RESEARCHERID
THOMSON REUTERS
[Inder et al.,11 (5), May 2026]

ISSN:2455-9679
SJIF Impact Factor: 6.190

IJRTSM

INTERNATIONAL JOURNAL OF RECENT TECHNOLOGY SCIENCE & MANAGEMENT

"EMBEDDED SYSTEM APPROACH TO VOICE-OPERATED DELIVERY ROBOT USING ESP32-S3 AND CAMERA INTEGRATION"

Inder Patel[1], Aryan nagle[2], Khushi Chouksey[3], Bhumika Pal[4], Sneha Badode[5], Yogesh Shrivastaval[6]

[1-5]*Department of Electronics & Communication Engineering, Sagar Institute of Science, Technology & Research (SISTec-R), Bhopal, Madhya Pradesh, India*

ayushinder877@gmail.com

ABSTRACT

Indoor transportation of medicines, documents, laboratory samples, food trays, and small packages is a repetitive activity in hospitals, educational institutions, offices, and small industrial environments. This paper presents the design and implementation of a low-cost autonomous indoor delivery robot based on the ESP32-S3 N16R8 microcontroller. The proposed system combines offline voice recognition, inertial heading estimation, ultrasonic obstacle detection, differential-drive motor control, grid-based path planning, and a finite state machine. A VC02 offline voice-recognition module is used to receive destination commands, while a BNO085 9-DOF inertial measurement unit provides heading information. An HC-SR04 ultrasonic sensor provides forward obstacle detection, and BTS7960 motor drivers control the differential-drive motors. The navigation system uses an A* pathfinding algorithm enhanced with a clearance map and turn-cost penalty to generate paths that reduce unnecessary turns and wall-hugging behaviour. A non-blocking finite state machine manages navigation, obstacle waiting, delivery, return-to-home, and fault conditions. The robot operates on a predefined 50 × 150 occupancy grid and performs all major computation locally without cloud connectivity or an external computer. Experimental testing over 50 delivery cycles achieved a 94% delivery success rate, A* path-planning times of approximately 180–240 ms, voice-recognition accuracy of 93%, and battery operation of approximately 95 minutes. The total bill of materials was reported as ₹5,840. The results demonstrate the feasibility of implementing useful autonomous indoor logistics using a low-cost embedded platform.

Keywords: *Autonomous Mobile Robot, Indoor Delivery Robot, ESP32-S3, A* Algorithm, Voice Recognition, BNO085 IMU, HC-SR04, Finite State Machine, Obstacle Detection, Embedded Robotics.*

I. INTRODUCTION

The rapid development of embedded systems, sensors, and autonomous navigation algorithms has created new opportunities for low-cost mobile robotics. Indoor delivery is an important application because the movement of small objects between fixed locations is repetitive, time-consuming, and often performed manually. Hospitals, offices, laboratories, educational institutions, hotels, and warehouses can benefit from automated transportation of documents, medicines, samples, and lightweight packages. The uploaded project identifies manual indoor logistics as an important source of inefficiency and proposes a compact autonomous robot capable of performing delivery without continuous human teleoperation.

Commercial autonomous mobile robots generally employ technologies such as LiDAR, depth cameras, powerful processors, and sophisticated robotics frameworks. Although these technologies provide high navigation capability, their cost and implementation complexity can make them unsuitable for small institutions. The project therefore focuses on an alternative

ISSN:2455-9679
SJIF Impact Factor: 6.190

architecture based on a single ESP32-S3 microcontroller, an offline voice-recognition module, an IMU, and a low-cost ultrasonic sensor.

The proposed robot accepts a spoken destination command through the VC02 module. The command is converted into a destination code and transmitted to the ESP32-S3. The controller then calculates an appropriate route using the A* algorithm over a predefined occupancy grid. During navigation, the BNO085 provides heading information while the HC-SR04 continuously monitors the forward region for obstacles. If an obstacle is detected within 25 cm, the robot stops and enters a n obstacle-waiting state. After three consecutive clear readings, navigation resumes. If the obstacle remains for 30 seconds, the robot enters a safe HALT state.

The major objectives of the system are to develop a compact differential-drive robot, provide offline voice-based destination selection, implement clearance-aware A* path planning, maintain responsive non-blocking firmware, provide reliable obstacle handling, and keep the overall hardware cost below ₹6,000.

The proposed system has applications in healthcare, education, offices, hotels, warehouses, small factories, and assisted-living environments. It also contributes to SDG 3, SDG 9, SDG 11, and SDG 12 through contactless delivery, accessible robotics, energy-efficient indoor mobility, and responsible resource utilization.

II. REVIEW OF LITERATURE

Autonomous indoor navigation has been an established research area for several decades. Early mobile robots demonstrated the feasibility of autonomous perception, planning, and movement, while modern autonomous mobile robots have introduced more sophisticated sensors and computational platforms. Commercial systems commonly employ LiDAR, cameras, simultaneous localization and mapping (SLAM), and powerful processors to achieve autonomous navigation.

The A* algorithm remains an important path-planning technique for grid-based navigation. It combines the actual movement cost with a heuristic estimate of the remaining distance to the destination. Its relatively simple implementation and deterministic behaviour make it suitable for resource-constrained embedded systems. The project applies A* to a 50 × 150 occupancy grid and uses Manhattan distance as the heuristic.

The proposed approach extends conventional A* by introducing a clearance-based penalty. A clearance map is generated using multi-source breadth-first search, assigning each free grid cell a distance from the nearest obstacle. This information is incorporated into the path cost so that the planner avoids routes that pass too close to obstacles. A turn-cost penalty is also introduced to discourage excessive direction changes. The project reports that tuning the clearance weight from 600 to 80 and increasing the turn penalty from 8 to 40 substantially reduced zigzag behaviour and produced longer straight sections.

For orientation estimation, the project uses the BNO085 9-DOF IMU. The sensor performs internal sensor fusion and provides quaternion-based orientation information, reducing the computational burden on the host microcontroller. The HC-SR04 ultrasonic sensor is used as a low-cost forward obstacle detector. Although ultrasonic sensing is relatively simple, it can provide effective short-range detection in controlled indoor environments.

Offline voice recognition is another important aspect of the proposed system. Instead of depending on cloud-based speech services, the VC02 module recognizes predefined keywords and provides command information to the microcontroller. This makes the robot usable without continuous Internet access.

The literature and project background therefore indicate a clear design opportunity: combine low-cost embedded hardware, offline voice control, A* path planning, inertial navigation, and simple obstacle detection into a compact indoor delivery platform. The proposed work integrates these elements into single-microcontroller architecture.

III. PROBLEM IDENTIFICATION

Despite significant progress in autonomous mobile robots, three major problems restrict their adoption in small institutions.

1. High Cost Commercial autonomous mobile robots frequently require LiDAR, depth cameras, high-performance

ISSN:2455-9679
SJIF Impact Factor: 6.190

processors, and complex supporting hardware. Such systems can be financially difficult for small clinics, laboratories, schools, and start-ups. The proposed project addresses this issue with a bill of materials of approximately ₹5,840.

2. High System Complexity Modern ROS-based autonomous robots require mapping, calibration, localization, navigation configuration, and specialized technical knowledge. The proposed system instead uses a predefined indoor map and an embedded controller, reducing deployment complexity.

3. Dependence on Internet Connectivity

Many voice-controlled systems depend on cloud-based speech-recognition services. This introduces network dependency and can create privacy and reliability concerns. The proposed robot uses offline keyword recognition through the VC02 module.

4. Navigation and Obstacle Handling

A robot operating indoors must follow a predetermined route while responding safely to unexpected obstacles. A static path alone is insufficient because people, chairs, boxes, and other objects may temporarily block the route. The proposed system therefore combines planned navigation with real-time ultrasonic obstacle detection.

5. Positioning Accuracy

The current robot estimates forward movement using calibrated travel speed and elapsed time rather than closed-loop wheel odometry. Consequently, wheel slip and variations in motor speed can affect positioning accuracy. The project identifies this as one of its main limitations.

IV. Solution

The proposed solution is a low-cost, voice-activated autonomous indoor delivery robot based on the ESP32-S3 N16R8 microcontroller.

The system architecture consists of five major subsystems:

1. **Voice Input:** The VC02 module recognizes predefined spoken commands and transmits a destination code through UART.
2. **Navigation and Processing:** The ESP32-S3 executes the A* path planner and controls the finite state machine.
3. **Heading Estimation:** The BNO085 IMU provides yaw information for directional control.
4. **Obstacle Detection:** The HC-SR04 detects objects in front of the robot and triggers emergency stopping when an obstacle is detected within 25 cm.
5. **Actuation and User Interface:** BTS7960 motor drivers control the differential-drive motors, while a 16 × 2 I²C LCD displays robot status.

The complete system operates from an 11.1 V, 2200 mAh battery pack. A buck converter supplies the required lower voltage to the logic and sensor components. The hardware configuration includes an ESP32-S3 N16R8, BNO085 IMU, HC-SR04 ultrasonic sensor, VC02 voice module, two BTS7960 motor drivers, two DC geared motors, caster wheel, LCD, battery, and chassis.

The navigation environment is represented by a 50 × 150 occupancy grid. Each cell represents approximately 2 cm × 2 cm. Obstacles are represented as occupied cells, while free cells can be traversed by the robot. Before path planning, a clearance map is generated using multi-source BFS. A* then uses movement cost, clearance penalty, and turn penalty to determine the preferred route.

The robot follows a complete delivery cycle:

*Voice command → Destination selection → A path planning → Motion execution → Obstacle monitoring → Target arrival → Delivery waiting period → Return path planning → Return home.**

V. Proposed Methodology

ISSN:2455-9679
SJIF Impact Factor: 6.190

The methodology follows an embedded-systems design approach in which the system is divided into input, processing, and output layers. The input layer consists of voice, ultrasonic, and IMU sensing. The processing layer consists of A* path planning and finite state machine control. The output layer consists of motor control and LCD feedback.

5.1 System Initialization

At startup, the ESP32-S3 initializes the SPI, I²C, UART, GPIO, motor drivers, IMU, LCD, and voice-recognition module. The occupancy grid is loaded into memory and the clearance map is generated.

5.2 Voice Command Recognition

The user speaks a predefined command. The VC02 offline voice-recognition module recognizes the command and sends a corresponding code to the ESP32-S3 through UART. When the destination command is received while the robot is at the home state, path planning begins.

5.3 Grid-Based Environment Representation

The environment is represented as an occupancy grid:

- 0 = free cell
- 1 = obstacle
- Home and target locations are defined during map creation.

The 50 × 150 grid contains 7,500 cells and is stored in the microcontroller memory.

5.4 Clearance Map Generation

Before executing A*, a multi-source BFS starts from all obstacle cells. The algorithm propagates outward and calculates the Manhattan distance of each free cell from the nearest obstacle.

The resulting clearance information is used to penalize paths that pass too close to obstacles.

5.5 A* Path Planning

The proposed cost function is based on three components:

Cost = movement cost + clearance penalty + turn penalty

The final tuned parameters are:

Parameter	**Value**
Grid size	50 × 150
Movement cost	1
Clearance weight	80
Maximum clearance cap	5
Turn penalty	40
Obstacle-cell penalty	5000
Heuristic	Manhattan distance

These parameters were experimentally tuned to reduce oscillatory motion and encourage longer straight paths.

ISSN:2455-9679
SJIF Impact Factor: 6.190

5.6 Motion Control

The robot uses differential-drive motion. The path is converted into forward movement and 90° turning operations. The BNO085 provides heading information, while the motor drivers control the left and right motors.

5.7 Obstacle Detection and Recovery

The HC-SR04 is polled periodically during forward movement. If an obstacle is detected within 25 cm, the robot immediately stops and enters the obstacle-waiting state.

To avoid false triggering due to sensor noise, the robot requires three consecutive clear readings before resuming movement. If the obstacle remains for 30 seconds, the system enters a permanent HALT state requiring manual reset.

5.8 Delivery and Return

After reaching the destination, the robot remains stationary for 60 seconds to allow delivery. A new A* path is then calculated from the destination back to the home location. After reaching home, the robot displays an arrival message and returns to the idle state.

Algorithm

Algorithm: Voice-Activated Autonomous Delivery

1. Start the robot.
2. Initialize sensors, motors, LCD, UART, SPI, and I²C.
3. Load the predefined occupancy grid.
4. Generate the clearance map using multi-source BFS.
5. Calibrate the IMU heading.
6. Enter the idle-at-home state.
7. Wait for a valid voice command.
8. Decode the destination.
9. Run A* from the current position to the destination.
10. If no path exists, enter HALT state.
11. Otherwise, execute the planned path.
12. Continuously monitor the ultrasonic sensor.
13. If an obstacle is detected within 25 cm:
    - Stop motors.
    - Save the current navigation state.
    - Enter obstacle-wait state.
14. Check the obstacle every 100 ms.
15. Resume only after three consecutive clear readings.
16. If blockage exceeds 30 seconds, enter HALT state.
17. When the target is reached, wait for 60 seconds.
18. Calculate a new A* path to the home position.
19. Execute the return path.
20. Display "Home Reached".
21. Return to the idle state.
22. End/await the next voice command.

VI. RESULT

The proposed robot was tested through 50 complete delivery cycles in a controlled indoor laboratory environment. The experimental results demonstrate that the proposed embedded architecture can perform autonomous delivery with acceptable performance.

The A* path-planning time was measured between approximately 180 and 240 ms for the 50 × 150 grid, remaining below the

ISSN:2455-9679
SJIF Impact Factor: 6.190

specified 500 ms target. The BNO085 showed less than 1.5° yaw drift after 60 seconds of static operation. Ultrasonic measurements at a 1 m wall showed approximately ±2 cm noise within three standard deviations.

The measured forward speed was approximately 12 cm/s, while the 90° turn duration was approximately 750 ms. The battery provided approximately 95 minutes of continuous operation, exceeding the 90-minute design requirement. Voice recognition achieved 93% accuracy across 50 attempts.

The delivery success rate was **47 successful deliveries out of 50 trials**, corresponding to **94%**. Thus, the system was very close to the desired 95% reliability target but did not fully achieve it. Two failures were associated with transient ultrasonic sensor readings near reflective tiled surfaces, while another resulted from an obstacle that remained in the robot's path for more than the 30-second timeout period.

Performance Summary

Performance Parameter	Experimental Result
A* solve time	180–240 ms
Yaw drift after 60 s	<1.5°
Ultrasonic noise	±2 cm (3σ)
Forward speed	≈12 cm/s
90° turn duration	750 ms
Battery life	≈95 min
Voice recognition accuracy	93%
Delivery success	47/50 (94%)
Main-loop execution time	≈6–9 ms
Total BOM cost	₹5,840

The non-blocking firmware architecture kept the main-loop execution time at approximately 6–9 ms, which is significantly below the 20 ms requirement. This allowed sensor polling, motor control, voice processing, LCD updates, and state-machine operations to run responsively.

The main advantages of the system are its low cost, offline operation, deterministic navigation, modular firmware, clearance-aware path planning, and safe obstacle handling. However, the current system has limitations including a static map, open-loop forward motion, single-direction obstacle sensing, limited voice vocabulary, absence of SLAM, and limited payload capacity.

VII. CONCLUSION

This research presented the design and implementation of a low-cost autonomous indoor delivery robot using an ESP32-S3 N16R8 microcontroller. The system integrates offline voice recognition, inertial heading estimation, ultrasonic obstacle detection, differential-drive motion, A* path planning, and a non-blocking finite state machine into a compact embedded platform.

The experimental results demonstrate that sophisticated autonomous delivery behaviour can be achieved without expensive LiDAR, external computers, cloud services, or complex robotics frameworks. The robot achieved an A* planning time of approximately 180–240 ms, 93% voice-recognition accuracy, approximately 95 minutes of battery operation, and a 94% delivery success rate across 50 controlled trials. The total reported bill of materials was ₹5,840.

The most important contribution of the work is the integration of low-cost hardware with carefully tuned algorithms. In particular, the combination of a clearance-aware A* planner and turn-cost penalty produced smoother routes, while the finite state machine provided predictable handling of delivery, obstacles, and fault conditions.

Although the current prototype is designed for a predefined indoor environment, it establishes a practical foundation for low-cost autonomous logistics. Future improvements such as wheel encoders, closed-loop PID control, additional range sensors,

ISSN:2455-9679
SJIF Impact Factor: 6.190

LiDAR-based SLAM, wireless telemetry, multi-target voice commands, payload monitoring, automatic charging, and voice feedback could significantly improve the system's reliability and scalability.

Overall, the proposed robot demonstrates that affordable embedded robotics can provide useful autonomous delivery services for small hospitals, laboratories, educational institutions, offices, and other controlled indoor environments.

VIII. REFERENCES

1. Adafruit Industries, *Adafruit_BNO08x Library Documentation*, Version 1.2.5, 2024.
2. Arduino, *Arduino IDE 2.x Reference Manual*, Arduino LLC, 2025.
3. Bosch Sensortec and CEVA Hillcrest Labs, *BNO080/BNO085 Datasheet*, Revision 1.16, 2023.
4. Choset, H., Lynch, K. M., Hutchinson, S., Kantor, G. A., Burgard, W., Kavraki, L. E., and Thrun, S., *Principles of Robot Motion: Theory, Algorithms, and Implementations*, MIT Press, 2005.
5. Cormen, T. H., Leiserson, C. E., Rivest, R. L., and Stein, C., *Introduction to Algorithms*, 4th ed., MIT Press, 2022.
6. Elechouse, *VC-02 Offline Voice Recognition Module: User Manual and Command Reference*, 2023.
7. Espressif Systems, *ESP32-S3 Technical Reference Manual*, Version 1.5, 2025.
8. Hart, P. E., Nilsson, N. J., and Raphael, B., “A Formal Basis for the Heuristic Determination of Minimum Cost Paths,” *IEEE Transactions on Systems Science and Cybernetics*, vol. 4, no. 2, pp. 100–107, 1968.
9. Infineon Technologies, *BTS7960 High Current PN Half Bridge Datasheet*, Revision 1.4, 2022.
10. LaValle, S. M., *Planning Algorithms*, Cambridge University Press, 2006.
11. Marakkalage, S. H. and Sarkar, C., “A Survey on Indoor Positioning and Localisation Systems,” *ACM Computing Surveys*, vol. 54, no. 7, 2021.
12. Mukhopadhyay, S. C., works on sensor-based robotic systems and embedded sensing.