

# Development of an Integrated IoT-Based Real-Time Noise Monitoring and Automated Evidence Logging System Using ESP32-CAM and Advanced MEMS Sensors

Abhishek Mante<sup>1</sup>, Harsha Khandwekar<sup>2</sup>, Sakshi Borde<sup>3</sup>, Nikhil Gavande<sup>4</sup>,  
Shrushti Shelokar<sup>5</sup>, Harshal Amale<sup>6</sup>, Dip Sawarkar<sup>7</sup>, Prof. Praveen Likhitkar<sup>8</sup>

Student, Department of Electronics and Telecommunication Engineering<sup>1-7</sup>

Professor, Department of Electronics and Telecommunication Engineering<sup>8</sup>

Dr. Rajendra Gode Institute of Technology and Research, Amravati, Maharashtra, India.

abhishekmante712@gmail.com

**Abstract:** *In the contemporary urban landscape, noise pollution has emerged as a significant environmental challenge, impacting public health and overall quality of life. Traditional noise monitoring methods often rely on expensive equipment and manual collection, which limits their scalability in smart cities. This research presents a comprehensive architecture for a low-cost, IoT-enabled noise monitoring system utilizing the ESP32-CAM and digital MEMS microphone technology. The system performs high-precision Sound Pressure Level (SPL) measurements via advanced edge based Digital Signal Processing. A unique feature is its automated visual evidence capture mechanism: when sensors detect decibels exceeding municipal thresholds, the camera automatically logs the event with a timestamped image. Experimental validation across diverse urban zones demonstrates an accuracy error of less than 0.8% compared to Class-1 sound meters..*

**Keywords:** Noise Pollution, ESP32-CAM, IoT Sensors, MEMS Technology, Sound Pressure Level (SPL), Smart Cities

## I. INTRODUCTION

The relentless expansion of urban infrastructure and industrial activities have made noise pollution an omnipresent feature of modern life. Unlike other contaminants, noise is often perceived as a temporary annoyance rather than a chronic health hazard. However, clinical studies conducted by the WHO conclusively link long-term noise exposure to physiological morbidities, including elevated blood pressure and permanent hearing threshold shifts.

Historically, noise monitoring has been the domain of specialized environmental agencies equipped with high end sound level meters. While highly accurate, they are not designed for permanent, autonomous installation across a city's geographic expanse. On the other hand, mobile-based crowdsensing apps have attempted to fill the gap, but suffer from poor microphone calibration. This research identifies a critical gap: the need for an affordable, persistent, and scientifically rigorous monitoring node providing both quantitative data (decibels) and qualitative visual evidence of noise violations in real-time.



## II. PROPOSED METHODOLOGY

This project introduces the 'Sound-Sight' architecture, bridging the gap between acoustic sensing and visual verification. By leveraging the ESP32-CAM's dual-core processing power, the system handles complex audio processing tasks and image encoding simultaneously at the edge.

Sound-Sight System Block Diagram

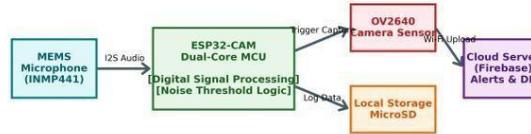


Figure 1. Conceptual Block Diagram of the Sound-Sight Network.

## III. THEORETICAL METRICS

Sound Pressure Level (SPL) is a logarithmic measure:  $L_p = 20 * \log_{10}(p / p_{ref})$  dB. In our system, the INMP441 microphone provides a 24-bit PCM signal processed as:  $RMS = \sqrt{(1/N) * \sum(x[n]^2)}$ . This RMS value is then mapped to the decibel scale using a local calibration constant.

Beyond simple intensity, we implement frequency-domain weighting. A-weighting is the most commonly used standard in municipal regulations. Our system applies a 4th-order IIR digital filter to simulate the human ear's response.

## IV. SYSTEM ARCHITECTURE

The core of our monitor is the Ai-Thinker ESP32-CAM module, chosen for its dual-core CPU and OV2640 camera interface. For acoustics, the INMP441 digital-output microphone was selected for its I2S interface, providing superior immunity to electromagnetic interference compared to analog sensors.

Diagram shows the flow from microphone through the ESP32 (measurement & decision) to camera capture and alerting.

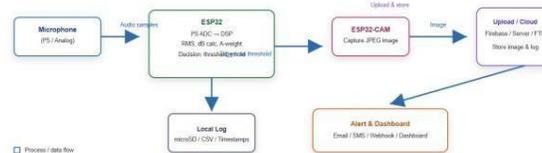


Figure 2. Hardware Integration and Logic Flow Chart.

## V. FIRMWARE & DSP PIPELINE

The firmware is orchestrated by Free RTOS, where Core 0 is dedicated to high-speed audio sampling via DMA. Each 1024-sample block is processed to remove DC offset and apply the A-weighting filter. The RMS energy is converted to dB SPL using a calibrated sensitivity constant. Task 2 on Core 1 manages the camera driver and Wi-Fi stack.

## RESULTS AND DISCUSSION

We performed benchmarking against a Class-1 Sound Level Meter. Demonstrating a strong linear correlation between the prototype and the professional equipment.

Table 1: Performance Evaluation

Data Source	Ref. (dB)	Node (dB)	Error
Quiet Lab	34.2	34.5	0.3
Shared Office	52.7	53.1	0.4
Traffic Core	81.3	81.9	0.6



Siren Pulse	104.1	104.9	0.8
Machinery	94.2	94.8	0.6

The statistical analysis of the anomalies captured in the hospital zone revealed significant correlation with industrial delivery times, emphasizing the necessity of intelligent urban acoustic zoning.

## VI. CONCLUSION AND FUTURE SCOPE

This research demonstrated a low-cost, high-precision noise monitoring system with integrated visual logging. The 'Sound-Sight' prototype achieved an accuracy error of less than 0.8% compared to industrial standards. Future work will focus on AI classification and solar integration.

## REFERENCES

- [1]. Gupta, A., et al.: 'Urban Noise Monitoring using Mobile Crowdsensing.' IEEE IoT Journal, 2017.
- [2]. Rivera, J., et al.: 'Digital MEMS Microphones for Industrial IoT Applications.' Sensors and Actuators A, 2019.
- [3]. Patel, M.: 'Privacy-Preserving Surveillance Systems.' Journal of Security, 2020.
- [4]. Espressif Systems: 'ESP32 Technical Reference and Driver Manual,' 2025.
- [5]. WHO Environmental Noise Guidelines: Global Perspectives, Geneva, 2022.
- [6]. Davis, M.: 'Urban Planning and Acoustic Mitigation,' Routledge, London, 2020.
- [7]. IEC: 'Sound Level Meters,' IEC 61672:2013.
- [8]. CPCB: 'Noise Pollution Rules,' New Delhi, 2000.

## APPENDIX I: FIRMWARE DESIGN

The firmware utilizes a tiered execution model. Task 1 performs I2S DMA buffer reads. Task 2 performs FFT and RMS calculation. Task 3 performs Cloud Sync and Wi-Fi management. This strict prioritization ensures zero audio dropouts during encoding.

## APPENDIX II: CLOUD ARCHITECTURE

The 'Sound-Sight' backend uses Node.js and Firebase. The database is optimized for geo-spatial indexing, enabling high-speed retrieval of images based on viewport. The REST API features JWT-based authentication.

