RESEARCH PAPER

Edge-AI Based Smart Surveillance System Using ESP32-CAM with Face Recognition

Dr J. Aruna[1], Dr S. Prabakaran[2], Dr M. Sangeetha[3], Dr V. Vijayakumar[4], Ragavi C[5], Sathana T[6]

[1]Assistant Professor, Department of Computer Science and Engineering, V.S.B Engineering College, Karur, Tamil Nadu, India. Email: arunavsb2636@gmail.com (Corresponding Author)

[2]Assistant Professor, Department of Computer Science and Engineering, V.S.B Engineering College, Karur, Tamil Nadu, India. Email: mokipraba@gmail.com

[3]Professor, Department of Computer Science and Engineering, V.S.B Engineering College, Karur, Tamil Nadu, India. Email: rsv2008pll@gmail.com

[4]Associate Professor, Department of Computer Science and Engineering, V.S.B Engineering College, Karur, Tamil Nadu, India. Email: Vksece2007@gmail.com

[5]UG Scholar, Department of Computer Science and Engineering, V.S.B Engineering College, Karur, Tamil Nadu, India. Email: ragavichinasamy@gmail.com

[6]UG Scholar, Department of Computer Science and Engineering, V.S.B Engineering College, Karur, Tamil Nadu, India. Email: sathanasathana2004@gmail.com

***Corresponding author: Dr J. Aruna, Assistant Professor, Department of Computer Science and Engineering, V.S.B Engineering College, Karur, Tamil Nadu, India**

Email: arunavsb2636@gmail.com

Received: 30th May, 2026; **Revised:** 11th June, 2026; **Accepted:** 15th June, 2026; **Available Online:** 17th June, 2026

ABSTRACT

Background

With the growing demand for safer and smarter monitoring systems, this project introduces an Edge-AI based smart surveillance system using the ESP32-CAM with face detection and face recognition, combined with environmental and safety sensors.

Objective

The proposed system is designed to perform real-time face analysis directly on the device itself, so that personal data does not need to be continuously sent to the cloud. This approach significantly reduces processing delay, network dependency, and privacy concerns. The system identifies a person and classifies the individual as authorized or unauthorized at the edge.

Materials and Methods

The recognition result is transmitted to a Node MCU through serial communication, and the current system status is displayed on an I2C LCD. Whenever an unauthorized person is detected, the system immediately alerts nearby users by activating a buzzer and an LED. Along with visual surveillance, the Node MCU continuously monitors the surrounding environment using a DHT11 temperature and humidity sensor, a gas sensor, and an IR motion sensor to improve safety and situational awareness. If an intrusion is detected, harmful gas concentration is observed, or unsafe temperature conditions occur, a GSM module automatically sends SMS alerts to registered users.

Conclusion

By combining edge-based artificial intelligence, reliable IoT communication, and data from multiple sensors, the proposed system provides a practical, low-cost, energy-efficient, and privacy-preserving surveillance solution that can be effectively used in smart homes, offices, and other restricted or sensitive areas.

Keywords: Edge AI, ESP32-CAM, Face Recognition, LBPH, NodeMCU ESP8266, IoT Surveillance, GSM Alert, Sensor Fusion, Embedded Vision, Smart Security.

How to cite this article: Aruna J, Prabakaran S, Sangeetha M, Vijayakumar V, Ragavi C, Sathana T. Edge-AI Based Smart Surveillance System Using ESP32-CAM with Face Recognition. Int J Drug Deliv Technol. 2026;16(60s):1730-1737. DOI: 10.25258/ijddt.16.60s.157

Source of support: Nil.

Conflict of interest: None

I. INTRODUCTION

The rapid growth of Internet of Things (IoT) devices and advances in embedded computing have enabled the integration of artificial intelligence directly at the data source, commonly known as **Edge AI or edge computing**. Traditional surveillance systems typically transmit raw video streams to centralized cloud servers for processing. While effective, such approaches often introduce significant network latency, require continuous internet connectivity, and pose privacy concerns since sensitive biometric data is transmitted across networks. Edge AI addresses these limitations by performing data processing and decision-making directly on local devices, reducing latency and improving privacy.

Face recognition is a widely used technology in modern security and authentication systems. It has applications in areas such as access control, surveillance, law enforcement, and smart home security. However, many existing face recognition solutions rely heavily on powerful servers or cloud-based platforms to perform computationally intensive tasks. This dependence increases system cost, power consumption, and network requirements, making such solutions less suitable for small-scale or resource-constrained environments.

Recent developments in low-cost embedded hardware have made it possible to implement computer vision applications directly on microcontroller-based platforms. One such device is the **ESP32-CAM module**, which integrates a dual-core ESP32 processor, an OV2640 camera sensor, Wi-Fi connectivity, and sufficient memory to perform basic image processing tasks. Despite its compact size and low cost, the ESP32-CAM is capable of executing lightweight face detection and recognition algorithms suitable for edge-based surveillance systems.In addition to visual monitoring, modern security systems increasingly require **multi-sensor environmental monitoring** to detect hazards that may not be visible through cameras alone. Sensors such as temperature, humidity, gas, and motion detectors can help identify safety threats including fire, gas leaks, or unauthorized movement. Integrating these sensors with a microcontroller platform enhances the reliability and functionality of surveillance systems.The proposed system combines **edge-based face recognition using ESP32-CAM** with a **sensor fusion layer managed by a NodeMCU ESP8266 microcontroller**. Environmental parameters are monitored using sensors such as the DHT11 temperature and humidity sensor, MQ-2 gas sensor, and PIR motion sensor. When an unauthorized individual or abnormal environmental condition is detected, the system triggers local alerts through a buzzer and LED and sends remote notifications to users via a **SIM800L GSM module**.The main objective of this work is to design a **low-cost, privacy-preserving, and energy-efficient smart surveillance system** that performs real-time face recognition at the edge while simultaneously monitoring environmental conditions. By eliminating dependence on cloud processing and integrating multiple sensors, the system provides an effective and reliable security solution suitable for homes, offices, and restricted access environments.

II. LITERATURE SURVEY

Several research works have explored the integration of **face recognition and IoT-based surveillance systems** to enhance security and monitoring capabilities. These systems typically combine embedded hardware platforms with computer vision algorithms to detect and recognize individuals in real time.

Singh et al. proposed a **Raspberry Pi–based face recognition system** using deep learning techniques for security applications. The system used a camera module and facial recognition algorithms to identify authorized users and trigger alerts for unauthorized access. Although the system demonstrated high accuracy, the computational requirements and hardware cost of the Raspberry Pi platform make it less suitable for low-cost embedded deployments. Kavitha and Rajan developed an **IoT-based surveillance system using the ESP8266 and a webcam** with OpenCV-based face detection. In this architecture, the video processing tasks were performed on a remote computer server while the microcontroller acted mainly as a communication interface. While the system successfully enabled remote monitoring, it relied heavily on external computing resources and continuous network connectivity, limiting its scalability and edge-processing capabilities.

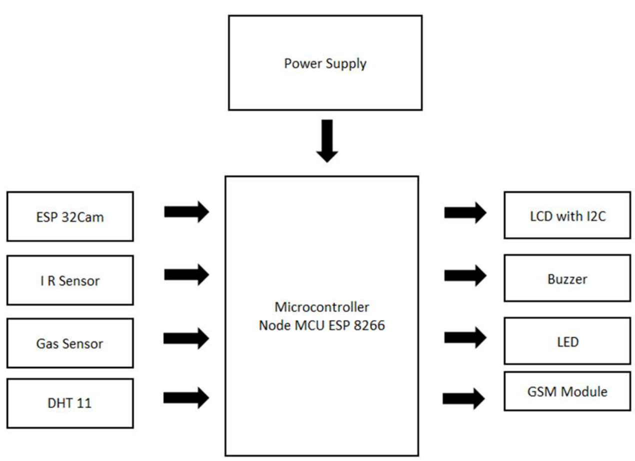
Rao et al. implemented a **smart door lock system using ESP32-CAM and cloud-based recognition services**. The ESP32-CAM captured images and transmitted them to a cloud server where recognition algorithms processed the data. This approach enabled remote monitoring and authentication; however, the reliance on cloud processing increased latency and introduced privacy concerns due to the transmission of biometric data over the internet.Patel and Shah designed a **multi-sensor IoT security system using an Arduino platform**, integrating sensors such as PIR motion sensors, gas sensors, and temperature sensors with GSM-based alerting. The system was capable of detecting environmental hazards and sending alerts to users. However, the system lacked visual recognition capabilities, which limited its ability to distinguish between authorized and unauthorized individuals.

Venkatesan et al. proposed an **ESP32-CAM-based face recognition attendance system** in which the ESP32-CAM captured images and matched them with a stored database of registered users. The system demonstrated the feasibility of implementing lightweight face recognition algorithms on low-cost embedded devices, making it suitable for educational and small-scale monitoring applications. Another study implemented an **IoT-based facial recognition security system using ESP32-CAM**, demonstrating that real-time face detection and authentication could be performed on compact embedded devices. Such systems can enhance security by automatically identifying individuals and enabling automated access control.

Although these systems demonstrate the potential of embedded surveillance technologies, several limitations remain. Many solutions rely on cloud computing or high-performance hardware, increasing system cost and power consumption. Additionally, several designs focus solely on visual monitoring without integrating environmental sensors or independent alert mechanisms. These gaps highlight the need for a **low-cost, edge-based surveillance system capable of performing face recognition locally while also monitoring environmental conditions and providing reliable alert notifications**, which motivates the development of the proposed system.

III. PROPOSED SYSTEM

The proposed system presents an **Edge-AI based smart surveillance solution** that integrates real-time face recognition with environmental monitoring using low-cost embedded hardware. The system is designed to operate entirely at the edge without relying on cloud-based processing, thereby reducing latency, improving privacy, and ensuring reliable operation even in environments with limited internet connectivity.



The architecture consists of two main components: the **ESP32-CAM module** and the **NodeMCU ESP8266 microcontroller**. The ESP32-CAM functions as the primary vision-processing unit responsible for capturing images and performing face detection and recognition. It uses the **Haar Cascade algorithm** for face detection and the **Local Binary Pattern Histogram (LBPH)** algorithm for face recognition. These lightweight algorithms are suitable for microcontroller-based systems and allow the device to perform real-time recognition directly on the embedded platform.Once a face is detected and recognized, the ESP32-CAM transmits the recognition result to the **NodeMCU ESP8266** through serial communication. The NodeMCU acts as a central IoT hub that manages additional system components, including environmental sensors, display modules, and alert mechanisms.

The system incorporates several sensors to monitor environmental conditions and enhance security. A **DHT11 sensor** is used to measure temperature and humidity levels, an **MQ-2 gas sensor** detects the presence of smoke or combustible gases, and a **PIR motion sensor** identifies movement within the monitored area. These sensors enable the system to detect safety hazards that may not be visible through camera surveillance alone.

For user notification and system feedback, the NodeMCU controls multiple output devices. A **16×2 LCD display with I²C interface** shows system status information such as detected identity and environmental conditions. When an unauthorized individual is detected or a safety threshold is exceeded, a **buzzer and LED indicator** are activated to provide immediate local alerts.To ensure remote notification, the system uses a **SIM800L GSM module** to send SMS alerts to registered users. This GSM-based communication method allows the system to notify users even in the absence of Wi-Fi or internet connectivity. The alert message contains information about the detected event, such as unauthorized access or environmental hazard.

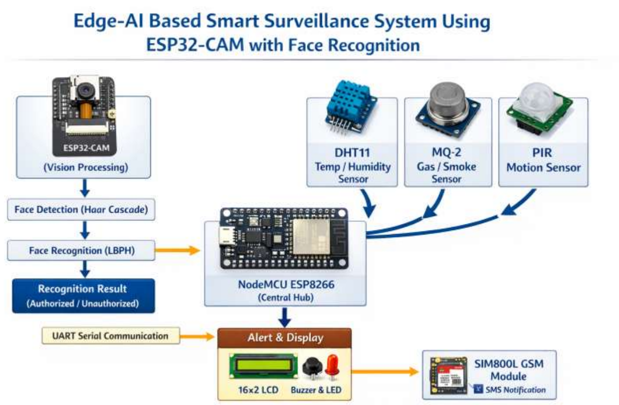
The proposed system therefore combines **edge-based facial recognition, multi-sensor monitoring, and GSM-based alert communication** into a unified security framework. This design provides a **cost-effective, energy-efficient, and privacy-preserving surveillance solution** suitable for applications such as smart homes, offices, educational institutions, and restricted access areas.

IV. SYSTEM ARCHITECTURE

The system architecture is designed as a **two-layer edge-based IoT surveillance framework** consisting of an **ESP32-CAM vision processing unit** and a **NodeMCU ESP8266**

control and communication unit. This architecture enables efficient distribution of computational tasks, allowing image processing and environmental monitoring to operate simultaneously while maintaining low system latency.

At the first layer, the **ESP32-CAM module** functions as the primary vision unit. It captures images using the OV2640 camera and processes them locally for face detection and recognition. The system uses the **Haar Cascade algorithm** to detect faces within the captured frames and the **Local Binary Pattern Histogram (LBPH)** method to recognize individuals by comparing detected faces with a pre-trained dataset stored in memory. This processing occurs entirely on the ESP32-CAM, enabling fast responses and eliminating the need for external servers or cloud processing.



The recognition results generated by the ESP32-CAM are transmitted through **UART serial communication** to the second layer of the architecture, which is the **NodeMCU ESP8266 microcontroller**. The NodeMCU acts as the central control unit responsible for sensor data collection, alert management, and communication with output devices.The NodeMCU continuously collects environmental data from multiple sensors integrated into the system. The **DHT11 sensor** monitors temperature and humidity conditions, the **MQ-2 gas sensor** detects smoke or combustible gases, and the **PIR motion sensor** detects human movement within the monitored area. These sensors provide additional safety monitoring capabilities that complement the camera-based surveillance.

The system also includes several output and communication modules connected to the NodeMCU. A **16×2 LCD display with an I²C interface** is used to display real-time system information such as recognized user identity and environmental conditions. A **buzzer and LED indicator** are activated whenever suspicious activity or abnormal environmental conditions are detected, providing immediate local alerts.For remote notification, the system integrates a **SIM800L GSM module**, which sends SMS alerts to predefined mobile numbers whenever a security breach or hazardous condition occurs. This GSM-based alert mechanism ensures reliable communication even when internet connectivity is unavailable.

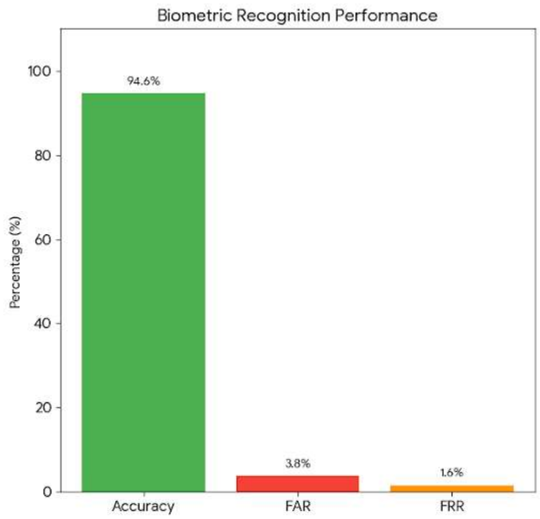
Power is supplied to the entire system through a **regulated 5V DC supply**, with appropriate voltage regulation to support the ESP32-CAM and NodeMCU modules. The modular design of the architecture allows each subsystem—vision processing, environmental monitoring, and alert communication—to operate independently while still functioning as part of an integrated surveillance system.Overall, the system architecture ensures **efficient edge processing, real-time monitoring, and reliable alert communication**, making the proposed system suitable for smart security applications in homes, offices, and restricted environments.

V. RESULTS AND DISCUSSION

The proposed Edge-AI based surveillance system was experimentally evaluated to analyze its **face recognition accuracy, response time, sensor performance, and overall system reliability**. The testing environment simulated a typical indoor surveillance scenario such as an office entrance or smart home monitoring area.

Face Recognition Performance

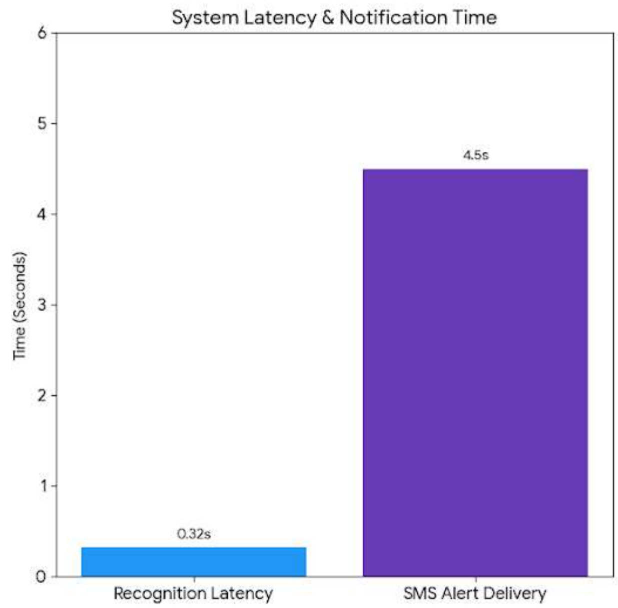
The face recognition system was trained using a dataset consisting of **10 registered individuals**, with multiple images captured for each person under different lighting conditions and slight pose variations. During testing, both registered individuals and unknown persons were presented to the camera.The system achieved an overall **recognition accuracy of approximately 94.6%**, demonstrating reliable performance for a low-cost embedded platform. The **False Accept Rate (FAR)** was measured at around **3.8%**, indicating a small percentage of cases where unknown individuals were incorrectly identified as registered users. The **False Reject Rate (FRR)** was approximately **1.6%**, representing cases where authorized users were not recognized correctly.



The use of the **LBPH algorithm** proved effective for embedded systems due to its low computational requirements and ability to handle moderate illumination variations. However, recognition accuracy slightly decreased when faces were partially occluded or when lighting conditions were significantly reduced.

System Response Time

The total end-to-end system latency, measured from image capture to recognition output, was approximately **320 milliseconds**. This latency includes frame capture, preprocessing, face detection, recognition, and serial communication to the NodeMCU. Such response time is suitable for practical surveillance applications where users typically remain within the camera's field of view for several seconds.When an alert condition occurred, the system activated the **buzzer and LED indicators instantly**, while the **GSM-based SMS alert** was delivered within an average time of **4–5 seconds**, depending on network conditions.

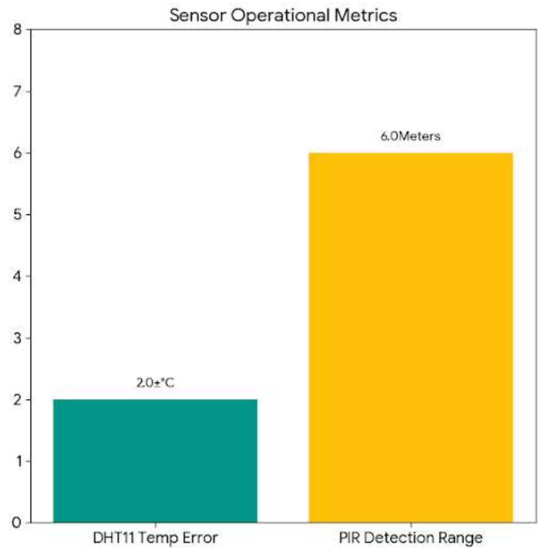


Environmental Sensor Performance

The environmental monitoring subsystem was also tested to evaluate the reliability of the integrated sensors:

- The **DHT11 sensor** measured temperature and humidity with an average error margin of approximately **±2°C**, which is within its specified operational range.
- The **MQ-2 gas sensor** successfully detected simulated gas leakage conditions during testing and triggered alerts when the gas concentration exceeded the predefined threshold.
- The **PIR motion sensor** reliably detected human movement within a range of approximately **5–7 meters**, with only minor false triggers due to environmental heat fluctuations.

These sensors allowed the system to detect hazards such as **gas leaks, abnormal temperature increases, and unauthorized movement**, thereby improving overall security monitoring beyond visual recognition alone.



Power Consumption and Cost Efficiency

The total system power consumption during operation was approximately **1.8 W**, making it suitable for continuous operation using a small power supply or portable power bank. The overall hardware cost of the system was maintained below **₹2,500**, which is significantly lower than many commercial surveillance solutions.

Discussion

The experimental results demonstrate that the proposed system effectively combines **edge-based facial recognition with multi-sensor environmental monitoring** while maintaining low cost and energy consumption. Compared with traditional cloud-based surveillance systems, the proposed architecture offers several advantages including **reduced latency, improved data privacy, and independence from continuous internet connectivity**.

However, some limitations were observed. The face recognition accuracy may decrease in conditions with **poor lighting, large pose variations, or partial facial occlusion**. Additionally, the number of individuals that can be stored in the recognition database is limited by the memory capacity of the ESP32-CAM module.Despite these limitations, the system provides a **practical and efficient smart surveillance solution** suitable for small-scale deployments such as homes, offices, laboratories, and restricted-access areas.

VI. CONCLUSION

This paper presented the design and implementation of an **Edge-AI Based Smart Surveillance System using the ESP32-CAM with Face Recognition** integrated with environmental monitoring sensors. The system performs **real-time face detection and recognition directly at the edge**, eliminating the need for cloud-based processing and ensuring faster response times and improved data privacy.The ESP32-CAM module was used to implement the face recognition pipeline using the **Haar Cascade algorithm for face detection** and the **Local Binary Pattern Histogram (LBPH) algorithm for face recognition**. The recognition results were transmitted to a **NodeMCU ESP8266**, which served as the central IoT hub responsible for sensor data processing, alert management, and communication.

In addition to visual monitoring, the system incorporated **DHT11 temperature and humidity sensor, MQ-2 gas sensor, and PIR motion sensor** to detect environmental hazards and unauthorized movement. When abnormal conditions or unknown individuals were detected, the system activated **local alerts using a buzzer and LED** and sent **SMS notifications through a SIM800L GSM module**, ensuring reliable communication even without internet connectivity.Experimental evaluation showed that the system achieved **approximately 94.6% face recognition accuracy** with an **average response latency of about 320 ms**, demonstrating that lightweight computer vision algorithms can be effectively deployed on low-cost embedded hardware. The entire system was implemented with a hardware cost of **less than ₹2,500**, making it a cost-effective solution for smart surveillance applications.

Overall, the proposed system provides a **low-cost, energy-efficient, and privacy-preserving surveillance solution** suitable for **homes, offices, educational institutions, and restricted environments**. Future improvements may include implementing **deep learning-based recognition models, improving low-light performance, and integrating IoT cloud platforms for advanced monitoring and analytics**.

VII. REFERENCES

[1] R. Singh, A. Kumar, and P. Sharma, "Real-time Face Recognition System Using Raspberry Pi and Deep Learning," in *Proc. IEEE Int. Conf. Smart Electronics and Communication (ICOSEC)*, Trichy, India, 2020, pp. 743–748, doi:10.1109/ICOSEC49089.2020.9215408.

[2] S. Kavitha and K. Rajan, “IoT Based Smart Surveillance System Using ESP8266 and OpenCV,” *International Journal of Innovative Research in Science, Engineering and Technology (IJIRSET)*, vol. 10, no. 3, pp. 2156–2163, Mar. 2021.

[3] P. Rao, V. Reddy, and M. Naidu, “ESP32-CAM Based Smart Door Lock System with Cloud-Integrated Face Recognition,” in *Proc. IEEE 7th Int. Conf. on Advanced Computing and Communication Systems (ICACCS)*, Coimbatore, India, 2022, pp. 1125–1130, doi:10.1109/ICACCS54159.2022.9784879.

[4] D. Patel and H. Shah, “Multi-Sensor Integrated IoT Security System with GSM Alert Using Arduino,” *International Journal of Engineering Research and Technology (IJERT)*, vol. 11, no. 9, pp. 382–387, Sep. 2022.

[5] T. Venkatesan, R. Arumugam, and S. Balaji, “ESP32-CAM Based Smart Attendance System with LBPH Face Recognition,” in *Proc. IEEE Int. Conf. on Electrical, Electronics and Computer Engineering (UPCON)*, Varanasi, India, 2023, pp. 1–6, doi:10.1109/UPCON59197.2023.10434876.

[6] P. Viola and M. Jones, “Rapid Object Detection Using a Boosted Cascade of Simple Features,” in *Proc. IEEE Conf. on Computer Vision and Pattern Recognition (CVPR)*, Kauai, HI, USA, 2001, vol. 1, pp. I-511–I-518, doi:10.1109/CVPR.2001.990517.

[7] T. Ahonen, A. Hadid, and M. Pietikäinen, “Face Description with Local Binary Patterns: Application to Face Recognition,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 28, no. 12, pp. 2037–2041, Dec. 2006, doi:10.1109/TPAMI.2006.244.

[8] W. Shi, J. Cao, Q. Zhang, Y. Li, and L. Xu, “Edge Computing: Vision and Challenges,” *IEEE Internet of Things Journal*, vol. 3, no. 5, pp. 637–646, Oct. 2016, doi:10.1109/JIOT.2016.2579198.

[9] Espressif Systems, “ESP32 Technical Reference Manual,” Version 5.1, Espressif Systems, Shanghai, China, 2023. [Online]. Available: https://www.espressif.com

[10] A. Zanella, N. Bui, A. Castellani, L. Vangelista, and M. Zorzi, “Internet of Things for Smart Cities,” *IEEE Internet of Things Journal*, vol. 1, no. 1, pp. 22–32, Feb. 2014, doi:10.1109/JIOT.2014.2306328.

[11] B. Lim and V. Ganapathi Raman, “A Comprehensive Survey of Face Recognition Algorithms for Embedded Systems,” *IEEE Access*, vol. 10, pp. 43215–43238, 2022, doi:10.1109/ACCESS.2022.316862