Global Journal of Engineering and Technology Advances
eISSN: 2582-5003
Cross Ref DOI: 10.30574/gjeta
Journal homepage: https://gjeta.com/

(RESEARCH ARTICLE)

Check for updates

AGRI-VISION: INTELLIGENT ROBOTIC ARM FOR SMART TOMATO HARVESTING

Udhyami M B, Honneshwar B *, B N Sai Teja, M Rakshith Yadav and Srinivas H S

Department of Electrical and Electronics Engineering, Rajarajeswari College of Engineering, Bengaluru, Karnataka, India.
* Corresponding Author

Global Journal of Engineering and Technology Advances, 2026, 29(01), 366–372

Publication history: Received on 20 September 2026; revised on 27 October 2026; accepted on 29 October 2026

Article DOI: https://doi.org/10.30574/gjeta.2026.29.1.0311

Copyright © 2026 Author(s) retain the copyright of this article. This article is published under the terms of the Creative Commons Attribution License 4.0

ABSTRACT

Agri-Vision is a computer-vision-guided robotic system for smart tomato harvesting that uses an Arduino UNO Q, an ESP32-CAM, servo motors, and image-processing techniques. The ESP32-CAM captures images of tomato plants, which are analysed to identify candidate ripe tomatoes using colour and shape information. When a target tomato is detected, the robotic arm moves toward it, and a scissor-based cutting mechanism is activated to separate the fruit from the plant. The mobile platform uses four DC motors controlled by an L293D motor driver. A 16 × 2 LCD displays system status and operating updates, while an ISD sound module provides voice alerts during harvesting operations. The system integrates computer vision, robotics, and embedded control to demonstrate a low-cost approach to smart agriculture. Its performance and benefits should be established through testing of the physical prototype.

Keywords: Smart Agriculture, Automated Tomato Harvesting, Arduino UNO Q, ESP32-CAM, Image Processing, Colour and Shape Detection, Robotic Arm, Servo Motors, Precision Agriculture.

1 INTRODUCTION

Agriculture contributes significantly to food security and economic development. However, harvesting agricultural products remains labour-intensive, particularly for crops such as tomatoes that require careful handling. Tomatoes at different stages of maturity may be present on the same plant, making selective harvesting important. Manual harvesting requires workers to identify suitable fruits, approach them carefully, and separate them from the plant without causing unnecessary damage.

The development of embedded systems, computer vision, and robotics provides opportunities to automate selected agricultural operations. Camera-based systems can capture images of crops, while image-processing algorithms can analyse visual characteristics to identify target fruits. Robotic manipulators can then position a harvesting tool near a selected fruit.

Agri-Vision proposes a low-cost robotic harvesting system that combines an ESP32-CAM, Arduino UNO Q, servo-controlled robotic arm, cutting mechanism, and mobile platform. The camera supplies visual information, the controller

processes or receives the detection result, and the robotic arm performs the required harvesting sequence. The mobile platform supports movement between harvesting locations.

1.1 Problem Statement

Traditional tomato harvesting depends on human observation and manual handling. The process can become difficult when large numbers of plants must be inspected and harvested within a limited period. In addition, inconsistent selection of ripe fruits and careless cutting can affect crop quality.

Therefore, a system is required that can identify suitable tomatoes from camera images and coordinate a robotic harvesting mechanism. The system should use affordable components, provide clear operating feedback, and allow further development without requiring a complete redesign of the hardware.

Objectives

The objectives of Agri-Vision are:

- To capture tomato plant images using the ESP32-CAM.
- To identify tomatoes suitable for harvesting using the selected vision algorithm.
- To coordinate the detection output with the Arduino UNO Q.
- To control the robotic arm through servo motors.
- To operate a cutting mechanism for separating selected tomatoes from the plant.
- To provide mobility through DC geared motors and a motor-driver circuit.
- To display operational information and provide voice alerts.
- To evaluate the prototype through detection, movement, and harvesting tests.

1.2 Scope of the Project

The project focuses on camera-based tomato identification and robotic harvesting. Its scope includes image acquisition, fruit detection, communication between electronic modules, servo control, cutting-mechanism operation, mobile-platform control, and system-status reporting.

The prototype is intended to demonstrate the feasibility of the concept. Full commercial deployment would require additional validation under different lighting conditions, plant arrangements, fruit sizes, and outdoor environments.

2 METHODS

The proposed methodology consists of hardware development, software development, system integration, and experimental evaluation. The hardware is assembled to support camera-based detection, arm movement, cutting, and mobile operation. The software coordinates these functions through a defined sequence of operations.

The development process begins with the selection and connection of electronic components. The camera is configured to capture tomato images, and the vision software is prepared to identify the required target. A communication method is then established between the detection module and the main controller. The controller uses the received result to initiate the corresponding movement sequence.

The robotic arm is tested independently before it is integrated with the camera and mobile platform. Once the individual subsystems operate correctly, the complete prototype is tested using tomato images and, where appropriate, physical tomatoes. The collected observations are used to assess system performance.

2.1 Related Work

Previous research has investigated low-cost robotic manipulators, servo-controlled arms, computer-aided design, and robot simulation. The uploaded reference article discusses a robotic-arm development workflow involving SolidWorks, the Unified Robot Description Format. It also describes the use of motion planning and the comparison of simulated and physical robot movement.

Such work demonstrates the importance of mechanical design, joint control, trajectory planning, and experimental validation when developing robotic manipulators. Educational robotic platforms also show how accessible components can be used to investigate automation and motion control.

Agri-Vision applies the general principles of robotic manipulation to a different task: selective tomato harvesting. Rather than claiming the same CAD and simulation architecture, this project focuses on camera-based fruit identification, embedded-controller communication, servo movement, and a cutting mechanism.

Relevant literature on vision-based fruit detection, agricultural harvesting robots, tomato ripeness classification, and robotic end-effectors can be compared in terms of detection methods, hardware requirements, harvesting approaches, cost, and experimental performance.

2.2 System Design

The system is divided into five major subsystems:

- **Vision subsystem:** The ESP32-CAM captures images of tomato plants. The selected image-processing method identifies potential target tomatoes and produces a detection result.
- **Controller subsystem:** The Arduino UNO Q coordinates the system's control sequence. Depending on the implemented architecture, image processing may run on the controller or the camera module may communicate a detection result generated elsewhere.
- **Robotic-arm subsystem:** Servo motors control the arm joints to move the harvesting tool toward the selected fruit.
- **Harvesting subsystem:** A scissor-based cutter separates the tomato from the plant. Its movement must be controlled to reduce unnecessary force and avoid damaging nearby stems or fruits.
- **Mobility and feedback subsystem:** DC geared motors move the platform, while an LCD and voice module provide operating information.

The design uses modular subsystems so that camera detection, arm movement, and platform control can be tested separately before complete integration.

2.3 Tomato Detection Method

The detection method is responsible for identifying tomatoes that are suitable for harvesting. The ESP32-CAM captures images under the available lighting conditions. The images are then analysed by the selected algorithm.

A basic image-processing approach may examine colour and shape. For example, colour information can help distinguish red tomatoes from green or partially mature fruits. Shape information can help distinguish approximately round fruit regions from surrounding leaves and stems. These characteristics are useful but are not sufficient in every situation, because lighting, shadows, occlusion, and overlapping fruits can affect detection.

Alternatively, a trained object-detection model can identify tomatoes in an image. If YOLO, FOMO, or another trained model is used, the manuscript should describe the model version, dataset, training procedure, image size, and deployment platform.

After identifying a candidate fruit, the system must generate an output that the controller can interpret. This may include a detection flag or additional information about the target's location. A simple detection flag indicates that a tomato was found, whereas accurate arm positioning may require image coordinates, calibration, or another method of estimating the fruit's position.

The final harvesting decision should depend on the implemented detection criteria and should not be described as a verified ripeness classification unless testing supports that claim.

3 SYSTEM ARCHITECTURE

The proposed architecture connects image acquisition, target identification, embedded control, robotic movement, cutting, and feedback.

The operational sequence begins with image capture. The vision subsystem identifies a suitable tomato and sends the relevant result to the main controller. The controller checks the detection result and initiates the programmed movement sequence. The arm moves towards the target, and the cutting mechanism is activated when the tool is positioned appropriately. The system then updates its status and returns the arm to a safe position.

The mobile platform can move to another harvesting location when the required movement conditions are satisfied. The LCD and voice module provide feedback during operation.

3.1 System overview

The overall system consists of the following functional blocks:

Tomato plant → ESP32-CAM → Image processing or trained detection model → Arduino UNO Q → Robotic arm and cutter

The Arduino UNO Q also coordinates the mobile platform and the status-reporting modules according to the implemented control logic.

The ESP32-CAM serves as the visual input device. The controller coordinates the hardware. Servo motors position the arm, while the cutting mechanism performs the harvesting action. DC motors provide platform movement, and the display and voice module communicate system status.

The exact communication protocol and allocation of image-processing tasks must be documented according to the actual circuit and software.

3.2 Mechanical Design

The mechanical structure consists of a mobile base, robotic arm, joints, and harvesting end-effector. The mobile base supports the electronic components and provides movement using DC geared motors and wheels.

The robotic arm contains servo-controlled joints that allow the harvesting tool to change position and orientation. The number of controlled joints depends on the actual arm construction. The servo angles and movement limits must be selected according to the mechanical geometry and the available workspace.

The end-effector incorporates a scissor-based cutting mechanism. It must approach the target without colliding with neighbouring stems or fruits. The cutting operation should occur only after the arm reaches the intended position. After cutting, the arm should return to a safe position before another target is approached.

Mechanical alignment, servo torque, structural stiffness, cutter clearance, and the weight of the arm affect the overall performance. The physical prototype should be tested to identify joint limits and movement sequences that are repeatable and safe.

3.3 Hardware and Software Integration

The electronic architecture integrates the ESP32-CAM, Arduino UNO Q, servo motors, DC motor driver, LCD, and voice module. The communication interface between the camera and controller must use compatible electrical levels and a defined data format.

The controller interprets the detection output and generates the required movement commands. Servo-control signals determine the arm's joint positions, while the motor driver supplies the switching interface needed to control the mobile platform's DC motors.

The LCD can display messages such as system ready, tomato detected, harvesting in progress, harvesting completed, or error. The voice module can provide corresponding audio alerts if its playback and triggering functions are implemented.

Separate power arrangements may be necessary for the controller, camera, servos, and motors. The power supply must be sized for the peak load, and the circuit should use appropriate voltage regulation and a common reference ground where required by the interface design. High-current motors and servos should not be powered directly from a controller's GPIO pins.

4 IMPLEMENTATION AND VALIDATION

Implementation involves assembling the mobile platform, mounting the robotic arm, installing the camera, connecting the electronic modules, and developing the control software. Each subsystem is first checked independently to simplify troubleshooting.

The camera is tested to verify image capture and detection output. The arm is tested to confirm that each servo reaches the intended position. The cutter is checked separately to ensure that it opens and closes correctly. The mobile platform is tested for forward movement and turning if those functions are supported by the design.

After individual tests are completed, the subsystems are integrated. The complete sequence is tested using a known target and then under more realistic conditions. Observations should be recorded for successful detections, incorrect detections, movement errors, cutting success, operation time, and visible fruit damage

4.1 Detection Testing

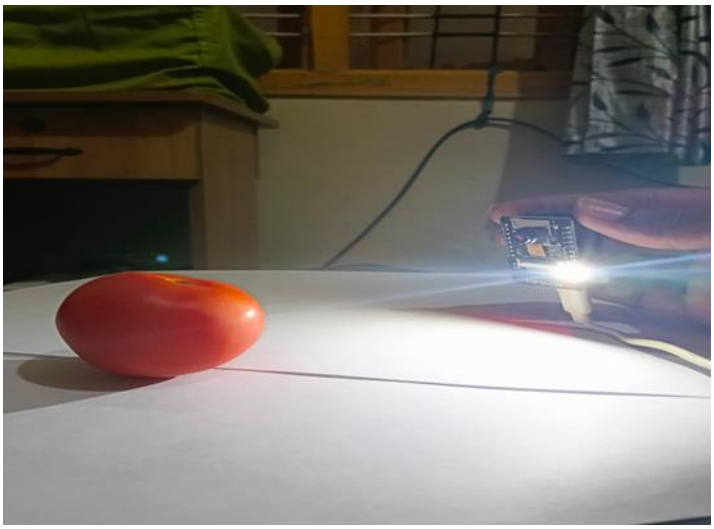


Figure 1: Tomato Detection Using the ESP32-CAM

Detection testing evaluates whether the vision subsystem correctly identifies tomatoes in camera images. The test set should include images with different fruit sizes, backgrounds, lighting conditions, and degrees of occlusion

Each test image is labelled according to the expected detection outcome. The system's output is compared with the reference labels to determine the number of correct and incorrect detections.

For a trained object detector, precision, recall, and, where appropriate, mean average precision can be reported. For a simple binary detector, accuracy can also be measured if the test set contains correctly labelled positive and negative examples.

The number of test images, the dataset composition, the detection threshold, and the evaluation method should be reported. Performance should not be estimated from a few successful demonstrations alone.

4.2 Robotic Arm and Harvesting Testing

The robotic arm is tested by commanding each joint to predefined positions. The observations should include whether the arm reaches the target area, whether movement is repeatable, and whether the arm returns to its intended starting position.

The harvesting test evaluates whether the cutter can separate a tomato under the tested conditions. Each trial should record whether the fruit was successfully detached, whether the cutter contacted the intended stem, and whether visible damage occurred.

The arm and cutter should be tested at different target positions when possible. The system should also be tested for false detections, missing detections, and communication failures. If no target is detected, the controller should not trigger an unplanned cutting operation.

5 RESULTS AND DISCUSSION

The results should be based on measurements collected from the physical prototype. The following tables provide a structure for recording the experimental findings.

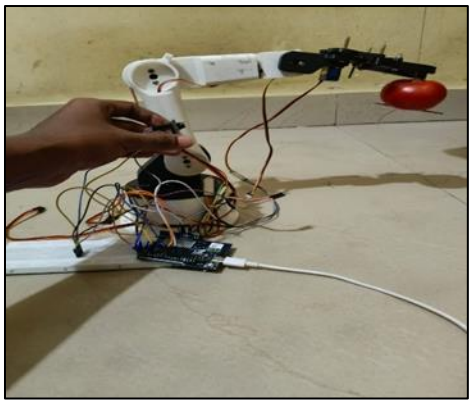


Figure 2: Robotic Arm Testing for Tomato Cutting



Figure 3: System Status Display

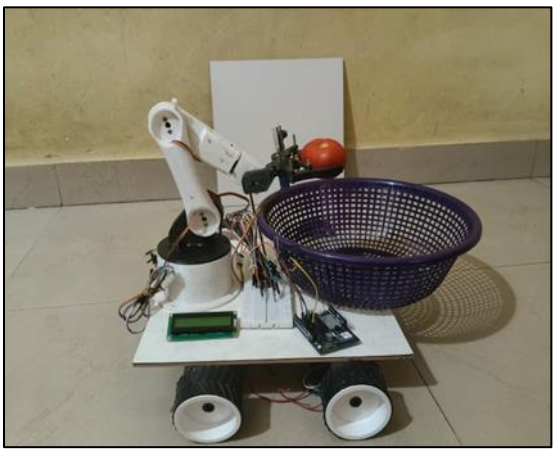


Figure 4: Final Agri-Vision Prototype

6 CONCLUSION AND FUTURE WORK

Agri-Vision proposes a vision-guided robotic system for supporting selective tomato harvesting. The design combines an ESP32-CAM, Arduino UNO Q, servo-controlled robotic arm, cutting mechanism, mobile platform, and status-reporting modules. Its purpose is to connect camera-based target identification with controlled physical harvesting.

The proposed architecture provides a modular approach in which image capture, detection, arm movement, cutting, and mobility can be developed and tested independently. Its potential benefits include reduced manual effort, more consistent target selection, and a practical platform for exploring embedded robotics in agriculture.

The actual effectiveness of the system must be established through physical experiments. Detection accuracy, harvesting success rate, operation time, repeatability, and fruit damage should be measured before making quantitative performance claims.

Future work may include improving detection under changing lighting conditions, adding depth or distance estimation, calibrating image coordinates to arm movement, implementing more precise trajectory control, and developing safer cutter actuation. Additional improvements could include obstacle detection, autonomous navigation, battery monitoring, and trials in real agricultural environments.

STATEMENTS ON COMPLIANCE WITH ETHICAL STANDARDS

Acknowledgment

The authors thank the Department of Electrical and Electronics Engineering, Rajarajeswari College of Engineering, for the support and facilities provided during the development of the Agri-Vision prototype.

Disclosure of conflict of interest

The authors declare that they have no competing interests related to the design, development, or reporting of this work.

REFERENCES

[1] Peter P. Ling, Reza Ehsani, K.C. Ting, Yu-Tseh Chi, Nagarajan Ramalingam, Michael H. Klingman, and Craig Draper, “Sensing and End-Effector for a Robotic Tomato Harvester,” 2004, Ottawa, Canada August 1 - 4, 2004, 2004.

[2] C. dela Peña, “Thinking Through the Tomato Harvester,” Boom, vol. 3, no. 1, pp. 34–40, 2013.

[3] F. Taqi, F. Al-Langawi, H. Abdulraheem, and M. El-Abd, “A cherry tomato harvesting robot,” 2017 18th International Conference on Advanced Robotics (ICAR), Jul. 2017. \

[4] R. Yenorkar and U. M. Chaskar, “GUI Based Pick and Place Robotic Arm for Multipurpose Industrial Applications,” 2018 Second International Conference on Intelligent Computing and Control Systems (ICICCS), Jun. 2018.

[5] [AUTHOR QUERY — Please verify this book reference and supply the author's correct name, exact title, publisher, year, and other required publication details before publication.]

[6] M. M. Hasan, K. N. W. Uddin, et al., "Smart Agriculture Robotic System Based on Internet of Things to Boost Crop Production," 2021, doi: 10.1109/ICREST51555.2021.9331132.

[7] D. Bini, D. Pamela, and S. Prince, "Machine Vision and Machine Learning for Intelligent Agrobots: A review," 2020, doi: 10.1109/ICCCSP49186.2020.9315242.

[8] A. Kumar, A. Raj, K. Singh, et al., "Precision Farming with Versatile Robotic Systems," 2023, doi: 10.1109/ICAC3N60023.2023.10530777.

[9] J. Tussupov, M. Yessenova, et al., "Analysis of Formal Concepts for Verification of Pests and Diseases of Crops Using Machine Learning Methods," *IEEE Access*, 2024, doi: 10.1109/IEEEAccess.2024.3411306.

[10] M. Ayaz, M. Ammad-Uddin, et al., "Internet-of-Things (IoT)-Based Smart Agriculture: Toward Making the Fields Talk," *IEEE Access*, 2019, doi: 10.1109/ACCESS.2019.2932662.