proposed a method for evaluating the tomato fruit harvestability using a tomato harvesting robot. In this paper, we propose a new algorithm to detect the hidden peduncle for gripping and cutting method.

A total harvesting robot system often includes robot mechanisms, end-effector tools, and vision algorithms. There are several research projects about designing tomato harvesting robots. The simulation of a tomato harvesting robot in ROS had been successfully implemented, integrating multiple components for a complete system [24]. In Gazebo, a virtual farm environment with tomato plants has been created for realistic testing. Singh *et al.* [25] developed a deep learning-powered robotic solution for strawberry detection and grasping in a simulated digital twin greenhouse environment based on YOLOv9-GLEAN deep learning model. This study was verified by simulation in ROS-Gazebo.

Besides researching simulation of harvesting robots, there are a few research experiments about harvesting tomato robots. Jun *et al.* [26] proposed an efficient tomato-harvesting robot that combines the principles of 3D perception, manipulation, and an end-effector which includes two parts, a grasping module and a cutting module. The data from sensors was used for building 3D environment to control the motion of the robot. Oktarina *et al.* [27] discussed the pilot project of employing robots as a harvesting robot. A small robot model was used to detect and harvest the tomatoes in the environment of the laboratory. These experimental results verify the proposed method and could be developed for real work. Rong *et al.* [28] suggested a harvesting robot capable of picking tomato clusters by cutting their fruit-bearing pedicels. The robot prototype was used to evaluate the working in real greenhouse environments.

However, in the above research, they just tested in lab environment, not real tomato trees in greenhouse. Another disadvantage is that the fixed base robot cannot move to another position to harvest another tomato. To move and collect tomatoes from trees in greenhouse, it needs the mobile platform and put the manipulator on this platform [29]. Many path planning and obstacle avoidance algorithms need to be considered [30], [31]. These methods have weak points when moving in greenhouse. To overcome this problem, it needs to design mobile platform as a track rail. A novel bot named AHPPEBot is designed for autonomous harvesting tomatoes [32]. The proposed method was based on crop phenotyping, pose estimation and a multi-task YOLOv5 model. In this paper, we use this method for harvesting tomato robots including a mobile platform and a manipulator.

This paper designs a beefsteak tomato harvesting robot in medium and small greenhouse. The system is designed as a complete modular structure, consisting of the cutting and holding tomato cluster mechanism, a manipulator system, a tomato detecting system, and the AGV system. The first contribution is that the authors present the new algorithm to first detect the tomato fruit itself, and if the stem is not visible, the system will automatically adjust the camera's viewing angle to provide a better perspective and uncover the hidden stem. The second contribution of this paper is that the author makes simulation of motion of the robot and the AGV platform to look for the optimal path for the robot system. From the simulation results, the author proposes suitable paths for control the robot. This enhancement ensures that the robot can accurately identify and locate both the tomatoes and their stems, improving the overall efficiency and reliability of the harvesting process. In this paper, section 1 shows the literature review about tomatoes harvesting robot and the contribution of the research. Section 2 shows the design of the tomatoes harvesting robot and the recognition algorithm. Section 3 shows the simulation and experimental results. Section 4 shows the conclusion of the paper.

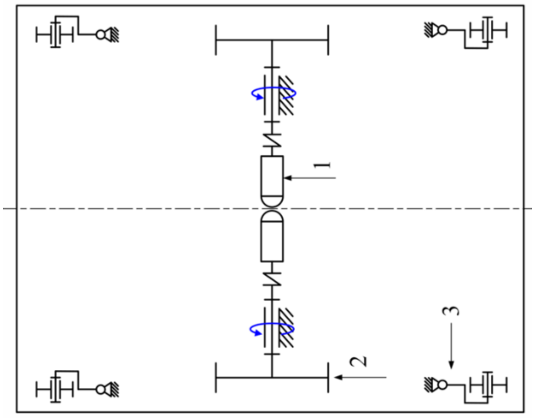
2. METHOD

The proposed harvesting tomato robot system includes the mobile platform, the manipulator, the cutting-tool, and the vision system. This part firstly presents the design of the tomato harvesting robot system. Next, the kinematic modeling of the robot is described. Finally, the detection algorithm of the tomato stem is proposed.

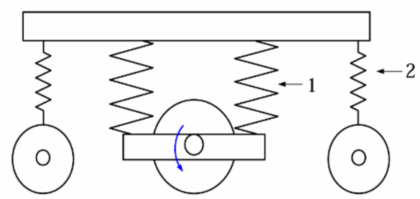
2.1. Design of mechanical parts

Figure 1 shows the design of the mobile platform. Figure 1(a) illustrates the schematic of the AGV wheel system. The input (1) is the drive motor, which drives the two wheels (2) through the motor, while the system is supported by four caster wheels (3). PU driving wheels are ideal for heavy-duty applications, offering superior wear resistance, high load capacity, and excellent traction. These wheels are widely used in industries such as automotive manufacturing, logistics, and warehousing, where Automated Guided Vehicles (AGVs) operate under demanding conditions. Rubber driving wheels provide outstanding shock absorption, traction, and grip on various surfaces. However, their wear resistance is generally lower than PU, and they are more vulnerable to damage from harsh chemicals and solvents. In this paper, PU wheels are chosen due to their abrasion resistance, suitability for movement on concrete surfaces, and low maintenance frequency. Figure 1(b) illustrates the schematic of the AGV suspension system, which consists of (1) the shock absorber of the driven wheel, and (2) the shock absorber of the caster wheel.

Figure 2 shows the design of the manipulator part. Figure 2(a) illustrates the schematic of the SCARA robot, which consists of (1) the actuator of the ball screw system (2), and (4) the actuator of the timing belt system (3). Figure 2(b) illustrates the Mantis gripper mechanism, which consists of (1) the actuator of the gripper, and gears Z1, Z2, Z2', and Z3 within the mechanism. A Mantis Gripper refers to a type of robotic gripper inspired by the structure and functionality of a praying mantis' forelegs. These grippers are designed to mimic the precise and efficient way that the praying mantis captures and holds its prey, using an arrangement of sharp, articulated segments that can adapt to various shapes and sizes of objects. Figure 3 shows the full design in 3D.

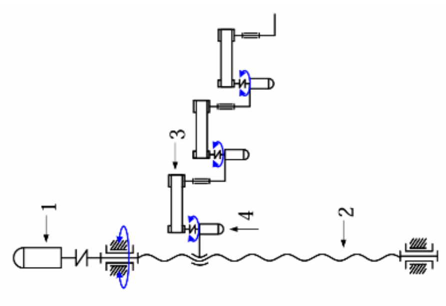


(a)

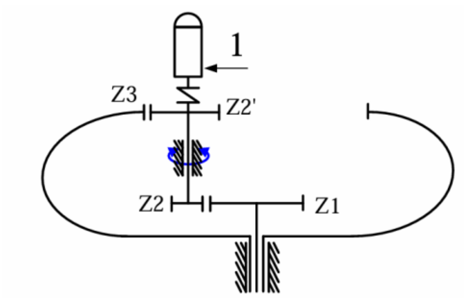


(b)

Figure 1. Design of mobile platform: schematic of (a) AGV wheel system and (b) the suspension system



(a)



(b)

Figure 2. Design of manipulator part (a) schematic of the SCARA robot and (b) schematic of Mantis gripper mechanism

2.2. Design of electrical and electronic parts

For high-level motion control and interfacing with advanced sensors such as cameras, we will use an embedded computer. The Raspberry Pi 4 Model B is selected as a compact and cost-effective embedded platform, equipped with a Quad-core CortexA72 processor, up to 8GB of RAM, and multiple USB ports for peripheral connections. It supports CSI camera modules, making it suitable for basic computer vision applications. While it does not offer the same GPU performance as industrial platforms like the Jetson series, the Raspberry Pi 4 can handle moderate image processing tasks and can run a full Linux OS, enabling integration with a wide range of development libraries and communication protocols such as I2C, SPI, UART, and CAN (via HATs). Its versatility and community support make it a strong choice for prototyping and educational robotics projects.

3D model of robot Model with attached coordinate frames and dimensions

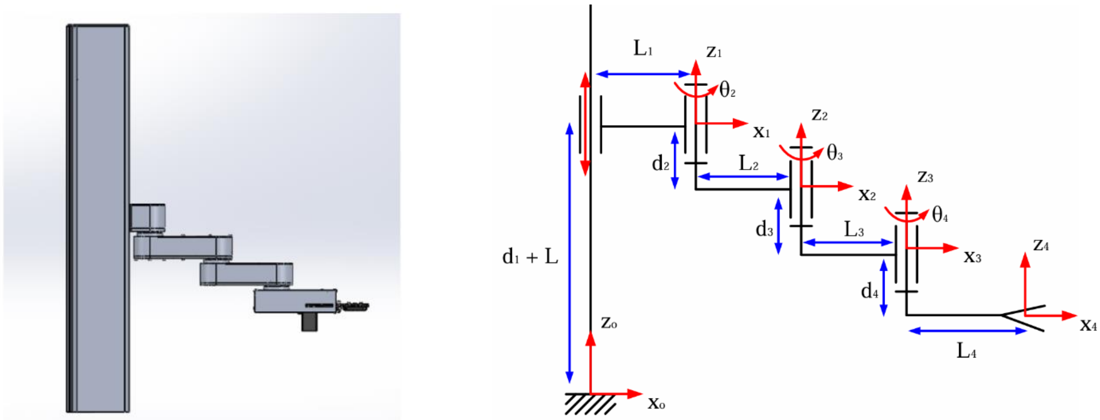


Figure 3. The model of the 4 DOF SCARA Robot

The computer runs the Linux kernel with PREEMPT-RT, enabling pseudo real time code execution required for precise timing in control and detection tasks. In the context of tomato detection, the system supports integration with computer vision libraries such as OpenCV for detecting tomatoes and sending appropriate commands to the robot arm for harvesting. The Linux-based environment also provides flexibility for customizing detection algorithms and optimizing performance for agricultural conditions.

2.3. Kinematic modeling of SCARA robot

The kinematic modeling of the SCARA robot is presented in this section. From Figure 3, the DH table is established. From that, the transformation matrix between two adjacent frames is calculated. Next, the transformation matrix from end-effector frame B_4 to base frame B_0 is calculated. Finally, the inverse kinematic of the robot can be solved. The DH parameters of the SCARA robot are shown in Table 1.

Table 1. DH parameters of SCARA robot

i	a_i (mm)	α_i (degree)	d_i (mm)	θ_i (degree)
1	L_1	0	$L+d_1$	0
2	L_2	0	$-d_2$	θ_2
3	L_3	0	$-d_3$	θ_3
4	L_4	0	$-d_4$	θ_4

The transformation matrix between two adjacent frames can be calculated and shown in (1) and (2).

$$ {}^{0}T_1 = \begin{bmatrix} 1 & 0 & 0 & L_1 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & d_1 + L \\ 0 & 0 & 0 & 1 \end{bmatrix}; \quad {}^{1}T_2 = \begin{bmatrix} \cos\theta_2 & -\sin\theta_2 & 0 & L_2\cos\theta_2 \\ \sin\theta_2 & \cos\theta_2 & 0 & L_2\sin\theta_2 \\ 0 & 0 & 1 & -d_2 \\ 0 & 0 & 0 & 1 \end{bmatrix} \tag{1} $$

$$ {}^{2}T_3 = \begin{bmatrix} \cos\theta_3 & -\sin\theta_3 & 0 & L_3\cos\theta_3 \\ \sin\theta_3 & \cos\theta_3 & 0 & L_3\sin\theta_3 \\ 0 & 0 & 1 & -d_3 \\ 0 & 0 & 0 & 1 \end{bmatrix}; \quad {}^{3}T_4 = \begin{bmatrix} \cos\theta_4 & -\sin\theta_4 & 0 & L_4\cos\theta_4 \\ \sin\theta_4 & \cos\theta_4 & 0 & L_4\sin\theta_4 \\ 0 & 0 & 1 & -d_4 \\ 0 & 0 & 0 & 1 \end{bmatrix} \tag{2} $$

The transformation matrix from end-effector frame B_4 to base frame B_0 can be calculated from these above transformation matrices and is shown in (3).

$$ {}^{0}T_4 = {}^{0}T_1\, {}^{1}T_2\, {}^{2}T_3\, {}^{3}T_4 = \begin{bmatrix} \cos(\theta_2+\theta_3+\theta_4) & -\sin(\theta_2+\theta_3+\theta_4) & 0 & p_x \\ \sin(\theta_2+\theta_3+\theta_4) & \cos(\theta_2+\theta_3+\theta_4) & 0 & p_y \\ 0 & 0 & 1 & p_z \\ 0 & 0 & 0 & 1 \end{bmatrix} \tag{3} $$

Here, p_x, p_y, and p_z are shown in (4), (5), and (6).

$$p_x = L_1 + L_2 \cos\theta_2 + L_3 \cos(\theta_2 + \theta_3) + L_4 \cos(\theta_2 + \theta_3 + \theta_4) \tag{4}$$

$$p_y = L_2 \sin\theta_2 + L_3 \sin(\theta_2 + \theta_3) + L_4 \sin(\theta_2 + \theta_3 + \theta_4) \tag{5}$$

$$p_z = L + d_1 - d_2 - d_3 - d_4 \tag{6}$$

For inverse kinematic, we need to solve for $d_1, \theta_2, \theta_3, \theta_4$ from (4), (5), and (6).

The first joint d_1 can be calculated from (6) and shown in (7).

$$d_1 = p_z - L + d_2 + d_3 + d_4 \tag{7}$$

The direction, the angle α, between the end-effector (X_4-axis) of the robot and the X_0-axis is shown in (8).

$$\alpha = \theta_2 + \theta_3 + \theta_4 \tag{8}$$

Two new equations show the relationship between θ_2 and θ_3 can be received from (4) and (5) and shown in (9) and (10).

$$p_x - L_1 - L_4 \cos\alpha = A = L_2 \cos\theta_2 + L_3 \cos(\theta_2 + \theta_3) \tag{9}$$

$$p_y - L_4 \sin\alpha = B = L_2 \sin\theta_2 + L_3 \sin(\theta_2 + \theta_3) \tag{10}$$

Summing the square of each element in two sides of (9) and (10), we receive the equation that depends only on the unknown θ_3, shown in (11).

$$A^2 + B^2 = L_2^2 + L_3^2 + 2L_2L_3 \cos\theta_3 \tag{11}$$

Thus, θ_3 can be calculated from (11). Replace the value of θ_3 in (9) and (10), we can calculate θ_2. Finally, θ_4 also can be determined from (8) with known value of θ_2 and θ_3.

2.4. Proposed algorithm of detecting the tomato stem

The proposed algorithm keeps the stem when picking tomatoes instead of cutting it off because the stem helps the tomato stay fresh longer. It was discovered that some of the tomato stems were not detectable because they were hidden behind the leaves or other obstacles. To address this issue, the team developed an improved detection algorithm. This new algorithm is designed to first detect the tomato fruit itself, and if the stem is not visible, the system will automatically adjust the camera's viewing angle to provide a better perspective and uncover the hidden stem. Figure 4 shows the improved detection of stem algorithms when the stem is hidden by the obstacles. This enhancement ensures that the robot can accurately identify and locate both the tomatoes and their stems, improving the overall efficiency and reliability of the harvesting process.

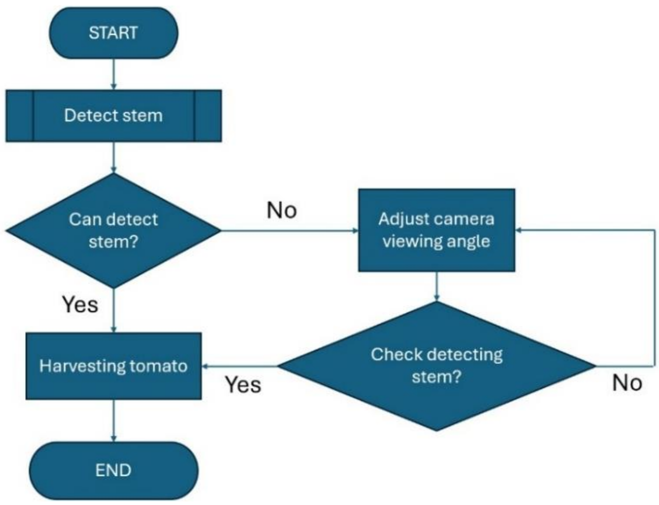


Figure 4. The proposed detection of the tomato stem algorithm

Next, the authors decided to collect 600 images of tomato stems as data for training. The data are collected via images on the internet and photos of the tomatoes the team bought and took the picture. Then, the authors used MakeSenseAI to label the data of tomato stems. When labeling, the authors adjusted the center of the bounding box coincide with the point where the end-effector of SCARA Robot will cut off when picking tomatoes. After that, the authors used GoogleColab and YOLOv8 for training data with 200 epochs. The hardware used is intel core i7 with 16 GB memory. The software used is YOLOv8 and Python v3.9. Several key hyper-parameter settings are as follows: learning_rates=[0.0001, 0.001, 0.01], batch_sizes=[32,64,128], optimizers=[‘adam’, ‘sgd’], epochs=[100, 200, 300]. The YOLOv8 is used for suitable with the current architecture of the hardware and software.

3. RESULTS AND DISCUSSION

3.1. Simulation results

Firstly, MATLAB program is used to simulate the operation of the SCARA robot. The objectives are calculating and simulating the forward and inverse kinematics of the SCARA robot, simulating the trajectory planning and motion of the SCARA robot, identifying possible configurations of the SCARA robot during key tasks, evaluating whether the robot can operate effectively within the required working space, and ensuring it can reach all necessary positions during operation. There are two positions of the SCARA robot that is needed to check are the position for scanning the tomato stems (SCAN pose) and the position for placing the tomatoes into the container (PLACE pose). Figure 5 showed respectively the configuration of SCARA Robot in SCAN pose and PLACE pose.

After scanning the tomato stem, the robot needs to move and control the cutting tool to the position of the stem. There are many directions to control the cutting tool depending on the value of the angle α. This direction angle affects the working ability of the robot, especially affecting the cycle time and the obstacle avoidance ability of robots. Assuming that the rotational velocity of each joint is the same, about 5 degrees per second. The calculation about the working cycle time for each direction angle are shown in Table 2. For this calculation, the values of SCARA robot parameters are as follows: L_1=200 mm, L_2=300 mm, L_3=300 mm, L_4=200 mm, d_2=50 mm, d_3=50 mm, d_4=50 mm.

SCAN pose PLACE pose

Figure 5. Two possible positions of SCARA Robot)

Table 2. The working cycle time for different direction angles of cutting tool

No.	α (degrees)	θ_2 (degrees)	θ_3 (degrees)	θ_4 (degrees)	Cycle time (seconds)
1	200	64.6633	33.0843	102.2524	20.45
2	190	54.6467	50.1513	85.2020	17.04
3	**180**	**46.8840**	**63.6122**	**69.5038**	**13.90**
4	170	40.6113	75.2336	54.1550	15.05
5	160	35.7153	85.5365	38.7482	17.11
6	150	32.2961	94.6655	23.0383	18.93
7	140	30.5587	102.5783	6.8630	20.52
8	130	30.7389	109.1071	-9.8460	21.82
9	120	33.0013	113.9974	-26.9987	22.80
10	110	37.2991	116.9647	-44.2638	23.39
11	100	43.2626	117.7848	-61.0475	23.56
12	90	50.2410	116.3878	-76.6288	23.28
13	80	57.5243	112.8908	-90.4151	22.58
14	70	64.5817	107.5436	-102.1253	21.51
15	60	71.1516	100.6304	-111.7820	22.36
16	50	77.1938	92.3872	-119.5810	23.92
17	40	82.8056	82.9526	-125.7582	25.15
18	30	88.1694	72.3284	-130.4978	20.10
19	20	93.5594	60.2996	-133.8590	26.77
20	10	99.4653	46.1661	-135.6313	27.13
21	0	107.3307	27.2660	-134.5968	26.92

From the Table 2, it can conclude that the working cycle time is smallest when the direction angle is 180^0. With this direction angle, the cycle time is 13.90 seconds. After that, the obstacle avoidance ability of the robot also was checked with several direction angles and shown in Figure 6. It can be concluded that the obstacle avoidance ability of the robot when the direction angle of the tool is 180^0 is better than other angles, including 150^0, 120^0, and 90^0.

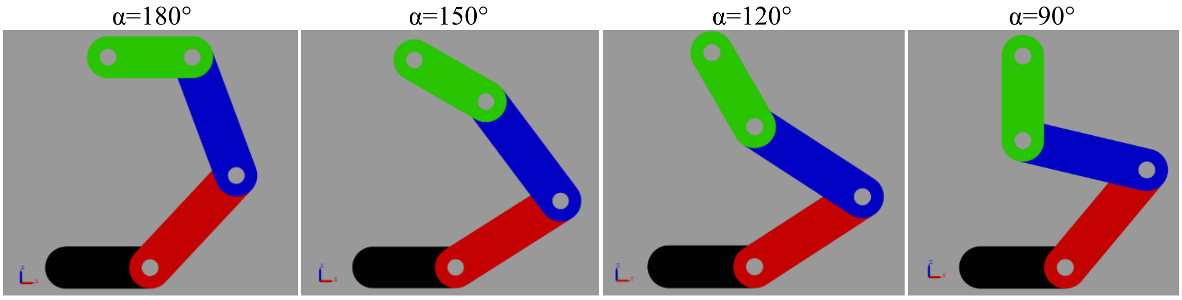


Figure 6. The obstacle avoidance ability of the robot for different direction angles of the cutting tool

Next, Coppelia Sim program is used for simulating the entire tomato harvesting process of the robotic system in a greenhouse environment. The simulation visualizes the complete flowchart of the tomato harvesting procedure, including detecting the tomato stem, determining the optimal trajectory and configuration for the SCARA robotic arm to maximize its operational efficiency in 3D space, picking the tomatoes, and placing tomatoes into the container. Figure 7 showed the motion of the robot to the picking position and showed task of picking the tomato stem. Figure 8 shows the motion of the robot to transfer the tomato to the container and shows the task of placing the tomato into the container.

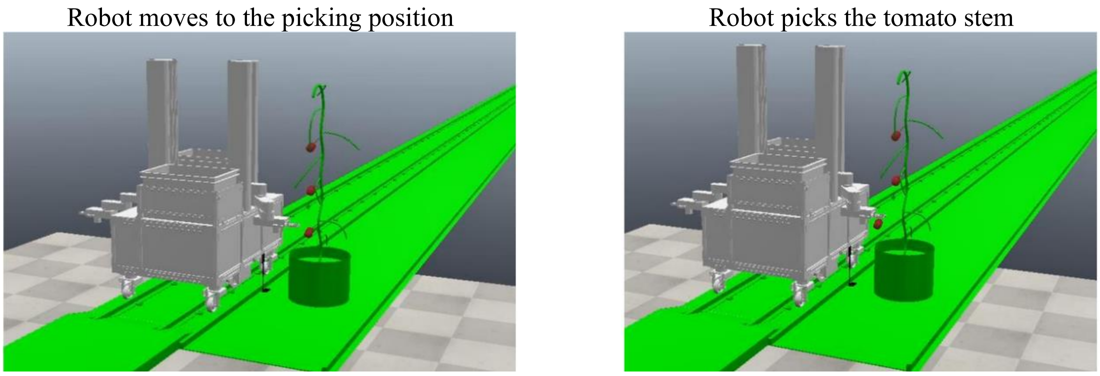


Figure 7. The SCARA robot moves to the picking position and picks the tomato stem

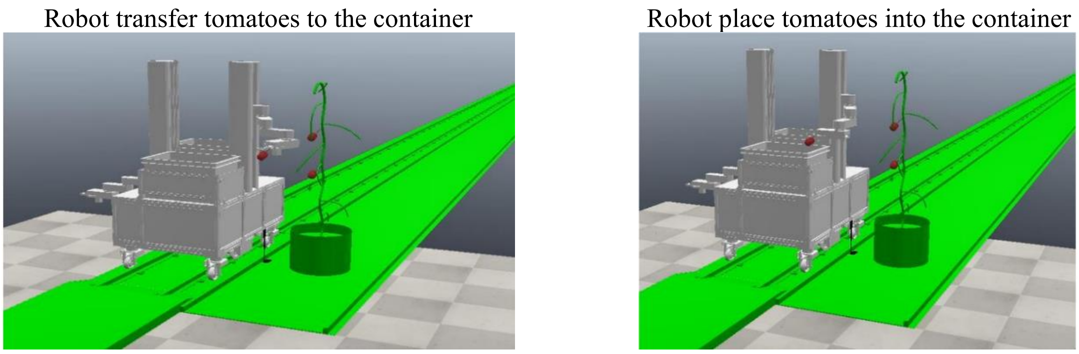


Figure 8. The SCARA robot transfers and places tomatoes to container

3.2. Experimental results

The objective of this experiment is to develop and evaluate a robotic system capable of performing tomato harvesting tasks, including cutting and securely holding tomato stems without causing damage to the fruit and detecting ripe tomatoes in real-world conditions and determining their precise 3D coordinates to enable accurate robotic arm movement and harvesting.

Initially, the research group tested the cutting capability of the robot to validate the calculations for cutting force and to ensure that the selected motor and blade were appropriate for the task. The robot is designed to hold the tomato stem securely using a gripping mechanism, and the experiment demonstrated that the holding mechanism works effectively in practice. The cutting test successfully verified that the system could handle the tomatoes, proving that both the force calculations and the chosen components perform as expected. Figures 9 and 10 showed the experimental results of before cutting and after cutting the stem.

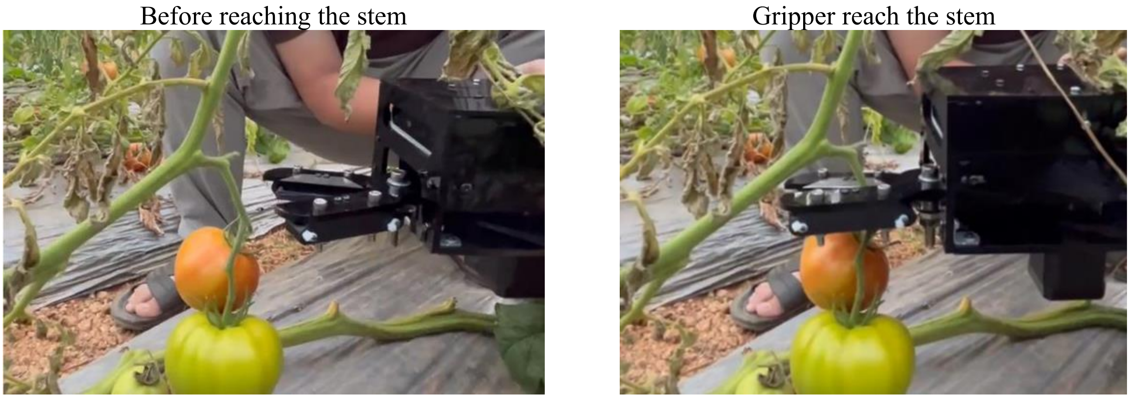


Figure 9. Experiment before cutting the stem

Figure 10. Experiment after cutting the stem

After successfully testing the cutting and holding of the tomatoes, the next phase of the experiment focused on testing the robot's ability to detect the tomatoes and determine their coordinates using the camera system. The camera was calibrated to capture the position of the tomatoes in real-time, and the robot used this data to calculate their exact 3D coordinates. This phase of the experiment aimed to verify the accuracy and reliability of the camera system in real-world conditions, ensuring that the robot could precisely identify and locate the tomatoes for harvesting. Figure 11 showed the experiment for detecting algorithms of tomato stem. The 3D coordinates of tomato stem are X=53.4 mm, Y=2.3 mm, Z=353.6 mm for experiment 1 and X=8.4 mm, Y=-36.9 mm, Z=337.8 mm for experiment 2. The authors used a file with 150 untrained images and Python code for testing the results. After testing 150 images, the training model can detect exactly 147/150 images, with the detection accuracy of 98%. For tomato detection, the accuracy is 100% for 150 images.

Figure 11. Experiment for detecting algorithm of tomato stem

3.3. Discussion

In this section, the number of epochs is discussed. In this research, the authors compare the performance plots when training 600 images with 100 epochs, 200 epochs and 300 epochs. Comparing the performance plots of models trained between 100 and 200 epochs, the authors found that increasing from 100 to 200 epochs improved the model's training effectiveness. Specifically, the loss decreased more significantly compared to the model trained with 200 epochs, and the precision, the recall and mAP remained high. Comparing the performance plots of models trained between 200 and 300 epochs, the authors found that increasing from 200 to 300 epochs was not as effective as 200 epochs because the loss stopped decreasing and tended to increase although the precision, the recall and mAP still high. When tested the file with 150 untrained images, the case with 600 images and 200 epochs achieved an accuracy of up to 98%, while the other cases resulted in an accuracy below 94.7% (142/150 images). Moreover, when tested with living detecting, the model trained with 600 images and 200 epochs also performed the best, even when rotating at various angles or moving.

4. CONCLUSION

This paper has successfully designed and implemented a mobile manipulator robot for beefsteak tomato harvesting in greenhouses. The system integrates mechanical, electrical, and control components, including a SCARA robotic arm and an AGV platform, to operate efficiently in confined greenhouse environments. To validate the design, simulations were conducted using MATLAB-Simulink and Coppelia Sim. These simulations helped verify the robot's performance and visualize its motion and control algorithms under various conditions, providing important insights before physical implementation. In addition, experimental tests were carried out in a real tomato greenhouse to evaluate the cutting and holding mechanism, as well as the camera-based detection algorithm. The detection accuracy of tomatoes is 100% and 98% for detecting the tomato stems. These experiments confirmed the effectiveness of the gripper and detection system and revealed several challenges in the harvesting algorithm. From these findings, it was able to propose improvements that enhance the system's overall reliability and performance.

ACKNOWLEDGMENTS

We acknowledge the support of time and facilities from Ho Chi Minh City University of Technology (HCMUT), VNU-HCM for this study.

FUNDING INFORMATION

This research is funded by Vietnam National University HoChiMinh City (VNU-HCM) under grant number DM2024-20-04.

AUTHOR CONTRIBUTIONS STATEMENT

This journal uses the Contributor Roles Taxonomy (CRediT) to recognize individual author contributions, reduce authorship disputes, and facilitate collaboration.

Name of Author	C	M	So	Va	Fo	I	R	D	O	E	Vi	Su	P	Fu
Thien An Dinh	✓	✓	✓	✓	✓	✓		✓	✓	✓				
So Nam Phung		✓				✓		✓	✓	✓	✓			
Tri Cong Phung	✓		✓	✓			✓			✓	✓	✓	✓	✓

C : **C**onceptualization
M : **M**ethodology
So : **So**ftware
Va : **Va**lidation
Fo : **Fo**rmal analysis
I : **I**nvestigation
R : **R**esources
D : **D**ata Curation
O : Writing - **O**riginal Draft
E : Writing - Review & **E**diting
Vi : **Vi**sualization
Su : **Su**pervision
P : **P**roject administration
Fu : **Fu**nding acquisition

CONFLICT OF INTEREST STATEMENT

Authors state no conflict of interest.

DATA AVAILABILITY

The data that supports the findings of this study are available from the corresponding author, [Tri Cong Phung], upon reasonable request.

REFERENCES

[1] E. Vrochidou, V. N. Tsakalidou, I. Kalathas, T. Gkrimpizis, T. Pachidis, and V. G. Kaburlasos, "An overview of end effectors in agricultural robotic harvesting systems," *Agriculture (Switzerland)*, vol. 12, no. 8, pp. 1–35, 2022, doi: 10.3390/agriculture12081240.

[2] P. Sengaphone *et al.*, "End-effectors of the robotic arms for tomato harvesting: A comprehensive review," *Journal of Computational and Applied Research in Mechanical Engineering*, vol. 14, no. 1, pp. 1–18, 2024, doi: 10.22061/jcarme.2024.10567.2397.

[3] W. Ye, L. Zhao, X. Luo, J. Guo, and X. Liu, "Perceptual soft end-effectors for future unmanned agriculture," *Sensors*, vol. 23, no. 18, pp. 1–28, 2023, doi: 10.3390/s23187905.

[4] F. Zhang *et al.*, "Research on flexible end-effectors with humanoid grasp function for small spherical fruit picking," *Agriculture (Switzerland)*, vol. 13, no. 1, pp. 1–18, 2023, doi: 10.3390/agriculture13010123.

[5] Z. Li, X. Yuan, and C. Wang, "A review on structural development and recognition–localization methods for end-effector of fruit–vegetable picking robots," *International Journal of Advanced Robotic Systems*, vol. 19, no. 3, pp. 1–29, 2022, doi: 10.1177/17298806221104906.

[6] J. Gao *et al.*, "Development and evaluation of a pneumatic finger-like end-effector for cherry tomato harvesting robot in greenhouse," *Computers and Electronics in Agriculture*, vol. 197, pp. 1–10, 2022, doi: 10.1016/j.compag.2022.106879.

[7] V. F. Tejada, M. F. Stoelen, K. Kusnierek, N. Heiberg, and A. Korsaeth, "Proof-of-concept robot platform for exploring automated harvesting of sugar snap peas," *Precision Agriculture*, vol. 18, no. 6, pp. 952–972, 2017, doi: 10.1007/s11119-017-9538-1.

[8] E. Vrochidou *et al.*, "An autonomous grape-harvester robot: Integrated system architecture," *Electronics (Switzerland)*, vol. 10, no. 9, pp. 1–22, 2021, doi: 10.3390/electronics10091056.

[9] Y. Park, J. Seol, J. Pak, Y. Jo, J. Jun, and H. Il Son, "A novel end-effector for a fruit and vegetable harvesting robot: mechanism and field experiment," *Precision Agriculture*, vol. 24, no. 3, pp. 948–970, 2023, doi: 10.1007/s11119-022-09981-5.

[10] A. Yeshmukhametov, K. Koganezawa, Y. Yamamoto, Z. Buribayev, Z. Mukhtar, and Y. Amirgaliyev, "Development of continuum robot arm and gripper for harvesting cherry tomatoes," *Applied Sciences (Switzerland)*, vol. 12, no. 14, pp. 1–22, 2022, doi: 10.3390/app12146922.

[11] Y. Yuan, J. Sun, and Q. Zhang, "An enhanced deep learning model for effective crop pest and disease detection," *Journal of Imaging*, vol. 10, no. 11, pp. 1–15, 2024, doi: 10.3390/jimaging10110279.

[12] J. Wu, C. Zhang, Z. Li, H. Fu, X. Peng, and J. T. Zhou, "dugMatting: Decomposed-uncertainty-guided matting," in *Proceedings of the 40th International Conference on Machine Learning*, 2023, pp. 37846–37859. doi: 10.48550/arXiv.2306.01452.

[13] Q. Zhang, Y. Zhu, F. R. Cordeiro, and Q. Chen, "PSSCL: A progressive sample selection framework with contrastive loss designed for noisy labels," *Pattern Recognition*, vol. 161, pp. 1–15, 2025, doi: 10.1016/j.patcog.2024.111284.

[14] Z. Wang *et al.*, "An improved Faster R-CNN model for multi-object tomato maturity detection in complex scenarios," *Ecological Informatics*, vol. 72, pp. 1–20, 2022, doi: 10.1016/j.ecoinf.2022.101886.

[15] M. Afonso *et al.*, "Tomato fruit detection and counting in greenhouses using deep learning," *Frontiers in Plant Science*, vol. 11, pp. 1–12, 2020, doi: 10.3389/fpls.2020.571299.

[16] T. Zheng, M. Jiang, Y. Li, and M. Feng, "Research on tomato detection in natural environment based on RC-YOLOv4," *Computers and Electronics in Agriculture*, vol. 198, pp. 1–15, 2022, doi: 10.1016/j.compag.2022.107029.

[17] Y. Ge *et al.*, "Tracking and counting of tomato at different growth period using an improving YOLO-Deepsort network for inspection robot," *Machines*, vol. 10, no. 6, pp. 1–20, 2022, doi: 10.3390/machines10060489.

[18] C. Song *et al.*, "TDPPL-Net: A lightweight real-time tomato detection and picking point localization model for harvesting robots," *IEEE Access*, vol. 11, pp. 37650–37664, 2023, doi: 10.1109/ACCESS.2023.3260222.

[19] W. Zheng, G. Dai, M. Hu, and P. Wang, "A robust tomato counting framework for greenhouse inspection robots using YOLOv8 and inter-frame prediction," *Agronomy*, vol. 15, no. 5, pp. 1–20, 2025, doi: 10.3390/agronomy15051135.

[20] Z. Qi *et al.*, "An improved framework based on tracking-by-detection for simultaneous estimation of yield and maturity level in cherry tomatoes," *Measurement: Journal of the International Measurement Confederation*, vol. 226, pp. 1–15, 2024, doi: 10.1016/j.measurement.2024.114117.

[21] J. Rong, G. Dai, and P. Wang, "A peduncle detection method of tomato for autonomous harvesting," *Complex and Intelligent*

Systems, vol. 8, no. 4, pp. 2955–2969, 2022, doi: 10.1007/s40747-021-00522-7.
[22] J. Rong, H. Zhou, F. Zhang, T. Yuan, and P. Wang, “Tomato cluster detection and counting using improved YOLOv5 based on RGB-D fusion,” *Computers and Electronics in Agriculture*, vol. 207, pp. 1–15, 2023, doi: 10.1016/j.compag.2023.107741.
[23] T. Fujinaga, S. Yasukawa, and K. Ishii, “Evaluation of tomato fruit harvestability for robotic harvesting,” in *2021 IEEE/SICE International Symposium on System Integration, SII 2021*, 2021, pp. 35–39. doi: 10.1109/IEEECONF49454.2021.9382603.
[24] V. Muthupandi, K. Kannan, J. Saroj Kanna, and K. Esakki Bala Karthik, “Tomato harvesting robot using Yolov8, Midas and inverse kinematics for smart farming automation,” *International Research Journal on Advanced Engineering and Management (IRJAEM)*, vol. 3, no. 04, pp. 1122–1128, 2025, doi: 10.47392/irjaem.2025.0183.
[25] R. Singh, L. Seneviratne, and I. Hussain, “A deep learning-based approach to strawberry grasping using a telescopic-link differential drive mobile robot in ROS-Gazebo for greenhouse digital twin environments,” *IEEE Access*, vol. 13, pp. 361–381, 2025, doi: 10.1109/ACCESS.2024.3520233.
[26] J. Jun, J. Kim, J. Seol, J. Kim, and H. Il Son, “Towards an efficient tomato harvesting robot: 3D perception, manipulation, and end-effector,” *IEEE Access*, vol. 9, pp. 17631–17640, 2021, doi: 10.1109/ACCESS.2021.3052240.
[27] Y. Oktarina, T. Dewi, P. Risma, and M. Nawawi, “Tomato harvesting arm robot manipulator; A pilot project,” *Journal of Physics: Conference Series*, vol. 1500, no. 1, pp. 1–10, 2020, doi: 10.1088/1742-6596/1500/1/012003.
[28] J. Rong, L. Hu, H. Zhou, G. Dai, T. Yuan, and P. Wang, “A selective harvesting robot for cherry tomatoes: Design, development, field evaluation analysis,” *Journal of Field Robotics*, vol. 41, no. 8, pp. 2564–2582, 2024, doi: 10.1002/rob.22377.
[29] N. Xu, Z. Li, J. Kang, Q. Meng, and M. Niu, “Agricultural vehicle automatic navigation positioning and obstacle avoidance technology based on ICP,” *IEEE Access*, vol. 12, pp. 85940–85954, 2024, doi: 10.1109/ACCESS.2024.3402743.
[30] Z. Wang, P. Li, Z. Wang, and Z. Li, “APG-RRT: Sampling-based path planning method for small autonomous vehicle in closed scenarios,” *IEEE Access*, vol. 12, pp. 25731–25739, 2024, doi: 10.1109/ACCESS.2024.3359643.
[31] A. Ghobadpour, G. Monsalve, A. Cardenas, and H. Mousazadeh, “Off-road electric vehicles and autonomous robots in agricultural sector: Trends, challenges, and opportunities,” *Vehicles*, vol. 4, no. 3, pp. 843–864, 2022, doi: 10.3390/vehicles4030047.
[32] X. Li, N. Ma, Y. Han, S. Yang, and S. Zheng, “AHPPEBot: Autonomous Robot for Tomato Harvesting based on Phenotyping and Pose Estimation,” in *Proceedings - IEEE International Conference on Robotics and Automation*, 2024, pp. 18150–18156. doi: 10.1109/ICRA57147.2024.10610454.

BIOGRAPHIES OF AUTHORS

Thien An Dinh obtained his Bachelor of Science in mechatronics engineering from the Ho Chi Minh University of Technology in 2025. His main research directions include industrial manipulator robots, collaborative robots, mobile robots, embedded systems, and automation control systems. He can be contacted at an.dinhthien@hcmut.edu.vn.

So Nam Phung obtained his Bachelor of Science in mechatronics engineering from the Ho Chi Minh University of Technology in 2025. His main research directions include industrial manipulator robots, collaborative robots, mobile robots, embedded systems, and automation control systems. He can be contacted at nam.phung1612@hcmut.edu.vn.

Tri Cong Phung is a senior lecturer at Ho Chi Minh University of Technology (HCMUT) in Vietnam. Tri Cong Phung obtained his Bachelor of Science in Mechatronics Engineering from the Ho Chi Minh University of Technology in 2004, his Master of Science in Intelligent Robotics and Mechatronic System Laboratory, Mechanical Engineering from Sungkyunkwan University, Korea in 2007 and his Ph.D. also from Sungkyunkwan University in 2011. His main research directions include industrial manipulator robots, collaborative robots, mobile robots, embedded systems, and automation control systems. He can be contacted at ptcong@hcmut.edu.vn.