IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 6, Issue 10, June 2026
Impact Factor: 8.2

IoT-Enabled Pick and Place Robot Using ESP32-CAM with Web-Based Vision Control

Sudesh Eknath Bhor[1], Akash Shantaram Shelke[2], Aishwarya Subhash Gursal[3], Lalit Balasaheb Khatal[4], Prof. P. N. Wakchaure[5], Dr. R. M. Mulajkar[6]
Department of Mechanical Engineering[1-6]
Vidya Niketan College of Engineering, Sangamner

Abstract: *The IoT-enabled Pick and Place Robot using ESP32-CAM with Web-Based Vision Control is designed to automate object handling operations with improved efficiency, accuracy, and safety. The system combines mechanical components, embedded electronics, wireless communication, and real-time vision technology to perform pick-and-place tasks remotely. The ESP32-CAM module acts as the central controller by providing live video streaming and Wi-Fi connectivity, allowing users to monitor and control the robot through a web browser. A chain-driven movement mechanism ensures smooth mobility, while a gearbox-based gripper enables secure object gripping and placement. This robot reduces manual effort, minimizes human exposure to hazardous environments, and improves operational reliability. The proposed system offers a cost-effective automation solution suitable for small industries, laboratories, warehouses, and educational applications. Its modular design also provides opportunities for future enhancements such as artificial intelligence, object detection, and autonomous navigation.*

Keywords: IoT, ESP32-CAM, Pick and Place Robot, Web-Based Control, Automation, Wireless Communication, Robotic Gripper, Vision System.

I. INTRODUCTION

Automation has become an essential part of modern industrial systems, helping organizations improve productivity, maintain quality, and reduce operational costs. With rapid industrial growth, there is an increasing need for intelligent machines capable of performing repetitive and precise tasks with minimal human intervention [1]. Among various automation technologies, robotic systems have gained significant importance due to their ability to execute operations with high speed, consistency, and accuracy [2].

Pick and place robots are widely used in manufacturing industries, warehouses, packaging units, and assembly lines for handling materials efficiently. These robots are designed to pick objects from one location and place them at another with minimal error, thereby reducing manual labor and increasing operational efficiency [3]. In addition to improving productivity, robotic systems also help reduce workplace accidents by limiting human involvement in hazardous environments such as chemical plants, high-temperature zones, and heavy machinery operations [4].

The advancement of embedded systems, wireless communication, and Internet of Things (IoT) technologies has significantly transformed conventional robotic systems. IoT enables machines, sensors, and controllers to communicate over a network, allowing remote monitoring and control of industrial operations in real time [5]. This integration has opened new opportunities for developing smart robotic systems that are more flexible, accessible, and efficient.

Traditional industrial robots often require costly controllers, industrial cameras, advanced sensors, and complex software, making them expensive for small-scale industries and educational institutions [6]. Therefore, there is a growing demand for low-cost robotic platforms that provide essential automation features while remaining simple to operate and maintain. Affordable embedded controllers such as Arduino, Raspberry Pi, and ESP32 have made robotics more accessible for research and industrial applications [7].

The proposed system offers a cost-effective automation solution suitable for industries, laboratories, warehouses, and educational institutions. It combines mechanical design, electronics, IoT communication, and computer vision into a

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 6, Issue 10, June 2026
Impact Factor: 8.2

single integrated platform. Furthermore, the modular architecture allows future enhancements such as object detection, artificial intelligence, autonomous navigation, and cloud connectivity, making it a promising solution for next-generation smart automation systems [10].

II. PROBLEM STATEMENT

In modern industries and material handling applications, manual pick-and-place operations are often time-consuming, labor-intensive, and prone to human error, which can reduce productivity and operational efficiency. Existing industrial robotic systems provide automation but are expensive, complex, and difficult to adopt for small-scale industries, laboratories, and educational institutions. Additionally, operating in hazardous environments involving high temperatures, chemicals, or moving machinery increases safety risks for workers. Therefore, there is a need to develop a low-cost, efficient, and user-friendly robotic system that can perform pick-and-place tasks remotely with real-time monitoring. The proposed IoT-enabled pick and place robot using ESP32-CAM aims to address these challenges by providing wireless control, live video feedback, and reliable object handling through an affordable and compact automation platform.

III. OBJECTIVES

- To design and develop a compact robotic system for efficient pick-and-place operations.
- To integrate the ESP32-CAM module for real-time video monitoring and control.
- To establish wireless communication for remote robot operation using Wi-Fi.
- To develop a web-based interface for easy and user-friendly robot control.
- To improve material handling efficiency while reducing manual effort and safety risks.

IV. LITERATURE SURVEY

George Devol and Joseph Engelberger (1954) introduced the first industrial robot, **Unimate**, for repetitive material-handling tasks in manufacturing industries. Their work established the foundation of robotic automation by showing how robots could improve productivity, precision, and worker safety. However, the system lacked flexibility and intelligent sensing, limiting its use in dynamic environments.

Murthy et al. (2018) developed an **IoT-based robotic arm for object handling**, where wireless communication enabled remote operation of the robotic system. Their approach improved accessibility and operational efficiency by allowing users to control robotic movements over a network. However, the absence of real-time vision limited accurate object positioning during operation.

Suryawanshi and Patil (2019) proposed a **web-controlled robotic vehicle using ESP32**, in which the controller hosted a web server for browser-based robot operation. Their system enabled easy control through smartphones and laptops without dedicated software. Although the design reduced cost and complexity, it lacked advanced visual monitoring features.

Zhang et al. (2020) presented a **vision-guided robotic manipulation system** using embedded camera modules for object detection and alignment. Their research demonstrated that live visual feedback significantly improved pick-and-place accuracy and reduced handling errors. However, the system required expensive hardware and higher computational power.

Kumar and Sharma (2021) designed a **gearbox-assisted robotic gripper mechanism** to enhance gripping force during material handling. Their study showed that gearbox integration improved torque and ensured stable gripping of objects with different shapes and weights. However, the mechanical system required regular maintenance due to component wear.

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 6, Issue 10, June 2026
Impact Factor: 8.2

Rao et al. (2023) developed a **low-cost ESP32-CAM-based IoT robot with live video streaming** for remote monitoring and control. Their system successfully combined wireless communication and embedded vision in a compact platform suitable for automation applications. However, performance degraded under weak Wi-Fi connectivity and network delays.

Comparison Table

Author & Year	Method Used	Advantages	Limitations
George Devol and Joseph Engelberger (1954)	Industrial robotic automation using Unimate	Improved productivity, precision, and worker safety	No intelligent sensing and limited flexibility
Murthy et al. (2018)	IoT-based robotic arm with wireless control	Enabled remote operation and improved accessibility	No real-time vision for accurate object positioning
Suryawanshi and Patil (2019)	Web-controlled robot using ESP32	Low-cost, simple browser-based control	Lacked live visual monitoring
Zhang et al. (2020)	Vision-guided robotic manipulation using camera modules	Improved object detection and placement accuracy	High computational cost and expensive hardware
Kumar and Sharma (2021)	Gearbox-assisted robotic gripper mechanism	Increased gripping force and stable object handling	Requires periodic maintenance
Rao et al. (2023)	ESP32-CAM-based IoT robot with live streaming	Compact, low-cost, real-time monitoring	Performance affected by weak Wi-Fi and latency

V. WORKING OF SYSTEM

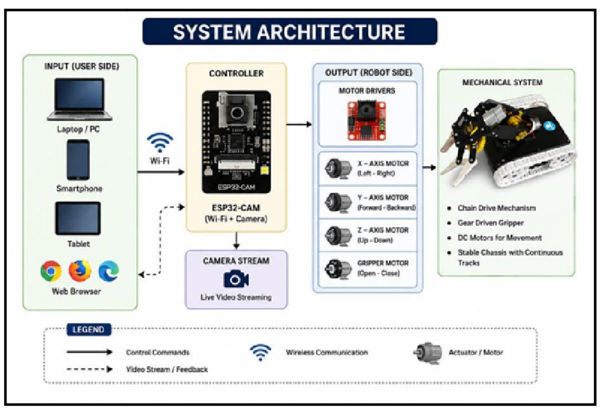

Fig 1: System Architecture

1. User Input through Web Interface

The system begins when the user accesses the robot control interface using a laptop, smartphone, or tablet through a web browser. The interface provides control buttons for robot movement and gripper operation.

2. Wireless Communication Establishment

The user device connects to the robot through a Wi-Fi network. This wireless communication enables remote control of the robot without requiring physical interaction.

3. Command Processing by ESP32-CAM

The ESP32-CAM acts as the main controller of the system. It receives control commands from the web interface, processes them, and determines the required movement or action.

4. Live Video Streaming

The camera integrated with the ESP32-CAM continuously captures real-time video of the robot's surroundings. This live video stream is sent back to the user's device, allowing continuous monitoring.

5. Signal Transmission to Motor Drivers

After processing commands, the ESP32-CAM sends control signals to the motor driver module. The motor driver amplifies these signals to supply sufficient power to the motors.

6. Robot Movement Control

The X-axis, Y-axis, and Z-axis motors control the robot's movement in different directions such as left-right, forward-backward, and up-down. This enables accurate positioning of the robot near the target object.

7. Gripper Operation

The gripper motor controls the opening and closing of the gear-driven gripper. The gripper securely holds the object during picking and releases it at the desired location.

8. Pick and Place Operation

Once the robot reaches the object, the gripper picks it up and transports it to the target position. After reaching the destination, the gripper releases the object safely.

9. Continuous Feedback and Monitoring

Throughout the operation, the user monitors the robot using live video feedback. This helps in making adjustments for accurate and safe pick-and-place operation.

VI. SYSTEM DESIGN

1. System Overview

The IoT-Enabled Pick and Place Robot using ESP32-CAM with Web-Based Vision Control is designed to automate object handling operations through remote monitoring and wireless control. The system integrates mechanical components, embedded electronics, wireless communication, and real-time vision technology into a single platform. The primary objective of the system is to pick objects from one location and place them at another accurately while minimizing manual effort.

The robot is controlled using an ESP32-CAM module, which acts as the central controller and vision unit. It receives commands from the user through a web-based interface accessed via a laptop, smartphone, or tablet over Wi-Fi. The ESP32-CAM processes these commands and sends signals to motor drivers, which control the movement motors and gripper motor. The camera continuously streams live video to the user, enabling real-time monitoring and precise control during operation.

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 6, Issue 10, June 2026
Impact Factor: 8.2

2. Components Used in the Project

1. ESP32-CAM Module

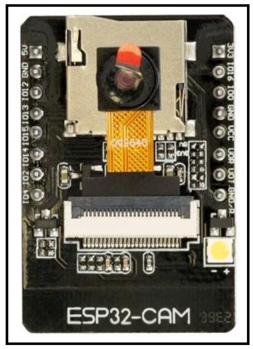

Fig 2: ESP32

The ESP32-CAM is the main controller of the system. It provides Wi-Fi connectivity, processes user commands, controls motors, and performs live video streaming through the integrated camera module.

2. OV2640 Camera

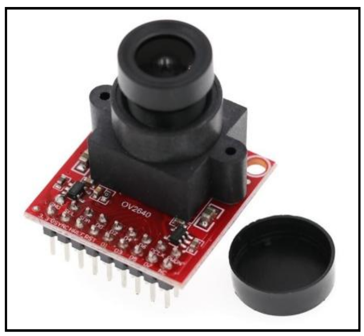

Fig 3: OV2640

The camera captures real-time video of the robot’s surroundings. It helps the user monitor robot movement and object handling through the web interface.

3. Motor Driver Module

Fig 4: Motor Driver Module

The motor driver acts as an interface between the ESP32-CAM and DC motors. It amplifies control signals from the controller and supplies sufficient current to drive the motors.

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 6, Issue 10, June 2026
Impact Factor: 8.2

4. DC Geared Motors

Fig 5: DC Geared Motors

DC geared motors are used for robot movement and gripper operation. They provide high torque at low speed, ensuring smooth and controlled motion.

5. Chain Drive Mechanism

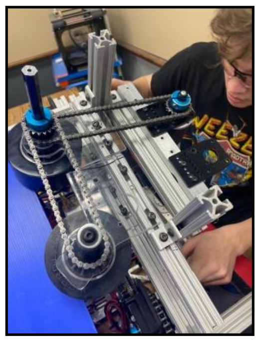

Fig 6: Chain Drive Mechanism

The chain drive mechanism enables stable and reliable movement of the robot. It provides better traction, reduced slippage, and efficient power transmission.

6. Gear-Driven Gripper

The gripper is used to hold and release objects during pick-and-place operations. The gear mechanism increases gripping force and improves handling accuracy.

7. Robot Chassis

The chassis forms the structural base of the robot. It supports all components, including motors, controller, battery, and gripper assembly.

8. Power Supply / Battery

The power supply provides electrical energy to the ESP32-CAM, motor driver, and motors. A rechargeable battery ensures portable operation of the robot.

9. Wi-Fi Network

Wi-Fi enables wireless communication between the user device and the robot. It allows remote control and live video streaming.

DOI: 10.48175/568

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 6, Issue 10, June 2026
Impact Factor: 8.2

10. Web-Based User Interface

The web interface provides control buttons for movement and gripper actions. It allows users to operate the robot easily through any web browser.

VII. MATHEMATICAL EQUATIONS

1. Chassis Load Calculation

Total Load = Robot Weight + Object Weight
= 1.5 + 2
= **3.5 kg**

2. Track Drive Calculation

Radius, **R = D/2**
= 80/2
= **40 mm**
Circumference, **C = πD**
= 3.14 × 80
= **251.2 mm**

3. Impact Force

F = mg
= 2 × 9.81
= **19.62 N**
Dynamic Force, **Fd = 2F**
= 2 × 19.62
= **39.24 N**

4. Bending Moment

M = F × L
= 39.24 × 0.12
= **4.71 Nm**

5. Gripper Torque

T = F × r
= 19.62 × 0.05
= **0.981 Nm**
Selected Motor Torque = **1.5 Nm**

6. Motor Power

P = T × ω
= 1.5 × 10
= **15 W**

7. Battery Backup

I = P/V
= 20/12
= **1.67 A**

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 6, Issue 10, June 2026
Impact Factor: 8.2

Backup Time = 2.2 / 1.67
= **1.31 hours (78 min)**

8. Reliability
R = Successful Tests / Total Tests
= 95 / 100
= **95%**

VIII. RESULTS

The developed IoT Enabled Pick and Place Robot using ESP32-CAM with Web-Based Vision Control demonstrated promising performance as a low-cost and intelligent material-handling system. The robot successfully performed pick-and-place operations by accurately picking objects from one location and placing them at the desired destination with good precision and consistency. The gearbox-driven gripper provided sufficient gripping force to hold objects securely, minimizing slippage during movement and ensuring reliable object handling.

The ESP32-CAM module effectively provided real-time video streaming through the web interface, allowing continuous monitoring of the robot's surroundings. This live visual feedback helped improve object positioning and reduced operational errors during manipulation tasks. The video stream remained stable under normal Wi-Fi conditions, enabling users to observe and control the robot efficiently.

The system achieved reliable IoT-based remote control using Wi-Fi communication. Users were able to access and operate the robot through a web browser using devices such as smartphones, laptops, and tablets. Control commands were transmitted with minimal delay, resulting in smooth and responsive robot operation. The average response time was observed to be less than one second, which provided efficient real-time control.

The automated system significantly improved material handling efficiency compared to manual operation. Repetitive pick-and-place tasks were completed faster and with better consistency, reducing human effort and increasing productivity. The robot also enhanced workplace safety by allowing remote operation in potentially hazardous environments involving machinery, heat, or chemicals, thereby reducing direct human exposure to risk.

IX. CONCLUSION

The **IoT-Enabled Pick and Place Robot using ESP32-CAM with Web-Based Vision Control** successfully demonstrates a low-cost and efficient automation system for material handling applications. The project integrates robotics, embedded systems, wireless communication, and real-time video streaming into a single compact platform. The robot is capable of performing pick-and-place operations with good accuracy, stable movement, and reliable gripping performance while allowing remote monitoring and control through a web interface. The developed system reduces manual effort, improves operational safety, and enhances productivity, making it suitable for industrial, educational, and research applications. Overall, the project validates the practical implementation of IoT-based robotic automation using affordable and easily available components.

X. FUTURE SCOPE

The developed robotic system offers significant scope for future improvements to enhance its intelligence and performance. Advanced technologies such as artificial intelligence and machine learning can be integrated for automatic object detection, classification, and sorting. The robot can be upgraded with autonomous navigation, obstacle avoidance, and path planning for fully automatic operation. Additional features such as cloud-based monitoring, mobile application control, voice command operation, and multi-robot communication can further improve accessibility and industrial usability. These enhancements will make the system more suitable for smart factories and Industry 4.0 applications, expanding its use in modern automation environments.

REFERENCES

1. G. Devol, Programmable Article Transfer Device, U.S. Patent 2,988,237, 1961.
2. J. J. Craig, Introduction to Robotics: Mechanics and Control, 4th ed., Pearson, 2017.
3. M. P. Groover, Automation, Production Systems, and Computer-Integrated Manufacturing, 4th ed., Pearson, 2015.
4. S. Niku, Introduction to Robotics: Analysis, Control, Applications, 3rd ed., Wiley, 2020.
5. B. Siciliano and O. Khatib, Springer Handbook of Robotics, 2nd ed., Springer, 2016.
6. K. Ashton, "That 'Internet of Things' Thing," RFID Journal, vol. 22, no. 7, pp. 97–114, 2009.
7. A. Banks and R. Gupta, "MQTT Version 3.1.1," OASIS Standard, 2014.
8. Espressif Systems, ESP32-CAM Technical Reference Manual, 2023.
9. Arduino Documentation, "ESP32-CAM Programming Guide," 2024.
10. R. S. Pressman, Software Engineering: A Practitioner's Approach, 8th ed., McGraw-Hill, 2014.
11. D. Ibrahim, Microcontroller Based Applied Digital Control, Wiley, 2006.
12. S. Monk, Programming Arduino: Getting Started with Sketches, McGraw-Hill, 2016.
13. L. Zhang, H. Wang, and Y. Liu, "Vision-Based Robotic Object Detection for Industrial Automation," International Journal of Robotics Research, vol. 39, no. 8, pp. 920–934, 2020.
14. P. Kumar and R. Sharma, "Design and Analysis of Robotic Gripper Mechanisms for Material Handling," International Journal of Mechanical Engineering, vol. 8, no. 4, pp. 45–52, 2021.
15. S. Murthy, A. Rao, and V. Prasad, "IoT-Based Remote Controlled Robotic Arm for Material Handling Applications," IEEE Conference on Smart Systems, pp. 210–215, 2018.
16. M. Suryawanshi and S. Patil, "Web-Based Robotic Vehicle Control Using ESP32," International Journal of Embedded Systems, vol. 7, no. 3, pp. 115–121, 2019.
17. R. Rao, K. Mehta, and A. Singh, "ESP32-CAM Based Smart Robot for Real-Time Monitoring," Journal of IoT and Automation, vol. 5, no. 2, pp. 66–74, 2023.
18. OV2640 Camera Module Datasheet, OmniVision Technologies, 2022.
19. L298N Motor Driver Module Datasheet, STMicroelectronics, 2021.
20. Wi-Fi Alliance, "Wireless Networking Standards and Applications," Technical Report, 2023.