Development of an Intelligent Edge AI Surveillance System for Industrial Safety

Bhukya Jagadeeshwar[1], Dr.B.Srikanth[2]
[1]M.Tech Student, [2]Assistant Professor
1,2, Department of ECE, KUCET
Kakatiya University, Warangal, Telangana, India, 506009
[1]bhukyajagadeesh9324@gmail.com, [2]bsrikanthiitkgp@kakatiya.ac.in

Abstract
Industrial environments require continuous surveillance and rapid response mechanisms to ensure safety, security, and operational efficiency. Conventional surveillance systems rely heavily on manual monitoring and lack intelligent decision-making capabilities. This project presents an IoT-X powered intelligent industrial surveillance and autonomy system using an AI-enabled ESP32-CAM module for real-time visual monitoring and automated event detection. The proposed system integrates edge artificial intelligence, IoT connectivity, and cloud analytics to detect unauthorized access, human presence, and abnormal activities within industrial premises. Lightweight deep learning models are deployed at the edge to enable low-latency inference, while the IoT-X framework ensures secure data communication and device management. Detected events trigger autonomous responses and real-time alerts to supervisors through mobile and web interfaces. The proposed solution offers a scalable, cost-effective, and intelligent surveillance platform suitable for Industry 4.0 and smart manufacturing applications.

Keywords: Artificial Intelligence (AI), Internet of Things (IoT), ESP32-CAM, Edge AI, Industrial Surveillance, Object Detection.

I. Introduction

Industrial environments such as manufacturing plants, warehouses, refineries, and power stations require continuous surveillance to ensure operational safety, security, and efficiency. Conventional CCTV-based surveillance systems primarily rely on human operators for monitoring, making them prone to delayed responses, human error, high operational costs, and limited scalability, particularly in large industrial facilities. Recent advancements in Artificial Intelligence (AI), the Internet of Things (IoT), and Edge Computing have enabled the development of intelligent surveillance systems capable of real-time monitoring, automated event detection, and autonomous decision-making. This work presents an AI-enabled IoT-based Intelligent Industrial Surveillance System using ESP32-CAM, which integrates edge AI processing, IoT communication, and cloud-based monitoring to provide a comprehensive surveillance solution. The ESP32-CAM captures real-time images and performs on-device object detection to identify human presence, unauthorized access, and abnormal activities using lightweight deep learning models. Detected events are securely transmitted through IoT protocols to a cloud platform for storage, visualization, and remote monitoring, while automated alerts and control actions enhance system responsiveness. The proposed system improves industrial safety, reduces dependence on manual surveillance, supports real-time intelligent decision-making, and aligns with Industry 4.0 requirements. Owing to its low-cost, scalable, and energy-efficient architecture, the system is well suited for applications in

Doi:10.48047/ijiee.2026.16.2.60

manufacturing industries, warehouses, power plants, smart factories, and other industrial security environments, with future potential for integration with advanced AI models, robotics, and predictive maintenance systems.

II. Related Works

The advancement of Internet of Things (IoT) and Artificial Intelligence (AI) has significantly improved surveillance systems by enabling real-time monitoring and intelligent decision-making. Traditional surveillance systems lack automation and rely on human intervention, leading to inefficiencies. Recent research focuses on integrating IoT, AI, and embedded systems to develop smart surveillance systems capable of detecting events and responding automatically (Islam et al., 2015; Gubbi et al., 2013).

IoT-based surveillance systems enable devices to communicate over the internet and provide real-time monitoring capabilities. A study by Islam et al. (2015) explains that IoT-based healthcare and monitoring systems allow continuous data collection and remote accessibility. Similarly, Gubbi et al. (2013) highlighted that IoT architecture supports real-time communication between devices and cloud platforms. Recent work on ESP32-CAM-based surveillance systems shows that IoT enables real-time alerts and remote monitoring, improving system efficiency (Sharma et al., 2025; Kumar et al., 2025).

The ESP32-CAM module has gained popularity due to its low cost and integrated Wi-Fi capability. According to Kumar et al. (2025), ESP32-CAM provides an efficient platform for real-time video streaming and IoT integration. Another study by Sharma et al. (2025) demonstrated that ESP32-CAM can be used for surveillance applications with remote monitoring and alert systems. These studies confirm that ESP32-CAM is suitable for industrial surveillance due to its compact size and wireless capabilities.

Artificial Intelligence enhances surveillance systems by enabling automated detection of objects and activities. A study by Redmon et al. (2016) introduced the YOLO (You Only Look Once) algorithm for real-time object detection. Similarly, Howard et al. (2017) proposed MobileNet, a lightweight deep learning model suitable for embedded systems. Research shows that AI-based systems can detect human presence, identify objects, and recognize abnormal activities with high accuracy (Raza et al., 2019; Singh et al., 2021). Edge computing allows data processing near the source, reducing latency and bandwidth usage. According to Shi et al. (2016), edge computing improves system responsiveness and reduces dependency on cloud infrastructure.

Recent studies demonstrate that combining edge computing with AI enables real-time detection and faster response, especially in industrial environments (Zhang et al., 2022; MDPI Sensors, 2025).

The integration of AI and IoT, known as AIoT, enables intelligent and autonomous systems. A survey by Zhou et al. (2019) highlights that AIoT systems combine data collection (IoT) and intelligent analysis (AI) to improve automation and decision-making. Similarly, Gubbi et al. (2013) emphasized that IoT combined with intelligent processing enables smart environments and autonomous systems.

Autonomous systems are capable of detecting events and triggering actions without human intervention. A study by Raza et al. (2019) demonstrated that AI-based surveillance systems can automatically generate alerts and notifications. Recent research using ESP32-CAM shows that autonomous surveillance systems can perform real-time monitoring

Doi:10.48047/ijiee.2026.16.2.60

and remote alerting, improving industrial safety (Kumar et al., 2025).

The literature review indicates that IoT, AI, and edge computing have significantly improved surveillance systems. However, most systems focus on individual technologies rather than integrated solutions. The proposed system addresses these gaps by combining ESP32-CAM, edge AI, IoT-X communication, and cloud analytics into a unified architecture, enabling real-time monitoring and autonomous decision-making.

III. Proposed System

The system design of the IoT-X Powered Intelligent Industrial Surveillance System focuses on integrating AI-based image processing, IoT communication, and autonomous decision-making into a unified architecture. The system is designed to perform real-time image capture, AI-based object detection, secure data transmission and automated alert generation

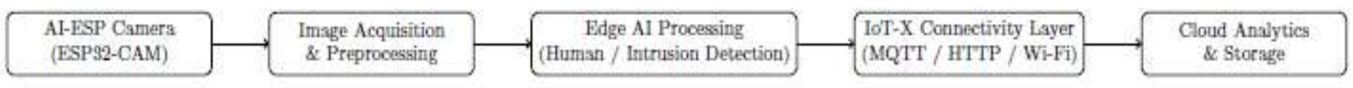

Figure 1 Image Acquisition and IoT-X Camera Nodes

Block Diagram of System

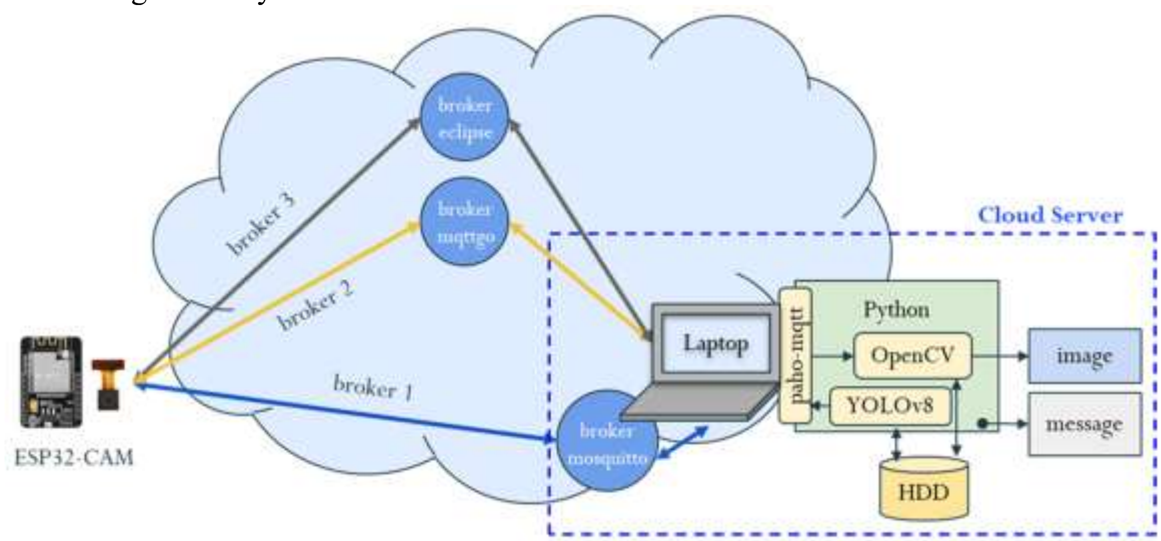

Figure 2 Block Diagram of the system with architecture

The working model of the proposed system integrates edge intelligence, IoT connectivity, and cloud analytics to enable real-time industrial surveillance with autonomous decision making. The system comprises multiple interconnected modules that collaborate to capture visual data, perform on-device and edge-level inference, communicate results securely, and trigger automated responses.

Doi:10.48047/ijiee.2026.16.2.60

AI Detection Flow

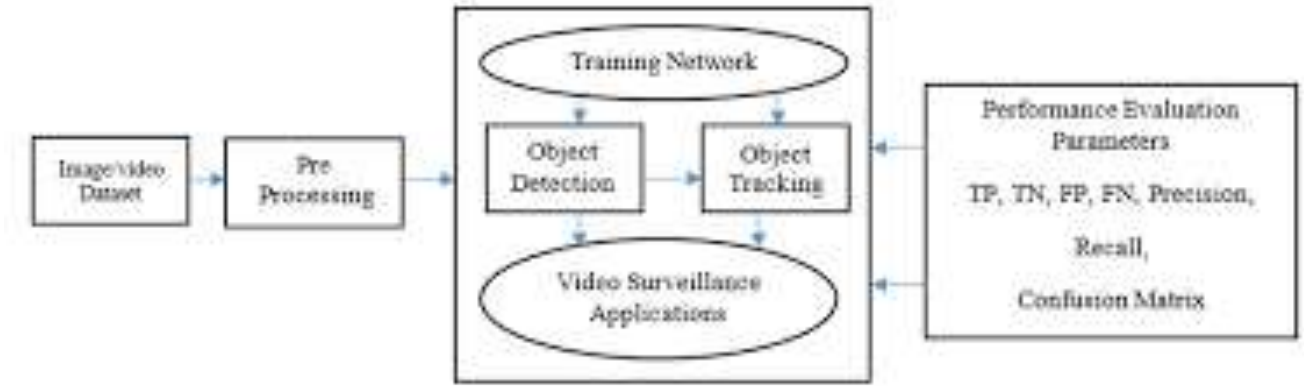

Figure 3 Object Detection work flow

1. Capture Image

The first stage of the AI pipeline is image acquisition using the ESP32-CAM equipped with the OV2640 camera sensor. The camera continuously captures images or video frames from the industrial environment at predefined intervals. These images serve as the raw input for the surveillance system. Continuous image acquisition ensures that the system monitors the environment in real time, allowing immediate detection of unauthorized access, human presence, or abnormal activities. The captured frames are converted into digital format and forwarded to the image processing module.

2. Preprocess Image

Before the captured image is analyzed by the AI model, it undergoes preprocessing to improve its quality and consistency. Image preprocessing includes operations such as noise removal, brightness and contrast adjustment, resolution normalization, and resizing to match the input dimensions required by the AI model. These operations remove unwanted distortions caused by poor lighting, camera noise, or environmental conditions. A clean and normalized image improves the accuracy of object detection while reducing computational complexity for the ESP32-CAM's limited hardware resources.

3. Feed to AI Model

After preprocessing, the image is supplied to a lightweight deep learning model running on the edge device. The project employs models such as MobileNet, Tiny-YOLO, or TensorFlow Lite that are optimized for embedded systems like the ESP32-CAM. The model performs inference directly on the captured image without requiring continuous cloud connectivity, thereby reducing network bandwidth usage and minimizing response time. Running AI at the edge enables the surveillance system to provide immediate detection even when internet connectivity is limited.

4. Extract Features

During inference, the AI model automatically extracts important visual features from the input image. Feature extraction identifies meaningful characteristics such as edges, shapes, textures, motion patterns, and color variations. Rather than analyzing every pixel individually, the neural network learns these distinctive features to recognize different objects accurately. These extracted features provide the basis for differentiating humans, machinery, vehicles, and other industrial objects while also identifying suspicious movements or abnormal behavior.

Doi:10.48047/ijiee.2026.16.2.60

5. Classify Object

Using the extracted features, the AI model determines the category of the detected object. It compares the learned features with patterns acquired during training and assigns an appropriate label such as person, machine, authorized object, or unknown object. In the proposed surveillance system, the classifier specifically detects human presence, unauthorized access, motion, and abnormal activities. Accurate object classification enables the system to distinguish between normal industrial operations and events requiring immediate attention.

6. Generate Output

Once the object has been classified, the AI model generates the inference output. This output generally includes the detected object label, confidence score, and event information. The system evaluates whether the detected object represents a normal situation or a potential threat. Detection results are stored locally or transmitted to the cloud through the IoT-X communication layer using MQTT or HTTP/HTTPS protocols for further analysis, visualization, and event logging. This output also becomes available on the monitoring dashboard for supervisors.

7. Trigger Action

The final stage converts the AI decision into an automated response. When the generated output indicates an abnormal event, the system immediately triggers predefined actions such as activating a buzzer, sending email notifications, updating the dashboard, logging the event, or initiating automatic control mechanisms. These actions are performed without human intervention, allowing the surveillance system to respond quickly to security threats and hazardous situations. The autonomous response capability significantly enhances industrial safety while reducing the need for continuous manual monitoring.

Edge AI Inference

Edge AI Inference is a key component of the proposed IoT-X Powered Intelligent Industrial Surveillance System. Instead of sending every captured image to a remote cloud server for processing, the AI model is deployed directly on the ESP32-CAM, allowing object detection and decision-making to occur at the edge of the network. This approach enables the system to analyze images immediately after they are captured, resulting in faster response times, lower communication delays, and reduced network bandwidth usage. By performing inference locally, the surveillance system continues to operate efficiently even in environments with limited or unstable internet connectivity.

On-Device AI Inference

The ESP32-CAM is equipped with a lightweight deep learning model optimized for embedded hardware. These models, developed using frameworks such as TensorFlow Lite, MobileNet, or Tiny-YOLO, are designed to perform efficient image analysis while consuming minimal memory and processing power. After capturing an image, the ESP32-CAM preprocesses it and feeds it directly into the AI model. The model analyzes the visual information and produces detection results within a short time, eliminating the need to transmit raw images to the cloud for every operation. This on-device inference significantly improves system efficiency and supports continuous real-time surveillance.

Real-Time Recognition Using Pretrained Deep Learning Models

The surveillance system employs pretrained deep learning models that have been trained to recognize various objects and activities commonly encountered in industrial environments. These models enable accurate identification of security-related events without requiring extensive computational resources.

Doi:10.48047/ijiee.2026.16.2.60

1. Human Presence and Motion Patterns

The AI model continuously analyzes captured images to detect the presence of people within the monitored area. In addition to identifying individuals, it examines movement patterns to distinguish between normal worker activities and unusual movements. Continuous monitoring of human presence helps improve workplace safety and ensures that restricted zones remain protected from unauthorized entry.

2. Unauthorized Access Attempts

One of the primary objectives of the surveillance system is to detect unauthorized personnel entering restricted industrial areas. The AI model compares the detected activity with predefined monitoring rules and identifies suspicious access attempts. Once an unauthorized entry is recognized, the system immediately records the event and initiates the appropriate alert mechanism, enabling rapid intervention by security personnel.

3. Object Classification

The deep learning model classifies detected objects into predefined categories such as person, machinery, equipment, or other industrial objects. Object classification allows the surveillance system to distinguish between normal operational equipment and human movement, reducing false alarms and improving the accuracy of event detection. Correct object identification is essential for making intelligent decisions regarding industrial safety and security.

4. Abnormal or Suspicious Activity Detection

Beyond simple object recognition, the AI model analyzes behavioral patterns to identify abnormal or suspicious activities. Examples include unexpected movement in restricted areas, unusual human behavior, prolonged presence in sensitive locations, or activities that deviate from normal industrial operations. Once such behavior is detected, the surveillance system automatically generates alerts and initiates predefined safety procedures to minimize potential risks.

Computer Vision (CV) is a branch of Artificial Intelligence (AI) that enables computers and embedded devices to acquire, process, analyze, and understand digital images and videos in a way that resembles human vision. It combines image processing, machine learning, and deep learning algorithms to identify objects, detect events, recognize patterns, and make intelligent decisions automatically. In industrial surveillance systems, computer vision eliminates the need for continuous manual monitoring by allowing cameras to automatically detect people, machinery, unauthorized access, and abnormal activities in real time.

Working Principle of Computer Vision

The computer vision pipeline consists of several stages:

1. Image Acquisition: The camera captures images or video frames from the surrounding environment.
2. Image Preprocessing: The captured image is enhanced by removing noise, adjusting brightness, normalizing resolution, and resizing.
3. Feature Extraction: The AI model identifies significant visual characteristics such as edges, shapes, textures, colors, and motion patterns.
4. Object Detection and Classification: The extracted features are analyzed to detect and classify objects such as humans, machinery, or unauthorized persons.
5. Decision Making: Based on the classification results, the system determines whether the detected event is normal or abnormal.
6. Action Generation: If an abnormal event is detected, the system automatically triggers alarms, notifications, dashboard updates, or other predefined actions.

Doi:10.48047/ijiee.2026.16.2.60

In the IoT-X Powered Intelligent Industrial Surveillance System, computer vision forms the core intelligence of the surveillance process. The ESP32-CAM captures live images from industrial areas, preprocesses them, and uses an embedded AI model to detect human presence, unauthorized access, and abnormal activities. Once an event is identified, the system sends alerts through MQTT/HTTP communication, updates the cloud dashboard, logs the event, and activates devices such as buzzers. This integration of computer vision, edge AI, and IoT provides a low-cost, scalable, and autonomous surveillance solution suitable for modern Industry 4.0 environments.

Object Detection Using Lightweight AI Models

The proposed industrial surveillance system employs lightweight deep learning models to achieve efficient real-time object detection on the resource-constrained ESP32-CAM platform. MobileNet, a lightweight Convolutional Neural Network (CNN), is utilized for feature extraction and object classification owing to its depthwise separable convolution architecture, which significantly reduces computational complexity, memory usage, and power consumption while maintaining high detection accuracy. Tiny-YOLO, a compact version of the YOLO (You Only Look Once) algorithm, performs fast object localization and classification in a single forward pass, enabling real-time detection of multiple objects such as human presence, industrial equipment, unauthorized intruders, and abnormal activities. To deploy these models efficiently on embedded hardware, TensorFlow Lite (TFLite) is used as the inference framework. TensorFlow Lite converts trained deep learning models into an optimized .tflite format through quantization, weight compression, and memory optimization, enabling low-latency inference with minimal RAM and processing requirements. During operation, the captured image is first preprocessed and then fed into the AI model, where the network scans the image, extracts meaningful visual features, identifies objects, and outputs the corresponding class labels (e.g., person or machinery) along with confidence scores. By performing inference directly on the ESP32-CAM, the system minimizes dependence on cloud computing, reduces communication bandwidth, lowers power consumption, and provides rapid autonomous decision-making. The combination of MobileNet, Tiny-YOLO, and TensorFlow Lite offers an effective balance between detection accuracy, computational efficiency, and real-time performance, making the proposed surveillance system highly suitable for edge AI-based Industry 4.0 applications.

Hardware Components Description

The proposed AI-enabled IoT-based Intelligent Industrial Surveillance System is built around a set of integrated hardware components that collectively perform image acquisition, edge AI processing, wireless communication, and automated event notification. The ESP32-CAM module serves as the central processing unit of the system, combining a dual-core ESP32 microcontroller with built-in Wi-Fi, Bluetooth, and an integrated camera interface. It captures images, executes lightweight AI models for edge inference, and communicates detection results to the cloud through IoT protocols. The integrated OV2640 CMOS camera sensor captures high-quality images with a resolution of up to 2 MP and supports JPEG compression, providing visual input for AI-based object detection while maintaining low power consumption. Edge AI processing is implemented directly on the ESP32-CAM using lightweight deep learning frameworks such as TensorFlow Lite, TinyML,

Doi:10.48047/ijiee.2026.16.2.60

MobileNet, and Tiny-YOLO, enabling real-time feature extraction, object classification, and detection of human presence, unauthorized access, and abnormal activities without relying on continuous cloud connectivity. The system also incorporates a buzzer/alarm to provide immediate audible alerts whenever suspicious events are detected and LED status indicators to display system operation, connectivity, and processing status. Additional sensors, including motion, temperature, and gas sensors, can be integrated to enhance environmental monitoring and improve industrial safety. A regulated 5 V power supply provides stable and reliable power to all hardware components, ensuring uninterrupted operation. Together, these components form a compact, energy-efficient, and scalable hardware platform capable of delivering intelligent real-time surveillance for Industry 4.0 and smart industrial applications.

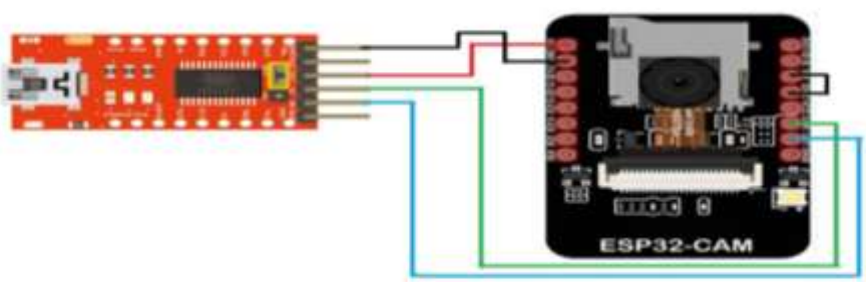

Figure 4: ESP32-CAM

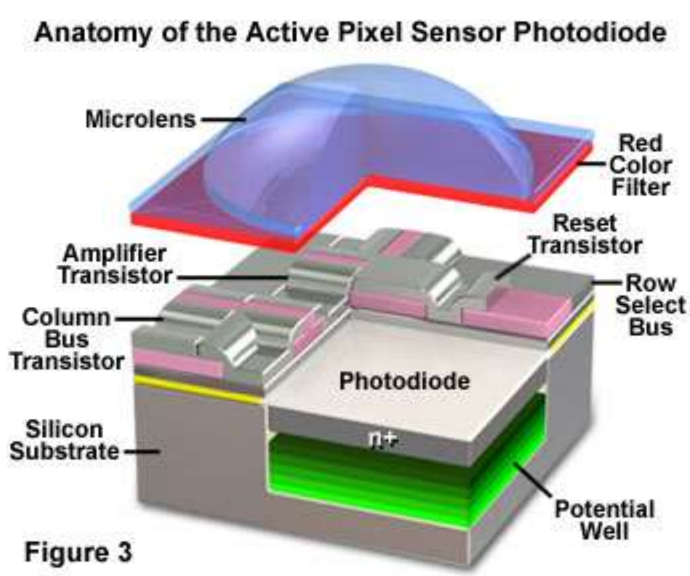

Figure 5: Camera Sensor (OV2640)

Doi:10.48047/ijiee.2026.16.2.60

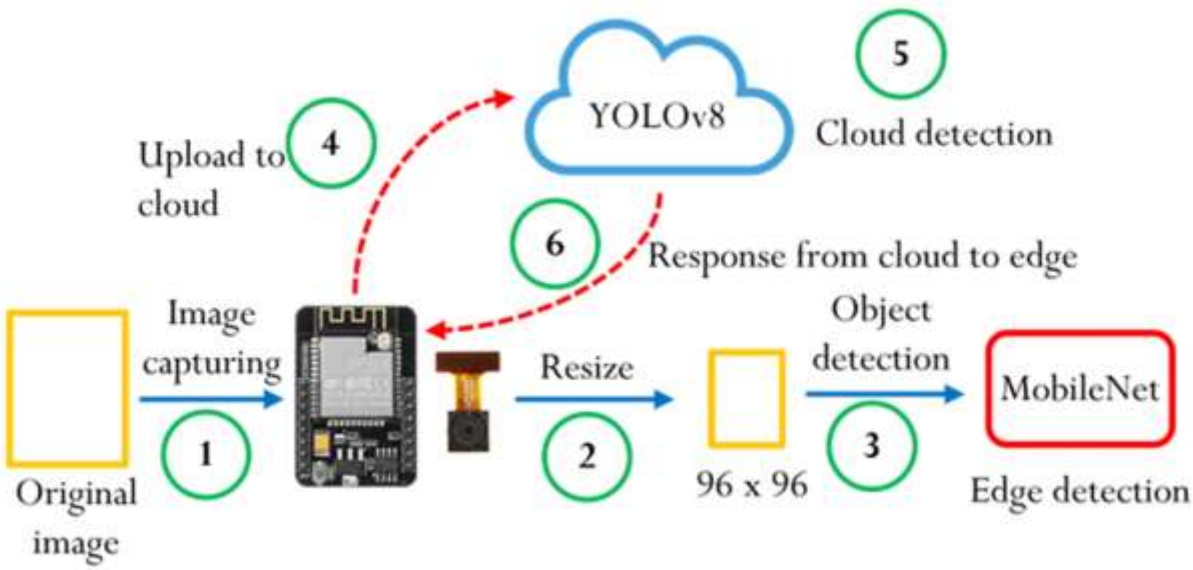

Figure 6: AI Processing Module (Edge AI on ESP32)

Cloud Analytics, System Design, and Performance

The proposed surveillance system integrates cloud analytics with a centralized web and mobile dashboard to enable remote monitoring, long-term data storage, event logging, and visual analytics. Detection results generated by the ESP32-CAM are transmitted to the cloud for archival, trend analysis, anomaly reporting, model updates, and system health monitoring, while stakeholders can access live camera feeds, alert history, device status, and analytical summaries through a unified dashboard. The working model offers several advantages, including real-time object detection with minimal latency, reduced bandwidth usage through edge AI processing, secure IoT communication, autonomous alert generation, and cost-effective deployment using low-power ESP32-CAM devices. The system is designed with key considerations such as low latency, high detection accuracy, secure communication, energy efficiency, and scalability. Although the architecture provides autonomous operation, efficient edge AI processing, and secure IoT connectivity, its performance is constrained by the limited processing capability of the ESP32, the need for reliable internet connectivity for cloud services, and the restricted size of AI models that can be deployed on embedded hardware. Overall, the integration of ESP32-CAM, Edge AI, IoT-X communication, and cloud analytics provides a robust, intelligent, and scalable surveillance solution for Industry 4.0 applications.

Conclusion

The IoT-X Powered Intelligent Industrial Surveillance and Autonomy with AI-ESP Camera project presents an intelligent, cost-effective, and scalable surveillance solution for modern industrial environments. By integrating the ESP32-CAM, Edge AI, IoT-X communication, and cloud computing, the system is capable of continuously monitoring industrial premises, detecting human presence, unauthorized access, object movement, and abnormal activities in real time. The deployment of lightweight AI models on the ESP32-CAM enables fast on-device inference, reducing latency, bandwidth usage, and dependence

Doi:10.48047/ijiee.2026.16.2.60

on cloud computing while ensuring prompt event detection and automated response. The integration of MQTT/HTTP communication, cloud-based monitoring, and autonomous alert generation enhances industrial safety, operational efficiency, and security. Overall, the proposed system demonstrates the potential of combining Artificial Intelligence, Internet of Things (IoT), and Edge Computing to develop smart surveillance systems that support Industry 4.0 and intelligent manufacturing environments.

Future Scope

The proposed surveillance system offers significant opportunities for future enhancement to further improve its intelligence, reliability, and automation capabilities. Advanced deep learning models with higher detection accuracy can be deployed as embedded hardware becomes more powerful. Additional sensors such as thermal cameras, gas sensors, smoke detectors, vibration sensors, and environmental monitoring devices can be integrated to provide comprehensive industrial safety monitoring. Future versions may incorporate facial recognition, predictive analytics, behavior analysis, anomaly prediction, and robotics for autonomous inspection and emergency response. Furthermore, integrating 5G communication, digital twin technology, blockchain-based security, and advanced cloud analytics can enhance scalability, cybersecurity, and real-time decision-making. These improvements will transform the proposed system into a fully autonomous smart surveillance platform capable of supporting next-generation Industry 4.0 and Industrial Internet of Things (IIoT) applications.

REFERENCES

1. Islam, S. R., et al., “The Internet of Things for Health Care,” IEEE Access, 2015.
2. Gubbi, J., et al., “Internet of Things: A Vision,” Future Generation Computer Systems, 2013.
3. Redmon, J., et al., “YOLO: Real-Time Object Detection,” 2016.
4. Howard, A., et al., “MobileNet,” Google Research, 2017.
5. Shi, W., et al., “Edge Computing: Vision and Challenges,” IEEE, 2016.
6. Zhou, J., et al., “AIoT: Artificial Intelligence of Things,” IEEE, 2019.
7. Raza, M., et al., “IoT-Based Smart Surveillance System,” IEEE, 2019.
8. Zhang, Q., et al., “Edge AI Systems,” IEEE, 2022.
9. Kumar, A., et al., “ESP32-CAM Surveillance Systems,” 2025.
10. Sharma, R., et al., “IoT Surveillance Using ESP32-CAM,” 2025.
11. Singh, V., et al., “AI-Based Monitoring Systems,” 2021.
12. Real-Time Video Surveillance Using ESP32-CAM, Journal of Embedded Systems, 2025
13. IoT-Based Home Monitoring System Using ESP32-CAM, 2025
14. Surveillance Robo Car Using ESP32-CAM Module, 2025
15. IoT and AI Based ESP32-CAM Smart System, Blockchain Frontier Technology, 2025
16. Edge Computing with ESP32 for AI Applications, MDPI Sensors Journal, 2025
17. Pervasive AI for IoT Applications

Doi:10.48047/ijiee.2026.16.2.60