

Vision-Based Positioning and Control of a Robotic Manipulator for Tomato Harvesting

Rustam Baratov¹ , Khurshid Sattarov^{1,2}  and Almardon Mustafoqulov¹ 

¹ “Tashkent Institute of Irrigation and Agricultural Mechanization Engineers” National Research University, Kary Niyaziy Str., 39, 100000, Tashkent, Uzbekistan.

² Tashkent University of Information Technologies named after Muhammad al-Khwarizmi, 108 A. Timur, Tashkent, 100200, Uzbekistan.

Submitted: 4 September 2026
Accepted : 5 October 2026
Online First: 12 October 2026

Corresponding author
Almardon Mustafoqulov
Mamatovich
a.mustafaqulov@tiame.uz

DOI:10.64470/elene.2026.41

RESEARCH ARTICLE

© Copyright, Authors,
Distributed under Creative
Commons CC-BY 4.0

Abstract: Automated harvesting of tomatoes requires reliable integration of machine vision, distance measurement, and robotic-manipulator control to achieve accurate and repeatable product positioning. This study presents a vision-based positioning and control system for a robotic manipulator designed for automatic tomato harvesting. The proposed system integrates a webcam, Raspberry Pi 4 for image processing, a microcontroller for control, an HC-SR04 ultrasonic sensor for distance measurement, and servo motors for manipulator actuation. Image-based tomato coordinates are used to generate the required angular positions of the manipulator joints, while distance information is used to support the approach motion toward the detected product. A closed-loop control concept based on the comparison between commanded and actual angular positions is also considered. Experimental evaluation was performed using red and green tomatoes to assess detection, harvesting, and complete work-cycle times. The detection and harvesting times were 4.932 s for the red tomato and 5.276 s for the green tomato, while the complete cycle times were 9.676 and 10.586 s, respectively. For an image resolution of 320×180 pixels, an example object position of $x = 15$ pixels and $y = 42$ pixels corresponded to servo angles of 115° for the base joint and 100° for the first joint. The results demonstrate the feasibility of integrating visual perception, distance measurement, and manipulator motion control using a relatively simple hardware architecture. The proposed approach provides a basis for further development of accurate, repeatable, and energy-efficient robotic tomato harvesting systems.

Keywords Agricultural robot, Computer vision, Robotic manipulator, Angular displacement, Closed-loop control, Raspberry Pi 4.

1. Introduction

A large part of agricultural fruit and vegetable production and harvesting still relies on manual labor. In particular, when harvesting soft and mechanically sensitive products such as tomatoes, determining the maturity level of the product, estimating its spatial position, and accurately bringing the harvesting mechanism to the required point are important technological tasks. Robotization makes it possible to reduce the need for manual labor, increase labor productivity, harvest products without damage, and improve the repeatability of technological processes (Baratov & Mustafoqulov, 2026; Mustafoqulov & Yoqubjonov,

2026).

In recent years, the combined use of computer vision, distance sensors, microprocessor-based control systems, and robotic manipulators has been expanding in agricultural robots. In such systems, the presence and position of the crop are determined from images acquired by a camera, while a distance sensor is used to estimate the distance between the manipulator and the product. In the robotic system investigated in this study, the camera transmits images to the Raspberry Pi 4 platform, and the processed information is used through a microcontroller to control the manipulator's servo motors. An HC-SR04 ultrasonic sensor is used to determine the position of the robotic manipulator relative to the product. However, automatic crop detection alone is not sufficient for high-quality harvesting. The manipulator's end-effector must approach the detected product along the required trajectory and perform the grasping or detachment operation at the specified position. Therefore, determining and controlling the angular positions of the manipulator joints is of particular importance. In open-loop control, the actual angular position may differ from the commanded value due to external influences, changes in mechanical load, actuator uncertainty, and wear of mechanical components. Closed-loop control, in contrast, makes it possible to monitor the actual state through sensor feedback and appropriately correct the control signal.

Another important aspect of controlling angular displacement in robotic manipulators is the relationship between motion accuracy and energy consumption. Excessive or unnecessary manipulator movements increase actuator operating time and consequently increase energy consumption. Therefore, determining the spatial position of the product, generating the corresponding angular position, and controlling manipulator motion based on feedback are important conditions for creating an energy-efficient robotic harvesting system.

This study considers a control system for a robotic manipulator designed for automatic detection and harvesting of tomatoes. The system consists of a camera, Raspberry Pi 4, microcontroller, HC-SR04 ultrasonic sensor, and servo motors. The position of the product detected by the camera is processed and transmitted to the control system to generate the corresponding manipulator movements. The manipulator's approach motion toward the product is organized by taking its spatial position and angular state into account.

The main objectives of this study are to develop a control approach based on accurate measurement and control of the angular displacement of robotic-manipulator joints during selective tomato harvesting and to experimentally evaluate its technological performance.

At the same time, the approach is aimed at adapting manipulator motion to the detected position of the product, reducing harvesting time, and increasing the possibility of harvesting the product without damage. In the preliminary experiments, 4.932 s was required to detect and harvest a red tomato, while 9.676 s was required to complete the harvesting operation and return the robot to its initial position. For a green tomato, the corresponding values were 5.276 s and 10.586 s. These results indicate that the manipulator operating time is related more to the product's distance relative to the robot than to the product color.

The main technical contribution of this study is the development and experimental implementation of a relatively simple, low-cost robotic control architecture that integrates machine vision, distance sensing, and servo-based manipulator control into a unified processing chain for tomato harvesting. Unlike approaches that primarily rely on complex 3D vision systems or dedicated visual-servoing architectures, the proposed system uses image-based object coordinates together with distance information obtained from an HC-SR04 ultrasonic sensor to generate manipulator motion commands. A further contribution is the practical integration of detected image coordinates with the angular positioning of the manipulator joints, providing a direct link between visual perception and mechanical actuation. The developed architecture demonstrates the feasibility of coordinating these functions using commercially available, low-cost hardware in a controlled experimental environment.

2. Analysis of Existing Research and Technological Solutions

The rapid development of research on the use of robotic systems in agriculture makes it possible to address, in an integrated manner, the automatic detection of fruits and vegetables, estimation of their spatial coordinates, manipulator motion planning, and harvesting with minimal mechanical impact. Modern harvesting robots generally consist of a mobile platform, robotic manipulator, vision system, and specialized end-effector, and their performance depends on the coordinated operation of these components (Droukas et al., 2023).

Machine vision technology is one of the principal tools for object detection and position estimation in fruit and vegetable harvesting (Hou et al., 2023; Y. Yang et al., 2023). Earlier systems used RGB cameras, stereo cameras, RGB-D cameras, and other sensors to acquire crop images and transmit their coordinates to the manipulator. Modern reviews of machine vision show that, in addition to fruit detection, increasing attention is being paid to determining the harvesting point, spatial localization, and visual servoing.

Robot-assisted tomato harvesting is particularly challenging because of variations in fruit size and shape, partial occlusion by leaves and stems, changing illumination conditions, and differences in maturity. Therefore, tomato detection systems must determine not only the presence of an object but also its precise position on the plant. A systematic review of computer vision published in 2023 analyzed a total of 1,285 articles were identified during the search process, and 121 articles were included in this review after screening and showed that deep-learning methods such as YOLO are widely used, while transferring laboratory results to real field conditions remains an important challenge (Guduru et al., 2023). Recent studies have focused on improving fruit recognition and spatial localization in harvesting robots through the use of different vision sensors and image-processing algorithms. Yang et al. (2023) reviewed vision-system hardware, sensor technologies, and algorithmic approaches for fruit recognition and positioning in harvesting robots. (Y. Yang et al., 2023). In that system, manipulator configuration, environment perception, and end-effector control were considered as a unified system.

Other studies have emphasized controlling the manipulator using visual feedback. For example, in a system developed for cherry-tomato harvesting, an RGB-D camera was mounted on the end-effector, the spatial state of the object was determined, and eye-in-hand visual servoing control was applied (Li et al. (2023) ,Thakur et al., 2023). The hybrid visual servoing method combines position-based and image-based control to improve manipulator accuracy and motion efficiency. Laboratory tests of this system recorded an average harvesting time of 9.40 s per product and a harvesting success rate of 96.25% (Baratov et al., 2025). At the same time, although open-loop control of a robotic manipulator provides a relatively simple hardware and software structure, its performance strongly depends on the accuracy of image-based coordinate estimation (Baratov & Mustafoqulov, 2023b). If deviations from the initially calculated coordinates, mechanical backlash, load changes, or partial object occlusion occur during manipulator motion, an open-loop control system cannot independently correct the actual state (Bai et al., 2023). A 2024 review of selective harvesting robots also identified the inability to compensate for visual detection errors during motion as a major limitation of open-loop control (Mustafoqulov et al., 2024).

In closed-loop control, the actual state of the robot is determined using sensors, and the control signal is regenerated based on the difference between the commanded and actual states. From this perspective, measuring and controlling the angular displacement of manipulator joints is an important element in determining the robot's spatial position. The accuracy of the angular state directly affects the end-effector approach trajectory toward the tomato, the grasping or cutting point, and the repeatability of the overall harvesting cycle.

The robotic-manipulator concept investigated in this study uses a relatively simple hardware architecture consisting of a camera, Raspberry Pi 4, microcontroller, servo motors, and an HC-SR04 ultrasonic sensor. The image acquired by the camera is processed on the Raspberry Pi 4 platform, information about the detected object is transmitted to the microcontroller, and the manipulator is moved through the

corresponding servo motors. The HC-SR04 sensor is used to determine the distance between the manipulator and the product. This approach reduces the hardware complexity of the robotic system while linking manipulator motion with sensor information. Preliminary studies also observed a dependence of manipulator motion on product position. An image plane of 320×180 pixels was used, and angular positions of the manipulator joints were specified according to the detected object coordinates. For example, at $x = 15$ pixels and $y = 42$ pixels, angular values of 115° for the base servo motor and 100° for the first joint were applied. This result demonstrates the possibility of generating manipulator kinematic motion parameters from image-based object coordinates.

Modern tomato harvesting robots increasingly integrate vision systems, advanced sensors, and manipulator control algorithms. However, system accuracy is not determined solely by object-detection accuracy. Converting detected coordinates into the actual angular state of the manipulator, detecting deviations that occur during motion, and compensating for them are also important issues. Therefore, integrating image-based object detection, distance measurement, and control of manipulator-joint angular displacement into a single closed-loop system is considered a major direction of this study.

The distinctive feature of this approach is that, rather than providing the manipulator only with a pre-calculated motion command, the product's position in the image is linked to the actual angular state of the manipulator. As a result, the control system can determine the error between the commanded and measured angular values and, when necessary, correct the control signal. This provides a basis for improving the robot's positioning accuracy, repeatability, and energy efficiency (Mustafoqulov et al., 2024).

Although high-precision systems such as RGB-D and LiDAR, deep-learning algorithms, and visual-servoing methods are actively developing in another line of research, organizing object detection, distance estimation, angular-displacement measurement, and manipulator control as a unified system based on a relatively simple hardware architecture remains practically and scientifically relevant (Malyshev et al., 2026).

In tomato harvesting robots, visual servoing technology enables real-time correction of the manipulator's position relative to the detected object. Liang et al. developed a 6+1 DOF tomato harvesting robot based on visual servo control by combining global and palm vision systems (Liang et al., 2025). In addition to tomato detection, determining the growth direction, stem, and cluster position also affects harvesting success. Zheng et al. developed a tomato-cluster localization and grasping strategy using multiple YOLOv5-based detectors and point-cloud data (Zheng et al., 2024). When optimizing the manipulator motion trajectory, not only time and accuracy but also energy consumption and motion smoothness should be considered. Ma et al. developed a multi-objective trajectory optimization method for a fruit-harvesting manipulator, jointly optimizing energy consumption and impact reduction (Ma et al., 2024). Successful grasping of tomatoes requires matching the detected object coordinates with the spatial pose of the end-effector. Wang et al. proposed a tomato harvesting robot using dual-camera image-based visual servoing to improve target–grripper alignment (Wang, L. et al., 2024). Adaptive pose-control approaches have also been proposed to address the problem of adjusting and controlling the spatial pose of the end-effector in tomato harvesting robots (Wang, D. et al., 2023). Other visual-servoing studies have improved harvesting accuracy by correcting the manipulator's position relative to the object in real time (Wang et al., 2023). In tomato harvesting robots, machine vision is used not only to detect the object but also to determine its spatial coordinates and generate end-effector motion (Li et al., 2023). Determining not only the central coordinates of the tomato but also its spatial pose and harvesting direction is necessary for correctly orienting the manipulator's end-effector (Rong et al., 2022; Yin et al., 2025).

3. Materials and Methods

3.1. Structural configuration of the robotic manipulator

The study considered a robotic-manipulator system for automatic detection and harvesting of agricultural products, particularly tomatoes. The main task of the proposed system is to estimate the position of the

product detected from an image, direct the manipulator toward the product, and perform the harvesting operation.

The robotic system consists of functional components including an image-acquisition device, image-processing platform, microcontroller-based control unit, distance sensor, and manipulator actuators. A webcam was used for image acquisition, a Raspberry Pi 4 platform for image processing, a microcontroller for controlling the manipulator servo motors, and an HC-SR04 ultrasonic sensor for determining the distance between the product and the manipulator.

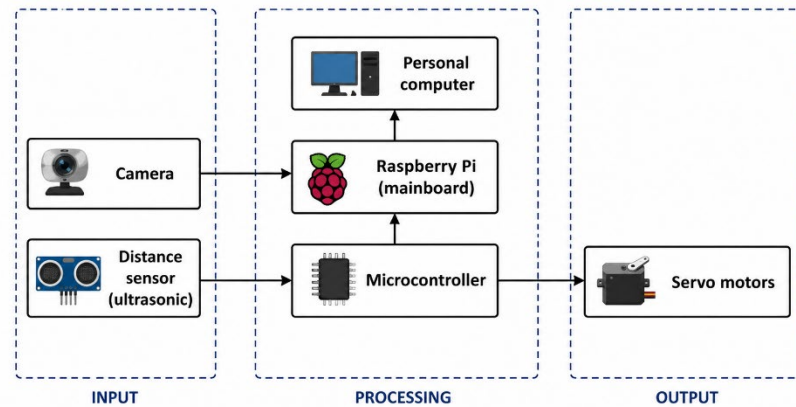


Figure 1. Structural diagram of the robotic manipulator

Figure 1 presents the functional structure of the robotic manipulator (Baratov et al., 2023a). The image acquired by the camera is transmitted to the processing platform. Based on information about the detected object, the microcontroller generates the corresponding control commands for the manipulator servo motors. The servo motors are mounted on the joints of the robotic manipulator and perform the required movement according to the position of the detected object.

The main feature of this architecture is the organization of object detection and manipulator motion as a single technological process. The camera serves as the sensing element of the system, the Raspberry Pi 4 performs image processing and object detection, and the microcontroller controls the actuating mechanisms.

3.2. Image acquisition and object detection system

The first stage of automatic tomato harvesting is image-based object detection. In this study, an image of the plant and its product is acquired using a webcam and transmitted to the Raspberry Pi 4 platform for further processing. According to the description of the system, the camera is positioned close to the manipulator's working part, allowing the product to be detected through the robot's vision system.

The image plane was set to a resolution of 320×180 pixels. The coordinates of the detected tomato in the image were used to generate the required manipulator motion parameters. Specifically, when $x = 15$ pixels and $y = 42$ pixels, the rotation angle of the base servo motor was set to 115° , while the rotation angle of the first-joint servo motor was set to 100° . As a result, the robot arm moved toward the left and approached the detected product. This relationship makes it possible to convert object coordinates in the image into mechanical motion parameters of the manipulator.

3.3. Distance measurement and sensor system

In addition to detecting the object from the image, it is necessary to determine the distance between the tomato and the manipulator's end-effector. An HC-SR04 ultrasonic sensor was used for this purpose. The sensor was mounted on the working part of the manipulator to determine the distance between the product and the harvesting device.

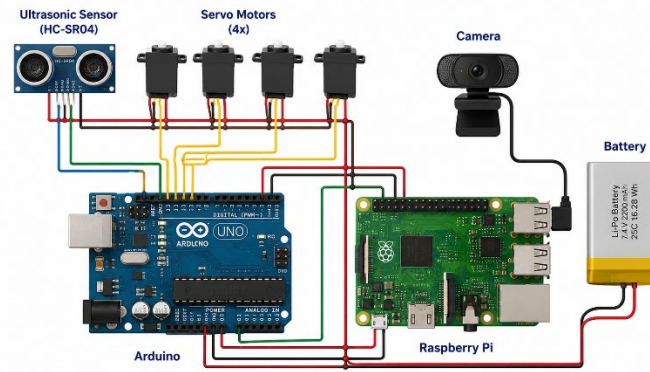


Figure 2. Connection diagram of the robotic-manipulator components

Figure 2 shows the main electrical components of the robot and their interconnections (Baratov et al., 2023b). The camera is connected to the Raspberry Pi 4 platform, while the HC-SR04 is used to determine the distance between the manipulator and the product. The specified operating range of the HC-SR04 sensor is 2–400 cm, and its sensing angle is 30°.

Distance information is used as additional sensor information for controlling the manipulator's approach to the product. Therefore, jointly considering the object's image coordinates and its distance is important for achieving more accurate manipulator positioning.

In this study, the HC-SR04 sensor is used primarily as an auxiliary distance-measurement device rather than as a general-purpose obstacle-detection sensor. Its function is to estimate the distance between the manipulator's working part and the detected tomato during the approach process. The sensor was selected for the laboratory prototype because of its low cost, simple hardware interface, and ease of integration with the microcontroller. However, its performance may be affected by environmental conditions and object characteristics in complex agricultural environments. Therefore, more advanced distance-sensing technologies, such as RGB-D cameras or LiDAR, may be considered in future versions of the system.

3.4. Generation of manipulator angular parameters

The manipulator can reach the product only if its joints move to appropriate angular positions. In this study, servo-motor rotation angles are specified based on the coordinates of the object in the image. For the identified experimental case with $x = 15$ and $y = 42$ pixels, the base servo motor was set to 115°, while the first-joint servo motor was set to 100°.

Under this approach, a functional relationship exists between the image coordinates and the angular state of the manipulator. In general, this relationship can be expressed as follows:

$$\theta_i = f_i(x, y, d) \quad (1)$$

where (x, y) are the coordinates of the object in the image, (d) is the distance between the object and the working body, (θ_i) is the angular position of the (i) -joint of the manipulator.

This equation expresses the general control idea of the study. However, additional experimental data are required to determine the exact mathematical form of the function (f_i) and its coefficients. Therefore, the relationships obtained in this work based on the available experimental points are analyzed separately in the following results section.

The kinematic model of the robotic manipulator was established to relate the Cartesian position of the tomato to the angular positions of the manipulator joints. Let L_1 , L_2 , and L_3 denote the characteristic link lengths of the manipulator, while θ_0 , θ_1 , and θ_2 represent the base, shoulder, and subsequent joint angles, respectively. For a target position represented by the Cartesian coordinates (x, y, z) , the horizontal radial distance from the base axis is expressed as

$$r = \sqrt{x^2 + y^2}$$

The base rotation angle is then obtained from $\theta_0 = \text{atan2}(x, y)$.

For the planar motion of the remaining links, the target coordinates can be represented by the forward kinematic relationships

$$r = L_2 \cos \theta_1 + L_3 \cos(\theta_1 + \theta_2), \quad (z - L_1) = L_2 \sin \theta_1 + L_3 \sin(\theta_1 + \theta_2).$$

Therefore, defining

$$D = \frac{r^2 + (z - L_1)^2 - L_2^2 - L_3^2}{2L_2L_3},$$

the second joint angle can be determined from $\theta_2 = \text{atan2}(\pm\sqrt{1 - D^2})$

The corresponding first joint angle is obtained as

$$\theta_1 = \text{atan2}(z - L_1, r) - \text{atan2}(L_3 \sin \theta_2, L_2 + L_3 \cos \theta_2).$$

The two possible signs in the expression for θ_2 correspond to the alternative manipulator configurations. The physically feasible solution is selected according to the mechanical joint limits and the required approach direction of the end-effector. Thus, the inverse kinematic model provides the joint-angle commands required to position the end-effector toward the detected tomato.

3.5. Automatic control algorithm of the robotic manipulator

In the automatic control system of the robotic manipulator, the command device, actuator, controlled object, and sensing element operate in an interconnected manner. In the proposed system, these elements are respectively designated as the command-generating device (TBQ), actuator mechanism (IM), control organ (BO), and sensing element (SE). The image-detection sensor is indicated as the sensing element.

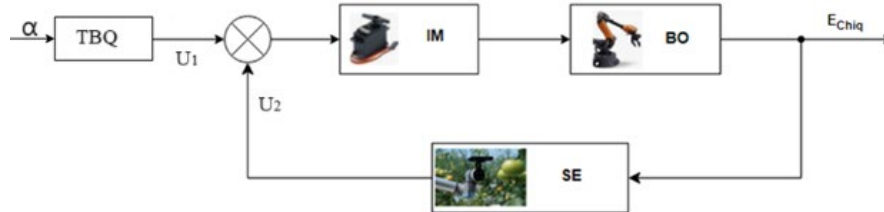


Figure 3. Functional diagram of automatic regulation of a robotic manipulator intended for harvesting agricultural products

Based on the functional diagram in Figure 3, the control process is organized in the following sequence: first, information about the object is obtained through the camera; the image is processed; information about the detected object is transmitted to the control system; and the motion of the actuating mechanisms is generated based on this information. The servo motors bring the manipulator joints to the required positions. According to the description, the detected tomato is assigned the binary value “1”, and the corresponding information is transmitted to the microcontroller. The servo motors then perform the required movement toward the detected object. This control element determines the magnitude of the voltage generated by the rotation ("Open"). A sensing element detects this voltage value and transmits the signal to a comparison element, which controls the rotation of the servo motor to the appropriate angle.

3.6. Closed-loop control and the PID approach

A feedback control system can be used to compare the actual state of the manipulator with the commanded state. In open-loop systems, the actual state of the robot is not monitored after the control signal is generated. As a result, the actual state may deviate from the commanded state due to mechanical load, external forces, or changes in mechanical components.

In closed-loop control, the actual state of the robot is measured using sensors and compared with the commanded state. The control signal is corrected based on the error detected. This approach can improve the positioning accuracy and repeatability of the manipulator.

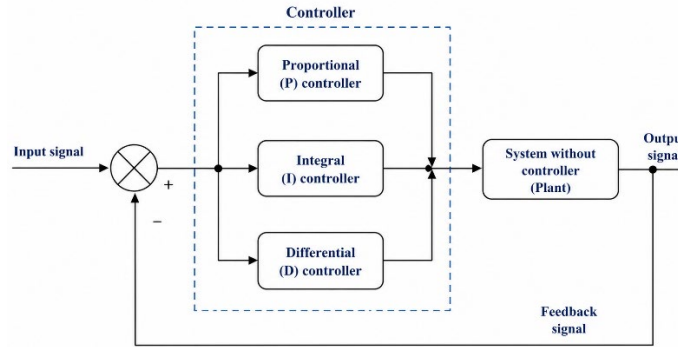


Figure 4. PID (Proportional–Integral–Derivative) control system

In a PID control system, the control signal is formed from proportional, integral, and derivative components. The proportional component depends on the current error value, the integral component on the time integral of the error, and the derivative component on the rate of change of the error. These components are represented by separate equations and by the general PID expression.

In general, the PID control signal can be expressed as follows:

$$u(t) = K_p \cdot e(t) + K_i \int e(t) dt + K_d \frac{de(t)}{dt}. \quad (2)$$

where $e(t)$ is the error between the commanded and measured states, and K_p , K_i and K_d are the proportional, integral, and derivative coefficients, respectively.

In this approach, the actual angular position θ_{meas} , is obtained from the angular-displacement sensor and compared with the commanded angular position θ_{ref} . The control error is calculated as follows: $e_\theta = \theta_{ref} - \theta_{meas}$.

Although the closed-loop and PID control principles are considered in the proposed control architecture, the present experimental study does not include a complete quantitative evaluation of PID tracking performance. Experimentally recorded PID gain parameters, commanded and actual joint-angle trajectories, and time-dependent tracking errors were not available for the present experiments. Therefore, the current results should be interpreted as an initial experimental demonstration of the integration of sensing, angular positioning, and manipulator control rather than as a complete quantitative validation of the PID controller. A more comprehensive evaluation will require experimental determination of the PID parameters and measurement of commanded and actual joint angles under different positions and loading conditions. The corresponding tracking errors can then be quantified using indicators such as MAE, RMSE, and maximum positioning error.

The control loop of the proposed robotic harvesting system consists of tomato detection, distance measurement, target-state comparison, control-command generation, servo actuation, and manipulator motion. The image-processing system determines the target tomato position, while the distance sensor provides additional information about the relative distance between the manipulator and the target. The measured information is compared with the desired target state, and the resulting control signal is used to generate the corresponding servo commands. The servo motors then adjust the manipulator configuration to approach and grasp the detected tomato.

The present experimental evaluation was performed using a lightweight manipulator prototype under controlled laboratory conditions. The effects of payload, tomato mass, gripper loading, joint inertia, and other dynamic characteristics were not independently measured or quantitatively modeled. Consequently,

the current results do not provide a complete dynamic analysis of the manipulator under different loading conditions. Further experimental studies will investigate the influence of payload and dynamic loading on positioning accuracy, response time, and closed-loop control performance.

4. Experimental Studies and Results

4.1. Experimental setup and research conditions

The operating performance of the robotic manipulator was evaluated through the processes of detecting the tomato, approaching it, harvesting it, and returning to the initial position.

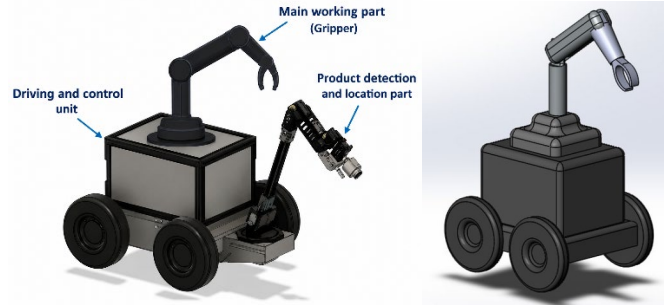


Figure 5. 3D view of the robotic manipulator

During the experiment, the robot's movement from the initial position to product detection and harvesting, as well as its return to the initial position after harvesting, was observed. Red and green tomatoes were used as experimental objects. According to the results presented in the refer to the cited study, the robot's operating time was related more to the product's distance relative to the manipulator than to its color (Baratov et al., 2023).

4.2. Tomato detection and harvesting time

One of the main experimental indicators was the time required to detect and harvest the tomato and the time required for the robot to return to its initial position after the harvesting operation. The measured detection and harvesting times and complete cycle times for red and green tomatoes are presented in Table 1.

Table 1. The times for tomato detection and harvesting and for the robot's return to the initial position

Tomato type	Detection and harvesting time, s	Complete cycle time, s
Red	4.932	9.676
Green	5.276	10.586

For the red tomato, the detection and harvesting process took 4.932 s, while the complete cycle, including the robot's return to the initial position after harvesting, took 9.676 s. For the green tomato, the corresponding values were 5.276 s and 10.586 s. Comparing the results, the detection and harvesting time for the green tomato was higher than for the red tomato by:

The absolute difference was: $\Delta t_{\text{harvesting}} = t_{\text{green}} - t_{\text{red}} = 5.276 - 4.932 = 0.344$ s

The relative difference was:

$$\delta_{\text{harvesting}} = \frac{t_{\text{green}} - t_{\text{red}}}{t_{\text{red}}} \times 100\% = \frac{5.276 - 4.932}{4.932} \times 100\% \approx 6.97\%. \quad (3)$$

For the complete cycle, the difference was $\Delta t_{\text{cycle}} = t_{\text{green.cycle}} - t_{\text{red.cycle}} = 10.586 - 9.676 = 0.910$ s, and the complete-cycle time for the green tomato was approximately:

$$\delta_{\text{cycle}} = \frac{t_{\text{green.cycle}} - t_{\text{red.cycle}}}{t_{\text{red.cycle}}} \times 100\% = \frac{10.586 - 9.676}{9.676} \times 100\% \approx 9.4\% \text{ higher.}$$

However, these differences cannot be attributed solely to tomato color. As noted above, the difference in

time values is related to the product's distance relative to the robot. Therefore, the time difference observed by color should be regarded as a result recorded under the given experimental conditions, and additional controlled experiments are required to determine the independent effect of color.

4.3. Generation of the manipulator angular state

The accuracy with which the robotic manipulator reaches the product is based on the relationship between the object coordinates in the image and the corresponding angular positions of the manipulator joints.

The image plane was set to 320×180 pixels. In one experimental case, the tomato coordinates in the image were $x = 15$ px and $y = 42$ px.

For these coordinates, the rotation angle of the servo motor at the base is set to: $\theta_0 = 115^\circ$ and the rotation angle of the servo motor at the first joint is set to: $\theta_1 = 100^\circ$.

where:

x — horizontal coordinate of the object in the image;

y — vertical coordinate of the object in the image;

θ_0 — rotation angle of the base servo motor;

θ_1 — rotation angle of the first-joint servo motor.

As a result of this control action, the robot arm moved to the left and approached the detected tomato. The corresponding image dimensions, detected object coordinates, and manipulator servo angles recorded during this experiment are presented in Table 2.

Table 2. The experimental values of the object coordinates and manipulator angles

Indicator	Value
Image width	320 px
Image height	180 px
Object (x) coordinate	15 px
Object (y) coordinate	42 px
Base servo angle	115°
First-joint servo angle	100°

This result shows that the required angular position of the manipulator can be generated using object coordinates determined in the image plane. However, because the available experimental data include only one coordinate case, additional experimental points are required to reliably determine the general mathematical relationship between object coordinates and servo angles.

The experimental results showed that the proposed robotic manipulator can automate tomato detection, approach, and harvesting. The robot's operating performance was evaluated mainly using two indicators: product detection and harvesting time and the duration of the complete work cycle. The possibility of generating the angular states of the manipulator servo motors based on the tomato coordinates in the image was also analyzed. The results showed that 4.932 s was required to detect and harvest the red tomato and 9.676 s for the complete work cycle. For the green tomato, these values were 5.276 s and 10.586 s, respectively. The time difference between the detection and harvesting stages was 0.344 s. The difference between the complete work cycles was $10.586 - 9.676 = 0.910$ s. In relative terms, the detection and harvesting time for the green tomato was approximately 6.97% higher than for the red tomato, while the complete work cycle was 9.40% higher.

The present experimental study primarily evaluates tomato detection and harvesting time, complete work-cycle time, and the generation of manipulator angular positions from image-based coordinates. However, the available experimental dataset does not provide sufficient repeated measurements to quantitatively determine the harvesting success rate, end-effector positioning accuracy, harvesting efficiency, or fruit damage rate. Therefore, these performance indicators are not assigned numerical values in the present study. Further experimental evaluation with repeated harvesting trials under systematically varied tomato positions, fruit sizes, distances, and environmental conditions is required to quantitatively assess these

parameters and provide a more comprehensive evaluation of the proposed robotic harvesting system.

However, it is not scientifically sufficient to explain this difference solely by the effect of tomato color. Our previous studies also showed that the time difference may be related to the product's position and distance relative to the robot (Baratov & Mustafoqulov, 2023a, 2023b). Therefore, it is appropriate to regard the observed color-related difference as a difference recorded under the experimental conditions.

This result shows that optimizing the robot's work cycle requires consideration not only of object detection but also of its spatial position relative to the manipulator.

When evaluating the performance of tomato harvesting robots, harvesting time as well as detection accuracy, spatial localization error, and harvesting success are important indicators. A cherry-tomato harvesting robot based on an RGB-D camera and hybrid visual servo control achieved an average harvesting time of 9.40 s/product and a harvesting success rate of 96.25%. This value is numerically close to the 9.676 s complete work cycle obtained in our study. However, these two values cannot be directly equated because the experimental conditions, robot construction, camera technology, control algorithm, tomato type, and methodology for defining “cycle time” are not the same. Another study published in 2023 used a YOLOv8-based machine-vision system to detect and localize tomatoes. It reported a vision-system accuracy of 91.5%, an absolute localization error of 1.55 cm, a harvesting success rate of 83.3%, and an average harvesting time of approximately 9.5 s/product (Baratov & Mustafoqulov, 2023a, 2023b).

These results show that integrating vision, localization, and end-effector control into a unified system is an important direction in modern tomato harvesting robots. A comparison of the proposed system with selected recent studies in terms of vision systems, manipulator control, harvesting time, and reported success rates is presented in Table 3.

Table 3. A comparison of the proposed system with selected recent studies

Study	Vision system	Control / manipulator	Time, s	Success
Hybrid visual servo robot	RGB-D	Hybrid visual servo	9.40/product	96.25%
Machine-vision robot	YOLOv8n	Vision-based	≈9.5/product	83.3%
This study	Camera	Microcontroller + servo	9.676/complete cycle*	—

The 9.676 s value in this study was obtained for the complete work cycle; the times reported in the studies above are harvesting-time measures defined according to their respective methodologies. Therefore, the values in the table are not intended to provide an absolute ranking of competitiveness, but rather to compare the technical approaches of different robotic systems in a general context.

Recent research on tomato harvesting robots is moving from simple object detection toward determining the spatial state of the object and adapting manipulator motion. A 2025 study developed a tomato harvesting robot that combined global and palm vision systems with visual servo control. In such an approach, camera data are used to correct the manipulator's pose relative to the object. The study reported an average harvesting time of 12.28 s indoors and 16.55 s outdoors, with successful harvesting rates of 85.7% and 80.5%, respectively. These results show that the experimental environment has a significant effect on robot performance. Leaf occlusion, changing illumination, objects located at different distances, and the limited workspace of the manipulator complicate the control process in field or greenhouse conditions. A comprehensive review published in 2026 also emphasizes that jointly determining 3D fruit position, stem position, and maturity, as well as integrating recognition and control algorithms in real environments, are major future directions (Malyshev et al., 2026). In our study, the image acquired by the camera, object coordinates, and information from the HC-SR04 distance sensor were linked to the generation of manipulator motion. In this respect, the proposed system is based on integrating visual perception and mechanical execution into a single control chain.

One of the main advantages of the proposed system is its relatively simple hardware architecture. The system uses a camera, Raspberry Pi 4, microcontroller, HC-SR04 ultrasonic sensor, and servo motors. Such an architecture simplifies the hardware structure of the device compared with complex 3D-sensing systems.

At the same time, camera-based object detection and additional distance information from the distance sensor can be integrated into the control system. The relationship between object coordinates and manipulator servo angles is also incorporated into the practical control mechanism. In the experiment, angles of 115° and 100° were specified for $x = 15$ px and $y = 42$ px.

The experimental validation of the relationship between image coordinates and manipulator joint angles was performed using a single coordinate case. Therefore, the obtained result demonstrates the feasibility of generating a manipulator motion command from image-based object coordinates, but it does not establish a general calibration model for the entire workspace. A comprehensive calibration procedure using multiple reference points distributed across the manipulator workspace is required to quantitatively determine the mapping accuracy and positioning error. This limitation will be addressed in future experimental studies.

The experimental evaluation presented in this study was conducted with a limited number of experimental cases. Therefore, the available dataset is not sufficient to provide a statistically robust assessment of repeatability and reliability under different operating conditions. In particular, repeated measurements covering different tomato positions, distances from the manipulator, fruit sizes, and environmental conditions were not systematically performed. Consequently, statistical indicators such as standard deviation, coefficient of variation, and confidence intervals are not reported in the present study. Further experiments with a larger number of repeated trials and systematically varied operating conditions are required to statistically evaluate the repeatability, reliability, and overall performance of the proposed robotic harvesting system.

The present experimental study primarily evaluates the operating time of the robotic system, including tomato detection and harvesting time and the complete work-cycle time, together with the generation of manipulator angular positions from image coordinates. However, the available experimental dataset does not provide sufficient measurements for a complete quantitative evaluation of tomato detection accuracy, spatial localization error, end-effector positioning error, harvesting success rate, fruit damage rate, and energy consumption. Therefore, these indicators are not assigned numerical values in the present study. Future experiments will address these performance measures using repeated and systematically controlled trials. For comparison with existing studies, only performance indicators with sufficiently comparable definitions and experimental conditions should be directly compared, while differences in robot configuration, sensing technology, control strategy, tomato type, and cycle-time definition should be explicitly considered.

The experiments in this study were conducted under controlled laboratory conditions. Therefore, the effects of partial occlusion by leaves or stems, complex backgrounds, and changes in illumination were not systematically evaluated. These factors may influence tomato detection, coordinate estimation, and consequently the positioning of the manipulator. The experimental results presented in this study should therefore be interpreted within the conditions of the laboratory setup and should not be directly generalized to uncontrolled greenhouse or field environments. Further experiments under different illumination levels, background configurations, and occlusion conditions are required to quantitatively evaluate the robustness of the proposed vision-based harvesting system.

The experimental results presented in this study were obtained under controlled laboratory conditions and using a limited positioning configuration. Therefore, the demonstrated image-coordinate-to-manipulator-angle relationship should not be considered a complete generalized solution for arbitrary three-dimensional tomato positions. For practical harvesting environments, the target position may vary continuously in three-dimensional space because of differences in tomato location, distance, plant structure, and manipulator configuration.

The proposed sensing and control architecture provides a basis for integrating visual target information and distance measurements into manipulator motion control. However, generalization to dynamic three-dimensional harvesting requires multi-point spatial calibration, a comprehensive inverse-kinematic model, and continuous updating of the target position during the harvesting process. These aspects will be

investigated in future experimental studies under varying tomato positions, distances, and environmental conditions.

5. Conclusions

This study investigated the structure, control principle, and experimental results of a robotic-manipulator system designed for automatic tomato detection and harvesting. The proposed system consists of a camera, Raspberry Pi 4, microcontroller, HC-SR04 ultrasonic distance sensor, and servo motors and is based on linking the positional information of the product detected from the image with manipulator motion.

The experiments recorded 4.932 s for detecting and harvesting a red tomato and 9.676 s for the complete work cycle. For a green tomato, the corresponding values were 5.276 s and 10.586 s. The results indicate that the duration of the work cycle may be related to the product's position relative to the robot and the manipulator motion trajectory.

For an image size of 320×180 pixels, an object detected at $x = 15$ px and $y = 42$ px was associated with a base servo angle of 115° and a first-joint servo angle of 100° . This result demonstrates that object coordinates obtained from images can be used to generate the mechanical motion of the manipulator.

The results show that a vision system, distance measurement, and manipulator control can be integrated into a single robotic system using relatively simple hardware. At the same time, because the number of available experiments is limited, additional experiments are required to statistically evaluate positioning accuracy, repeatability, and harvesting success.

Future research will include experimental calibration of manipulator angles for different object coordinates, determination of the error between commanded and actual angles, evaluation of work-cycle time at different distances, improvement of the vision algorithm, and quantitative determination of the robot's energy consumption.

Declaration of Ethical Standards

The authors declare that this study complies with all applicable ethical standards.

Credit Authorship Contribution Statement

Rustam Baratov: Supervision, Validation, Formal analysis, Writing – Original Draft.

Khurshid Sattarov: Methodology, Validation, Visualization.

Almardon Mustafoqulov: Data Curation, Investigation, Resources, Writing – Review & Editing, Supervision.

Declaration of Competing Interest

The authors declared that they have no conflict of interest.

Acknowledgements

The authors would like to express their sincere gratitude to all individuals and institutions who supported the development and experimental evaluation of this research. The authors also appreciate the support and contributions that facilitated the completion of this study.

Funding

The authors received no external funding for the research, authorship, and publication of this article.

Data Availability

The experimental data supporting the findings of this study, including tomato detection and harvesting times, image coordinates, and manipulator servo angles, are available from the corresponding author upon reasonable request.

References

- 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. *Computers and Electronics in Agriculture*, 205, 107584. <https://doi.org/10.1016/j.compag.2022.107584>
- Baratov, R., & Mustafoqulov, A. (2023a). Model of field robot manipulators and sensor for measuring angular displacement of its rotating parts. *E3S Web of Conferences*, 401, 04006. <https://doi.org/10.1051/e3sconf/202340104006>
- Baratov, R., & Mustafoqulov, A. (2023b). Smart angular displacement sensor for agricultural field robot manipulators. *E3S Web of Conferences*, 386, 03008. <https://doi.org/10.1051/e3sconf/202338603008>
- Baratov, R., & Mustafoqulov, A. (2026). Research on the sensitivity of an adaptive angle displacement sensor for agricultural robot-manipulator control systems. *Agro Ilm*, 124(3), 137–138.
- Baratov, R., Mustafoqulov, A., & Namozov, S. (2025). Technical parameters of the sensor for measuring the angle displacement of rotary mechanisms of vehicles. *Uzbekhydroenergetics*, 7(1), 35–44. <https://doi.org/10.5281/zenodo.15484039>
- Droukas, L., Doulgeri, Z., Tsakiridis, N. L., Triantafyllou, D., Kleitsiotis, I., Mariolis, I. G., Giakoumis, D., Tzovaras, D. K., Kateris, D., & Bochtis, D. (2023). A survey of robotic harvesting systems and enabling technologies. *Journal of Intelligent & Robotic Systems*, 107(2), 21. <https://doi.org/10.1007/s10846-022-01793-z>
- Guduru, D., Khatri, N., Kumar, S., & Shukla, P. K. (2023). A comprehensive review of machine vision systems and artificial intelligence algorithms for the detection and harvesting of agricultural produce. *Scientific African*, 21, e01798. <https://doi.org/10.1016/j.sciaf.2023.e01798>
- Hou, G., Chen, H., Jiang, M., & Niu, R. (2023). An overview of the application of machine vision in recognition and localization of fruit and vegetable harvesting robots. *Agriculture*, 13(9), 1814. <https://doi.org/10.3390/agriculture13091814>
- Li, Y.-R., Lien, W.-Y., Huang, Z.-H., & Chen, C.-T. (2023). Hybrid visual servo control of a robotic manipulator for cherry tomato harvesting. *Actuators*, 12(6), 253. <https://doi.org/10.3390/act12060253>
- Liang, Z., Li, X., Wang, G., Wu, F., & Zou, X. (2025). Palm vision and servo control strategy of tomato picking robot based on global positioning. *Computers and Electronics in Agriculture*, Volume: Volume 237, Part B. 110668. <https://doi.org/10.1016/j.compag.2025.110668>
- Ma, P., Zhu, A., Chen, Y., Tu, Y., Mao, H., Song, J., Wang, X., Su, S., Li, D., & Dong, X. (2024). Multi objective motion planning of fruit harvesting manipulator based on improved BIT algorithm. *Computers and Electronics in Agriculture*, 227, 109567. <https://doi.org/10.1016/j.compag.2024.109567>
- Malyshev, D., Filice, L., Mirabelli, G., & Carbone, G. (2026). Robotized tomato harvesting: A review of grasping process, grippers and vision systems. *Robotics and Autonomous Systems*, 201, 105441. <https://doi.org/10.1016/j.robot.2026.105441>
- Mustafoqulov, A., Baratov, R., Radjapov, Z., Kadirov, S., & Urinov, B. (2024). Angular displacement measurement and control sensors of agricultural robot-manipulators. *BIO Web of Conferences*, 105, 03003. <https://doi.org/10.1051/bioconf/202410503003>
- Mustafoqulov, A., & Yoqubjonov, A. (2026). Robototexnika va inson: hamkorlik yoki raqobat? [Robotics and humans: Cooperation or competition?]. *Science and Education*, 7(6), 56–64. <https://cyberleninka.ru/article/n/robototexnika-va-inson-hamkorlik-yoki-raqobat>
- Mustafoqulov, A., & Yoqubjonov, A. (2026). Robototexnika va inson: hamkorlik yoki raqobat? *Science and Education*, 7(6). <https://cyberleninka.ru/article/n/robototexnika-va-inson-hamkorlik-yoki-raqobat>
- Rong, J., Wang, P., Wang, T., Hu, L., & Yuan, T. (2022). Fruit pose recognition and directional orderly grasping strategies for tomato harvesting robots. *Computers and Electronics in Agriculture*, 202, 107430. <https://doi.org/10.1016/j.compag.2022.107430>
- Thakur, A., Venu, S., & Gurusamy, M. (2023). An extensive review on agricultural robots with a focus on their perception systems. *Computers and Electronics in Agriculture*, 212, 108146. <https://doi.org/10.1016/j.compag.2023.108146>
- Wang, D., Dong, Y., Lian, J., & Gu, D. (2023). Adaptive end-effector pose control for tomato harvesting robots. *Journal of Field Robotics*, 40(3), 535–551. <https://doi.org/10.1002/rob.22146>

- Wang, L., Chu, Y., Huang, Y., & Lian, F. (2024). Enhancement on target-gripper alignment: A tomato harvesting robot with dual-camera image-based visual servoing. In *2024 IEEE International Conference on Robotics and Automation (ICRA)* (pp. 13611–13617). IEEE. <https://doi.org/10.1109/ICRA57147.2024.10610772>
- Yang, Q., Du, X., Wang, Z., Meng, Z., Ma, Z., & Zhang, Q. (2023). A review of core agricultural robot technologies for crop productions. *Computers and Electronics in Agriculture*, 206, 107701. <https://doi.org/10.1016/j.compag.2023.107701>
- Yang, Y., Han, Y., Li, S., Yang, Y., Zhang, M., & Li, H. (2023). Vision based fruit recognition and positioning technology for harvesting robots. *Computers and Electronics in Agriculture*, 213, 108258. <https://doi.org/10.1016/j.compag.2023.108258>
- Yin, Z., Li, H., Zuo, Z., & Guan, Z. (2025). Key technologies of tomato-picking robots based on machine vision. *International Journal of Agricultural and Biological Engineering*, 18(3), 247–256. <https://doi.org/10.25165/j.ijabe.20251803.8954>
- Zheng, X., Rong, J., Zhang, Z., Yang, Y., Li, W., & Yuan, T. (2024). Fruit growing direction recognition and nesting grasping strategies for tomato harvesting robots. *Journal of Field Robotics*. Volume41, Issue2. March 2024. Pages 300-313. <https://doi.org/10.1002/rob.22263>