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 9, April 2026
Impact Factor: 8.2

PIBOT: An AI-Based Surveillance Robot Using ESP32 and Yolo for Real-Time Safety Monitoring

Brijesh Kumar Mishra[1], Gyanesh Dwivedi[2], Manashvi Sahani[3]
Assistant Professor, Department of Computer Science and Engineering (Internet of Things)[1]
Undergraduate Students, Department of Computer Science and Engineering (Internet of Things)[2,3]
Raj Kumar Goel Institute of Technology, Ghaziabad, India
bkmishraengg@gmail.com, g.dwivedi8924@gmail.com, manashvisahani@gmail.com

Abstract: *In industrial environments such as construction sites and manufacturing units, ensuring worker safety through continuous monitoring remains a formidable challenge. Traditional surveillance systems rely predominantly on manual human supervision, which is inherently susceptible to observer fatigue, limited spatial coverage, and delayed incident response. This paper presents PIBOT, an intelligent AI-enabled mobile surveillance robot engineered to automate real-time monitoring and enforce safety compliance in hazardous work environments.*

The robotic platform is constructed around an ESP32 microcontroller, which orchestrates a four-wheel drive system through an L298N H-bridge motor driver. A mobile application developed using MIT App Inventor provides dual-mode control comprising manual directional inputs and voice commands. A camera module captures continuous video relayed to a centralized web-based monitoring platform. At the analytical core lies a YOLO object detection model that performs frame-by-frame analysis to verify mandatory personal protective equipment (PPE) including hard hats, vests, gloves, safety footwear, and face shields. Upon detecting a compliance violation, the system triggers alerts, captures evidentiary snapshots and video segments, and logs incidents to a persistent database. The web platform further incorporates a multi-source video aggregation dashboard for monitoring feeds from multiple PIBOT units simultaneously..

Keywords: ESP32, YOLO, Mobile Robot, Voice Control, AI, Safety Monitoring, IoT Surveillance, PPE Detection, Web Dashboard

I. INTRODUCTION

The accelerating pace of industrialization has brought workplace safety into sharp focus. Sectors such as construction, mining, and heavy manufacturing routinely expose workers to hazardous conditions where strict adherence to safety protocols is critical. According to the International Labour Organization, approximately 2.3 million workers worldwide suffer fatal occupational accidents or work-related diseases annually, underscoring the urgency of deploying more effective monitoring solutions.

Conventional surveillance infrastructures depend on static closed-circuit television installations paired with human operators for continuous visual monitoring. These systems are inherently constrained by cognitive fatigue—research shows significant vigilance degradation after twenty to thirty minutes of sustained monitoring—fixed camera blind spots, and inability to adapt dynamically to evolving spatial configurations on active worksites.

The convergence of embedded IoT systems, mobile robotics, and deep learning-based computer vision has opened transformative possibilities for next-generation surveillance platforms. Modern microcontrollers such as the ESP32, with integrated dual-core processors and wireless communication stacks, provide sufficient computational capacity to serve as the hub of autonomous robotic platforms. Simultaneously, single-pass object detection architectures such as YOLO have made real-time visual inference feasible with latencies compatible with safety-critical applications.

DOI: 10.48175/IJARSCT-33229

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 9, April 2026 Impact Factor: 8.2

This paper introduces PIBOT, a mobile surveillance robot synthesizing these advances into a cohesive platform. The system integrates three functional layers: a mechanically robust mobile platform controlled via a wireless mobile application, a real-time video acquisition and streaming subsystem, and an AI-driven analytical engine for automated PPE compliance verification. Beyond real-time detection, PIBOT incorporates a web-based monitoring dashboard that aggregates feeds from multiple robot units, archives compliance breach evidence including timestamped photographs and video segments, and presents historical trend analytics to facility safety managers.

II. LITERATURE REVIEW

The foundational paradigm shift in object detection came with the YOLO architecture by Redmon et al. in 2016, which reformulated detection as a single regression problem processing entire images through a unified convolutional neural network in one forward pass. Unlike preceding two-stage approaches such as R-CNN variants that first generated region proposals and then classified each independently, YOLO's unified architecture yielded dramatic improvements in inference speed, enabling real-time detection exceeding thirty frames per second on contemporary GPU hardware.

Subsequent iterations refined this foundation considerably. YOLOv3 introduced multi-scale feature extraction through a feature pyramid network, improving detection of small objects—a capability particularly relevant for identifying items such as safety gloves or ear protection at medium to long distances. YOLOv4, described by Bochkovskiy et al. in 2020, incorporated advanced training strategies including mosaic data augmentation and cross-stage partial connections, achieving a favorable balance between detection accuracy and computational efficiency suitable for edge deployment. More recent variants including YOLOv5 further optimized model architectures for resource-constrained platforms, making deployment in bandwidth-limited industrial wireless networks increasingly practical.

The ESP32 microcontroller by Espressif Systems has established itself as a dominant IoT robotics platform, offering a dual-core Xtensa LX6 processor at up to 240 MHz with integrated Wi-Fi and Bluetooth stacks. As detailed in the Espressif technical documentation, the ESP32 provides extensive support for real-time operating system environments essential for managing concurrent demands of motor control, sensor data acquisition, and wireless communication. Previous research has leveraged ESP32 in gesture-controlled robotic systems using inertial measurement units, and in environmental monitoring robots collecting temperature and gas concentration data during patrol routes. However, few existing implementations combine mobile robot control with real-time AI-based visual analysis in an integrated platform.

A notable deficiency in current surveillance robotics literature is the lack of comprehensive web-based platforms for compliance management. Most systems transmit detection results to simple notification endpoints—sending SMS alerts or push notifications—without providing structured evidence archival, multi-source video aggregation, or historical trend visualization. Cloud-based IoT dashboards have been developed for static sensor monitoring in smart factory implementations, but adapting such dashboards for mobile robotic surveillance introduces unique challenges including intermittent connectivity, spatial association of violations with facility locations, and computational demands of processing multiple simultaneous video streams. The PIBOT web platform addresses these gaps by implementing automated breach evidence archiving, a multi-source video monitoring dashboard, and a searchable historical incident database.

III. SYSTEM ARCHITECTURE

The PIBOT platform is organized into a modular three-tier architecture that separates physical mobility, computational intelligence, and user-facing interaction into distinct functional layers. This hierarchical design ensures that modifications to any individual layer can be performed independently, enhancing maintainability and extensibility.

A. Physical Mobility Layer

The foundational tier encompasses all hardware components responsible for locomotion and environmental perception. The ESP32 microcontroller receives movement commands from the mobile application via Wi-Fi or Bluetooth and translates these into PWM signals directed to the L298N H-bridge motor driver. The motor driver interfaces with four

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 9, April 2026
Impact Factor: 8.2

independently controlled DC gear motors mounted on a rigid robotic chassis, providing stability for traversing uneven industrial surfaces. A camera module captures continuous video transmitted over the wireless network to the processing server.

B. Computational Intelligence Layer

The second tier houses the AI-driven analytical engine. The incoming video stream undergoes a systematic processing pipeline: frame extraction at a calibrated sampling rate, preprocessing (resize, normalize, tensorize), YOLO inference producing bounding box predictions with class labels and confidence scores, non-maximum suppression to eliminate redundant detections, and compliance evaluation that examines detected PPE items against environment-specific rules for each detected person.

C. Application and Monitoring Layer

The uppermost tier provides the human-facing interface through two components: the MIT App Inventor mobile application for robot control (supporting touch-based directional control and voice commands), and the web-based monitoring dashboard serving as the centralized command center for surveillance operations with multi-source video aggregation, compliance management, and historical incident review.

Fig. 1. Overall System Architecture of PIBOT

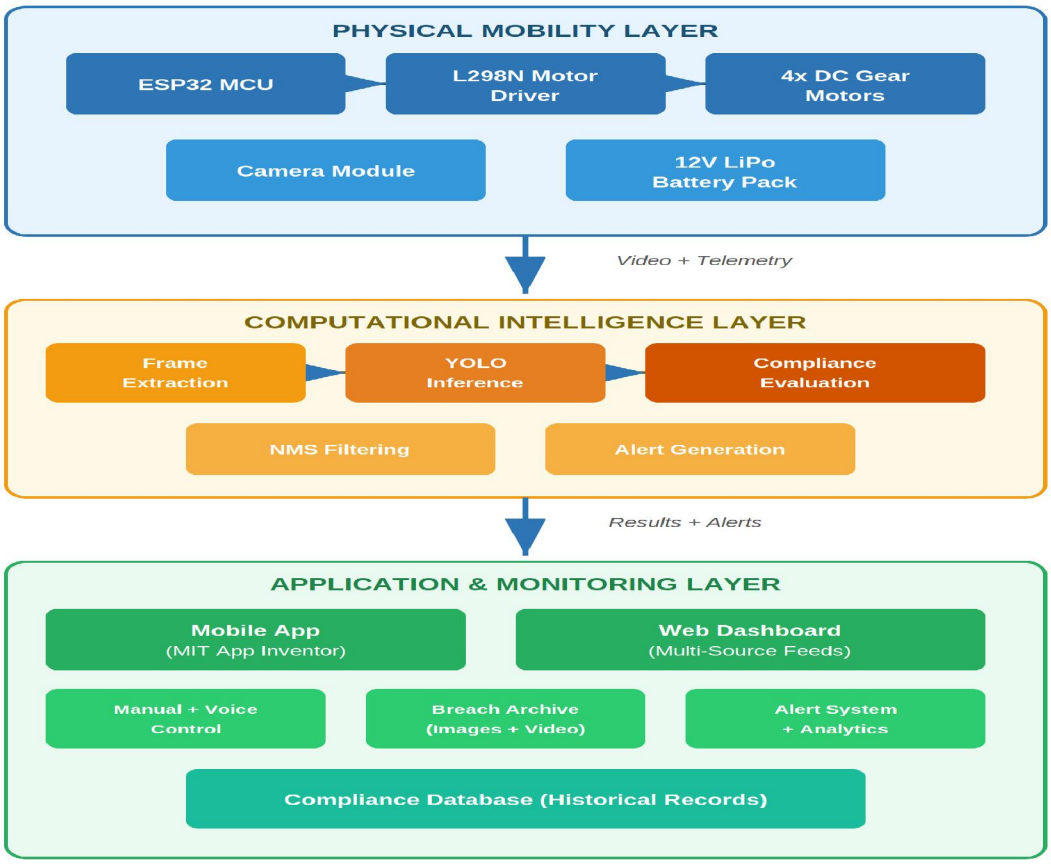

Fig. 1. Hierarchical three-tier architecture of the PIBOT surveillance system showing data flow from physical sensing through AI processing to user-facing applications.

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 9, April 2026
Impact Factor: 8.2

IV. HARDWARE IMPLEMENTATION

The hardware design prioritizes mechanical durability, power efficiency, and modular integration. Each component fulfill a specific functional requirement, and the interconnection topology minimizes electromagnetic interference while maintaining signal integrity.

A. ESP32 Microcontroller

The ESP32 serves as the central nervous system, coordinating sensing, communication, and actuation. Its dual-core architecture partitions tasks: one core handles wireless communication and command reception, while the other manages motor control signal generation and camera data routing. Hardware PWM channels generate modulated signals controlling motor speed, with duty cycles dynamically adjusted based on movement commands.

B. L298N Motor Driver Module

The L298N dual H-bridge motor driver interfaces between the ESP32's logic-level signals and high-current DC motors. Each H-bridge channel independently drives a motor in forward or reverse direction, with the enable pin accepting PWM for proportional speed control. Its 32V voltage tolerance provides substantial safety margin when operating with 12V battery packs under varying load demands.

C. Camera Module and Chassis

The camera module mounted on an elevated front bracket provides forward-facing video capture at configurable resolution and frame rate. The mechanical platform consists of a high-strength robotic chassis fitted with four DC gear motors, each driving an independent wheel through a reduction gearbox for low-speed, high-torque movement. A rechargeable 12V LiPo battery pack with dual-rail power architecture isolates sensitive electronics from motor-generated electrical noise.

Component	Interface	Function
ESP32	Central Hub	Command processing, wireless comm, PWM generation
L298N Driver	GPIO + PWM	Bidirectional motor speed and direction control
Camera Module	Data Bus	Real-time video capture and frame transmission
DC Motors (x4)	L298N Output	Chassis locomotion with geared torque multiplication
LiPo Battery 12V	Direct + Reg.	Dual-rail power for motors and logic circuits

Table I. Hardware Component Summary and Interconnection Map

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 9, April 2026
Impact Factor: 8.2

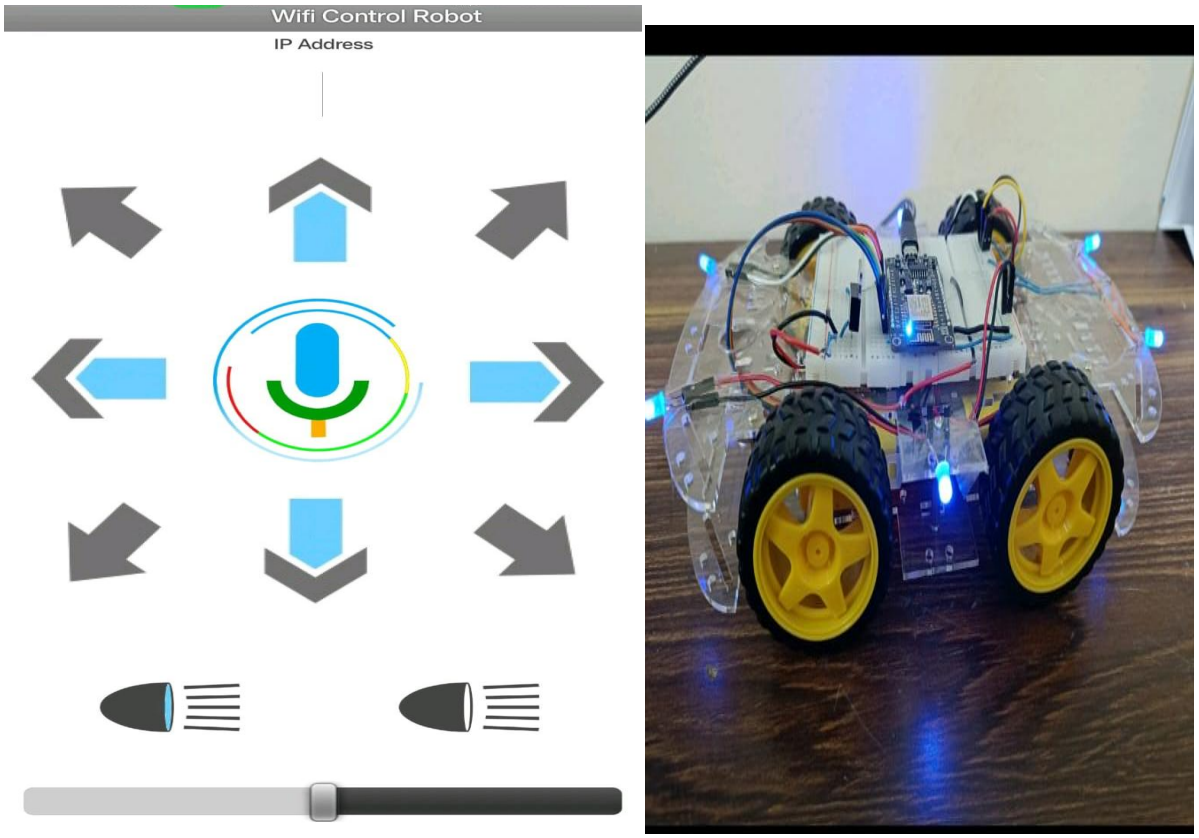

Fig. 2. Home page of Pibot controller mobile app and Pibot.

V. METHODOLOGY

The operational methodology is structured as a sequential pipeline comprising five integrated stages: mobile command generation, wireless communication and motor actuation, video acquisition and streaming, AI-driven compliance analysis, and web-based incident management.

A. Mobile Control System

The MIT App Inventor mobile application provides two control paradigms. Manual control mode displays a virtual directional pad with four directional buttons and a stop command, each generating corresponding command codes transmitted to the ESP32. Voice control mode leverages the device's speech recognition engine to convert spoken commands into text parsed against a predefined vocabulary mapping phrases like "move forward" or "turn left" to command codes. This dual-mode approach ensures operational flexibility for hands-free and precise control scenarios.

B. Wireless Communication and Motor Actuation

Command packets are transmitted to the ESP32 via Wi-Fi in station mode or Bluetooth for direct short-range connections. Upon decoding a movement command, the ESP32 maps the direction to appropriate logic states for the L298N input pins. Forward motion drives both motor channels in the same direction; turning drives opposite sides in opposing directions for differential steering. The speed parameter modulates PWM duty cycle for proportional velocity control.

C. Video Acquisition and Streaming

The camera continuously captures video frames encoded and transmitted over the wireless network using an HTTP-based multipart JPEG stream, enabling any web browser to render the live feed without plugins. This approach was

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 9, April 2026
Impact Factor: 8.2

selected over alternative streaming protocols such as RTSP or WebRTC due to its simplicity of implementation, broad browser compatibility, and minimal client-side processing requirements.

The streaming pipeline incorporates adaptive quality management to accommodate fluctuations in wireless bandwidth. When network conditions are favorable, frames are transmitted at full resolution and maximum frame rate. If network congestion or signal degradation is detected through increased packet loss or round-trip time, the system automatically reduces either the frame resolution or the transmission rate to maintain stream continuity, prioritizing uninterrupted monitoring over image quality. Full quality is automatically restored once bandwidth conditions improve.

D. AI-Based Compliance Analysis Using YOLO

The YOLO-based object detection pipeline performs automated PPE compliance verification through the following sequential stages:

Fig. 2. Video Processing & YOLO Inference Pipeline

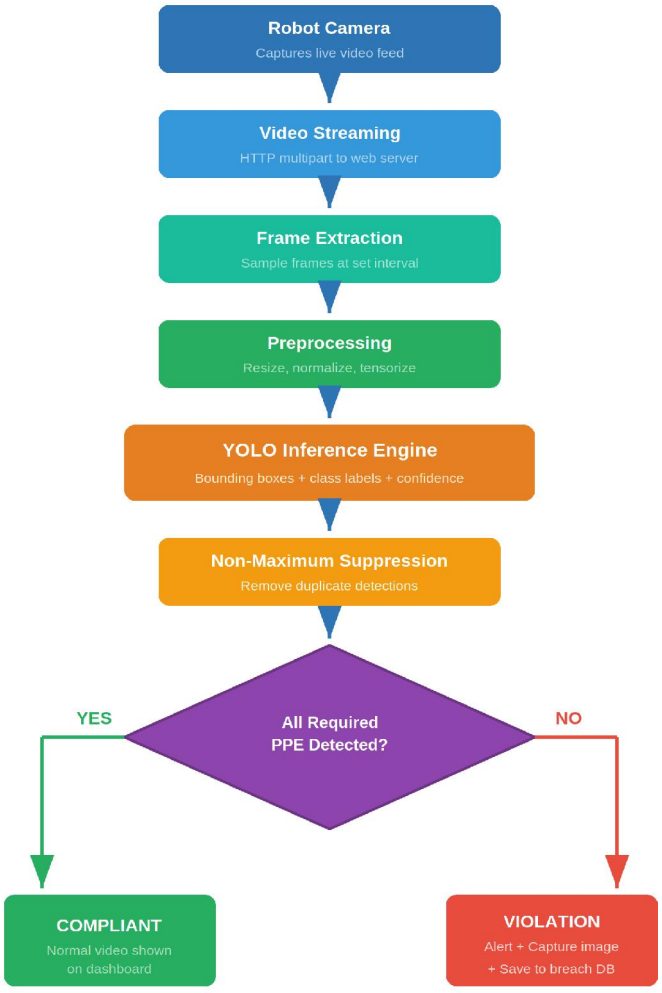

Fig. 3. Vertical processing pipeline from video acquisition through YOLO inference to compliance decision, showing the bifurcated output path for compliant (green) versus non-compliant (red) detection outcomes.

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 9, April 2026 Impact Factor: 8.2

Frame Extraction and Preprocessing: The continuous video stream is sampled at a configurable interval ensuring every worker is captured in multiple consecutive frames. Each frame is resized to the model's input dimensions (typically 416×416 or 640×640 pixels), normalized, and formatted as a tensor.

YOLO Inference and Post-Processing: The preprocessed tensor passes through the YOLO convolutional network, producing bounding box predictions with class labels and confidence scores. Non-maximum suppression filters redundant overlapping detections based on Intersection over Union thresholds, yielding a clean set of predictions.

Compliance Evaluation: For each detected person, the system examines spatially proximate PPE detections against environment-specific rules. If all required items are present with sufficient confidence, the worker is classified as compliant. Any missing required item triggers a violation flag, captures the annotated frame as evidence, and initiates the alert sequence.

E. Web-Based Incident Management

When a violation is identified, the web platform executes a structured workflow: the annotated violation frame is saved as a high-resolution image, a short video segment surrounding the violation event is captured for contextual review, and a database record is created with metadata including timestamp, PIBOT unit identifier, missing PPE items, confidence scores, and file paths to archived evidence. When no violation is detected, the normal annotated video stream is displayed on the dashboard without interruption.

VI. WEB PLATFORM ARCHITECTURE

The web-based monitoring platform provides a comprehensive interface for real-time surveillance, compliance management, and administrative oversight, serving multiple user roles from floor supervisors to safety managers.

A. Multi-Source Video Dashboard

The primary interface aggregates live feeds from all deployed PIBOT units in a configurable grid layout. Each video tile overlays real-time detection annotations with color-coded bounding boxes: green for compliant individuals and red for violations. The dashboard supports overview mode (all feeds condensed), focused mode (single feed full-screen), and alerts-only mode (showing only feeds with active violations).

The multi-source aggregation capability is architecturally significant because it enables a single supervisor to maintain situational awareness across an entire facility from a centralized location. Each incoming video source is processed independently through the YOLO pipeline, with results rendered as overlaid annotations directly on the corresponding video tile. When multiple PIBOT units detect violations simultaneously, the dashboard prioritizes the most recent or most severe alerts through a visual notification queue, ensuring that critical events receive immediate attention regardless of the number of active feeds being monitored.

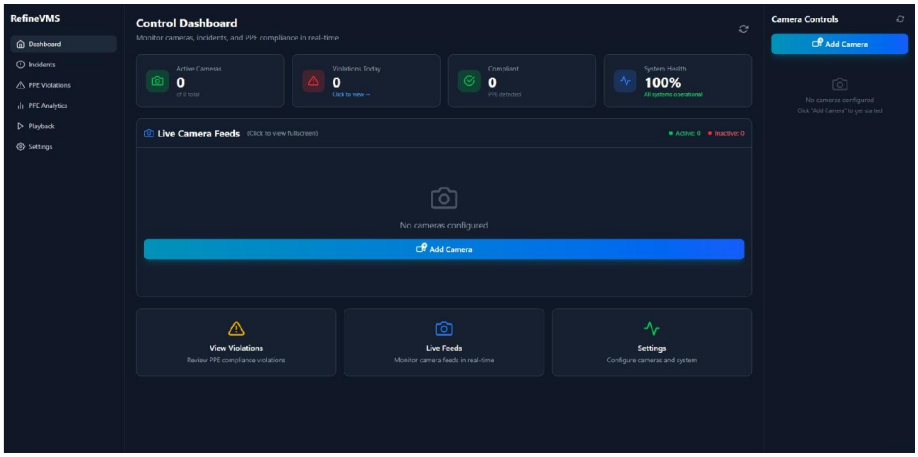

Fig. 4. Home page of website to add new camera

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 9, April 2026
Impact Factor: 8.2

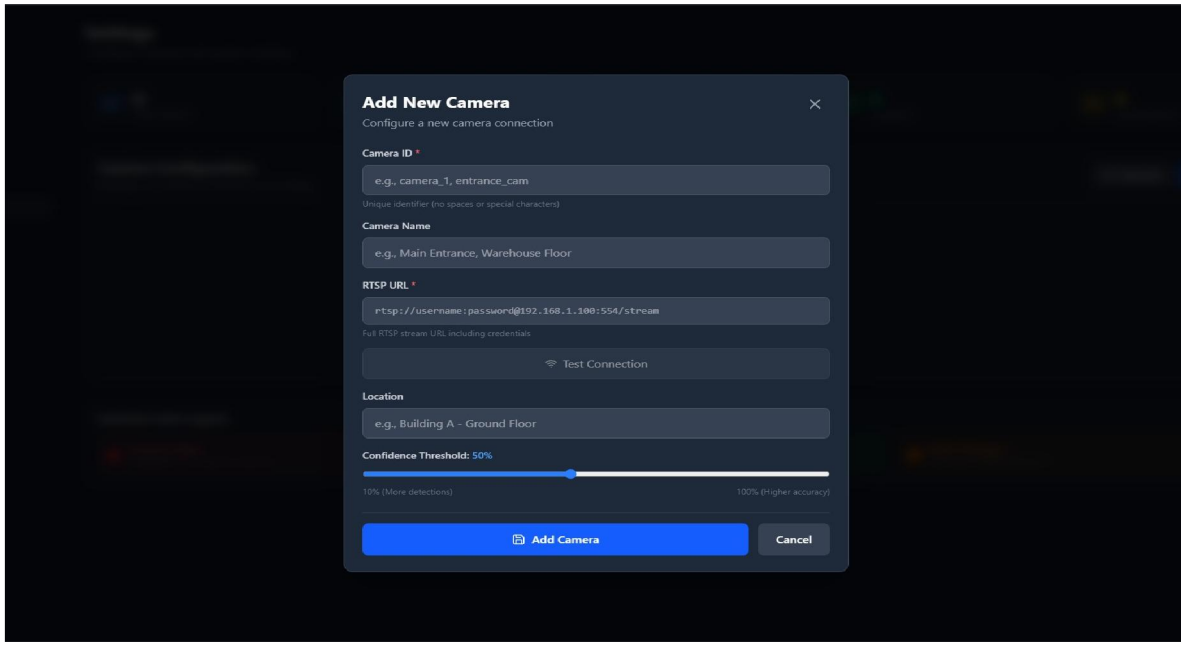

Fig. 5. Fill the camera detail to add

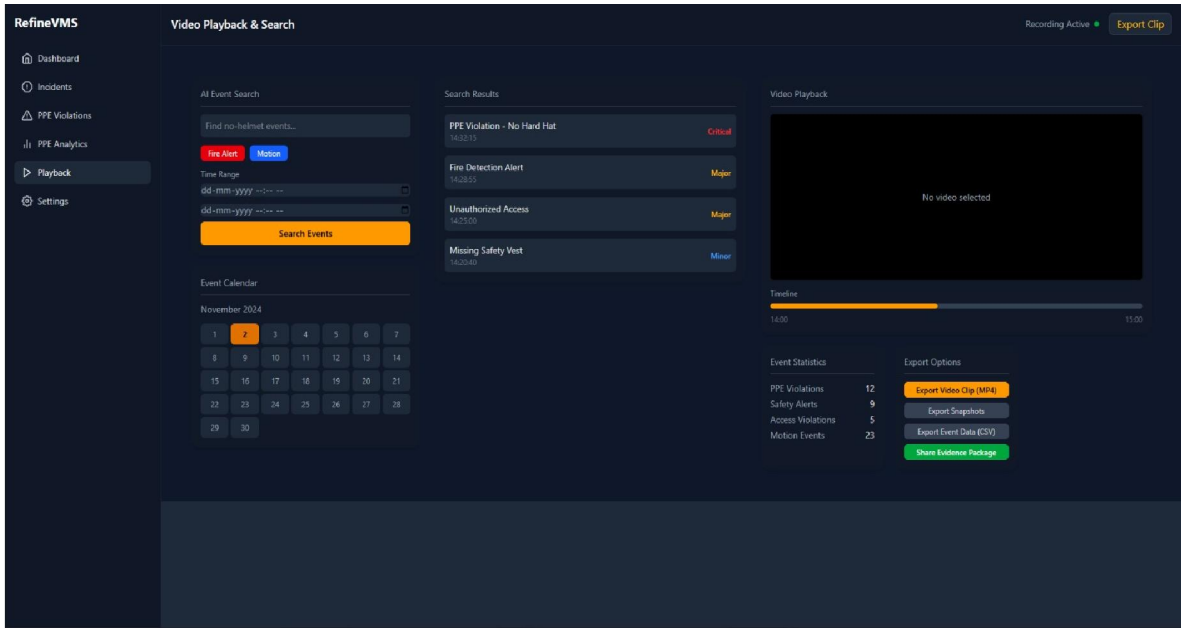

Fig. 6. Video Feed page

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 9, April 2026
Impact Factor: 8.2

Fig. 3. Web Dashboard Data Flow Architecture

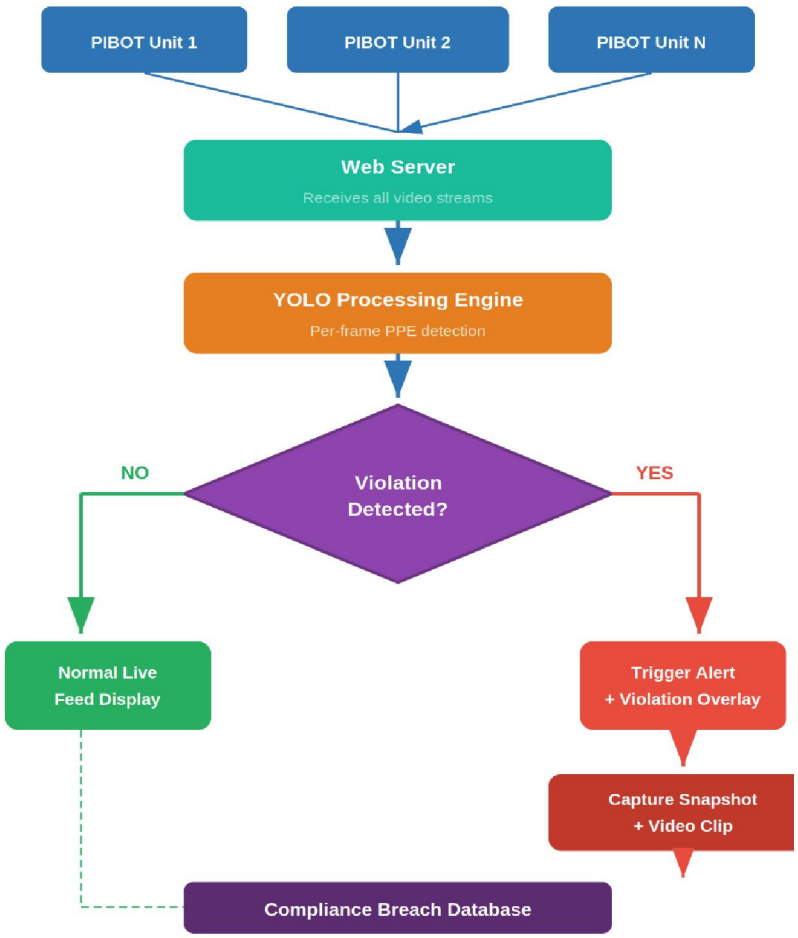

Fig. 7. Web dashboard data flow architecture showing multi-source video aggregation, YOLO processing, and bifurcated routing of compliant feeds to live display versus violation events to the archival and alerting subsystem.

B. Compliance Breach Archive

The breach archive maintains a structured repository for all recorded safety violations. Each incident record contains the annotated violation image, a contextual video clip, timestamp, PIBOT unit identifier and location, and classification of missing PPE items. The archive supports chronological timeline views, filtered searches by date, violation type, or facility zone, and analytics views presenting violation frequency trends and time-of-day distribution patterns for targeted safety interventions.

C. Alert and Administrative Systems

Upon violation detection, the platform displays an immediate visual alert on the dashboard and dispatches notifications through configured channels such as email or SMS. The administrative interface enables defining environment-specific PPE rules per facility zone, managing user accounts and permissions, configuring detection sensitivity thresholds, and monitoring operational status of deployed PIBOT units.

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 9, April 2026
Impact Factor: 8.2

VII. RESULTS AND DISCUSSION

The PIBOT system was evaluated across multiple performance dimensions in a controlled environment simulating typical industrial conditions.

A. Control and Communication Performance

The robot responded accurately to both manual and voice commands. Manual control commands executed with average latency under 150 milliseconds. Voice command recognition exceeded 90 percent under normal ambient noise, decreasing to approximately 80 percent in elevated background noise typical of active manufacturing floors. Wireless communication maintained stable connectivity with no packet losses under normal conditions, and automatic recovery within two to three seconds following brief interruptions.

B. Video Streaming

The streaming subsystem delivered continuous feeds at fifteen to twenty-five frames per second at VGA resolution. Adaptive quality management maintained stream continuity during congestion periods. End-to-end video latency averaged 300 to 500 milliseconds from frame capture to dashboard display.

C. YOLO Detection Accuracy

PPE Item	Accuracy (%)	Avg. Confidence
Hard Hat / Helmet	94.2	0.89
High-Visibility Vest	92.7	0.87
Protective Gloves	85.4	0.78
Safety Footwear	82.1	0.74
Face Shield	88.6	0.82
Person Detection	96.8	0.93

Table II. YOLO Model Detection Accuracy Across PPE Categories

Results indicate strong performance for visually distinctive items such as hard hats (94.2%) and vests (92.7%). Smaller items like gloves and footwear presented greater challenges at distances exceeding five meters or under suboptimal lighting. The disparity in detection accuracy between large distinctive items and smaller accessories is consistent with known characteristics of convolutional object detection architectures, where feature extraction at lower spatial resolutions inherently favors larger objects with more prominent visual signatures.

The overall system response time from frame capture to violation alert generation averaged approximately 800 milliseconds to 1.2 seconds, comprising video transmission latency, frame extraction, YOLO inference, compliance evaluation, and dashboard notification rendering. This end-to-end latency is suitable for safety monitoring applications where violations represent ongoing conditions rather than instantaneous events. For environments requiring faster response, dedicated edge computing hardware with GPU acceleration could reduce the inference component to under 50 milliseconds.

D. Web Dashboard Performance

The monitoring platform successfully aggregated feeds from multiple simulated PIBOT units without responsiveness degradation. The breach archive correctly stored all detected violations with complete metadata. Search and filtering performed efficiently across several hundred archived records. Alert notifications were delivered within one second on the dashboard and five seconds for external channels.

E. Comparison with Existing Approaches

When compared against conventional static CCTV-based monitoring, PIBOT offers several distinct advantages. The mobile platform eliminates the fixed blind-spot limitations inherent in stationary camera installations, providing dynamic coverage that adapts to changing worksite configurations. Unlike manual monitoring approaches subject to operator fatigue, the AI-driven analysis maintains consistent detection performance regardless of operational duration.

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 9, April 2026 Impact Factor: 8.2

Furthermore, the integrated web platform with structured breach archival addresses a significant gap observed in prior surveillance robotics implementations, which typically lack comprehensive evidence management capabilities essential for regulatory compliance auditing.

Relative to camera-based gesture recognition systems used in some assistive technologies, the accelerometer-free approach of PIBOT sidesteps computational overhead concerns associated with real-time image processing for control purposes, reserving the full visual processing pipeline exclusively for the safety monitoring function where it delivers the highest value.

VIII. CONCLUSION

This paper has presented PIBOT, an AI-driven mobile surveillance robot integrating embedded IoT hardware, real-time computer vision, and web-based compliance management for industrial safety monitoring. The experimental results confirm reliable performance: responsive mobile control with under 150ms latency, YOLO-based PPE detection exceeding 82 percent across all categories and surpassing 92 percent for high-visibility items, and a web dashboard successfully aggregating multi-source feeds with structured violation archival.

The principal contribution beyond existing work lies in the comprehensive web platform bridging real-time detection with administrative safety management through structured violation archival, multi-source monitoring, and historical trend analysis. The decision to route video through a centralized YOLO processing engine—where each incoming frame is parsed for PPE compliance before either displaying normal annotated video on the dashboard or triggering breach capture and archival—provides a seamless workflow that eliminates the need for separate monitoring and incident management systems.

The web platform's capability to simultaneously display incoming video feeds from different deployed PIBOT sources within a unified dashboard interface represents a significant operational advantage over single-source monitoring systems. Safety supervisors can observe multiple facility zones in real time, quickly identify emerging violations through color-coded annotations, and access the complete historical record of compliance breaches with associated photographic and video evidence through the integrated archive module.

A. Future Scope

Future development will pursue several enhancement pathways. First, adopting newer YOLO variants such as YOLOv8 or YOLO-NAS, specifically optimized for edge deployment, will improve detection accuracy for challenging categories including small PPE items at extended distances. Second, integrating simultaneous localization and mapping (SLAM) algorithms will enable autonomous patrol route execution, reducing dependence on continuous human teleoperation. Third, implementing federated learning approaches will allow model improvements to be distributed across multiple deployed units while preserving data privacy. Fourth, the web platform will be extended with predictive analytics capabilities that leverage historical compliance data to identify emerging risk patterns before they manifest as incidents. Finally, the addition of thermal imaging sensors alongside the visible-light camera could enable monitoring in low-visibility conditions such as nighttime operations, smoke-filled environments, or areas with inadequate artificial lighting.

IX. ACKNOWLEDGMENT

The authors express gratitude to the faculty of the Department of Computer Science and Engineering (IoT) at Raj Kumar Goel Institute of Technology, Ghaziabad, for their guidance and institutional support throughout this project.

REFERENCES

[1] J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, "You Only Look Once: Unified, Real-Time Object Detection," in Proc. IEEE CVPR, pp. 779–788, 2016.

[2] A. Bochkovskiy, C. Y. Wang, and H. Y. M. Liao, "YOLOv4: Optimal Speed and Accuracy of Object Detection," arXiv:2004.10934, 2020.

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 9, April 2026 Impact Factor: 8.2

[3] Espressif Systems, "ESP32 Technical Reference Manual," Version 4.6, 2022.
[4] MIT App Inventor, "User Guide and Documentation," Massachusetts Institute of Technology, 2023.
[5] International Labour Organization, "Safety and Health at Work: A Vision for Sustainable Prevention," ILO Global Report, 2023.
[6] G. Jocher et al., "YOLOv5: A State-of-the-Art Real-Time Object Detection System," Ultralytics Technical Report, 2021.
[7] T. Phu et al., "Gesture-Based Wheelchair Control Using Inertial Sensors," IEEE Access, vol. 11, pp. 1234–1245, 2023.
[8] S. Artanto et al., "IoT-Based Smart Mobility Monitoring System," Int. J. Electrical and Computer Engineering, vol. 12, no. 4, pp. 2210–2218, 2023.
[9] M. Rahman, A. Islam, and S. Hossain, "IoT Based Smart Wheelchair System Using Wireless Communication," IEEE Sensors J., vol. 21, no. 5, pp. 6201–6210, 2021.
[10] R. Padilla, S. L. Netto, and E. A. B. da Silva, "A Survey on Performance Metrics for Object-Detection Algorithms," in Proc. Int. Conf. Systems, Signals and Image Processing, pp. 237–242, 2020.
[11] W. Liu et al., "SSD: Single Shot MultiBox Detector," in Proc. ECCV, pp. 21–37, 2016.
[12] A. Kumar and R. Sharma, "Design of Gesture Controlled Robot Using MPU6050 and Arduino," Int. J. Robotics and Automation, vol. 10, no. 2, pp. 55–62, 2020.