IJIREEICE ISSN (O) 2321-2004, ISSN (P) 2321-5526

International Journal of Innovative Research in Electrical, Electronics, Instrumentation and Control Engineering

Impact Factor 8.414 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 14, Issue 3, March 2026

DOI: 10.17148/IJIREEICE.2026.14350

AI Based Audio Surveillance System

Ms. Aphasana Mulla[1], Vedanti Thakur[2], Anaya Kumbhar[3], Swanandi Sawant[4], Vanshika Vhatkar[5]

Lecturer, Electronic and Telecommunication, Bharti Vidyapeeth Institute of Technology, Navi Mumbai India[1]

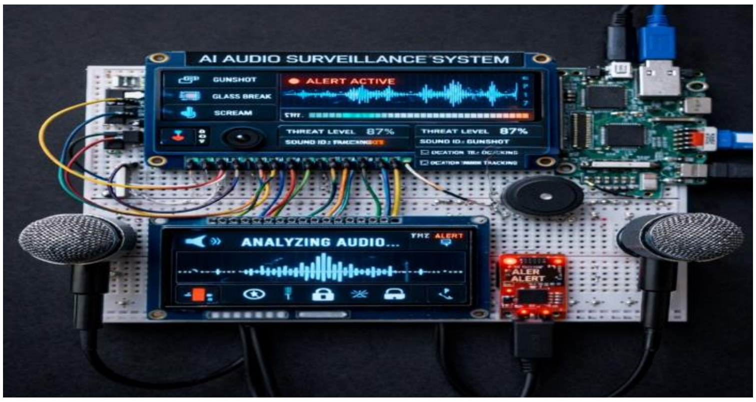
Students, Electronic and Telecommunication, Bharti Vidyapeeth Institute of Technology, Navi Mumbai India[2-5]

Abstract: The intelligent "AI-based audio surveillance system" is very good at listening to the sounds around us. It can hear what is going on and figure out what is important. This system can tell the difference between sounds and sounds that are not normal. Even sounds like gunshots, screams, or an alarm going off can be picked up by it. The audio surveillance system works well even when we cannot see what is happening. When it hears something it will send out an alert right away. This means that people do not have to sit and listen all the time. The audio surveillance system helps us respond quickly when something bad happens. It is really helpful, during emergencies. The audio surveillance system is very useful because it can detect sound events and send alerts. Overall, AI-based audio surveillance enhances security and public safety

I. INTRODUCTION

In the modern world, safety and security have become critical concerns due to rapid urbanization, increasing population, and rising crime rates. Public places such as railway stations, airports, shopping malls, educational institutions, hospitals, and residential societies require continuous monitoring to ensure the safety of people and property. Surveillance systems play an important role in maintaining security by observing activities and detecting suspicious behavior. Traditionally, surveillance systems mainly rely on video cameras (CCTV) for monitoring. Although video surveillance has proven useful, it has several limitations such as dependency on lighting conditions, camera blind spots, occlusion, and the requirement of continuous human supervision.

To overcome these limitations, intelligent surveillance methods using artificial intelligence (AI) have been introduced. Audio surveillance is an emerging technology that focuses on monitoring environmental sounds to identify abnormal or suspicious events. By combining audio sensing with artificial intelligence, surveillance systems can automatically detect dangerous situations even when visual data is unavailable. The AI Powered Audio Surveillance System is designed to provide an intelligent, automated, and reliable solution for enhancing security and safety.



II. EXISTING SYSTEM AND ITS LIMITATION

Existing surveillance systems use traditional monitoring techniques to ensure safety and security in public and private areas.

- CCTV-Based Video Surveillance: CCTV systems use cameras to monitor activities visually. The video is recorded or observed live by security personnel.
- Manual Human Monitoring: Security staff manually monitor video screens to detect suspicious activities. This method depends completely on human attention.

IJIREEICE ISSN (O) 2321-2004, ISSN (P) 2321-5526
International Journal of Innovative Research in Electrical, Electronics, Instrumentation and Control Engineering
Impact Factor 8.414 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 14, Issue 3, March 2026
DOI: 10.17148/IJIREEICE.2026.14350

- Motion Detection Systems : Motion sensors detect movement and trigger alarms. These systems are mainly used in restricted areas and homes.
- Simple Audio Alarm Systems : These systems detect loud sounds and trigger alerts but cannot classify different types of sounds as they do not use AI.

III. PROBLEM STATEMENT

In many public and private places, security mainly depends on CCTV cameras and manual monitoring by security personnel. These traditional surveillance systems require continuous human attention and may not work effectively in low light, crowded areas, or locations with limited camera coverage. As a result, some dangerous situations may not be detected on time.

Many critical incidents such as screams for help, gunshots, explosions, or aggressive arguments are first indicated through sound rather than visual observation. However, existing surveillance systems are not capable of monitoring and analyzing audio signals intelligently. Continuous manual monitoring is tiring and can lead to human errors, increasing the risk of missed incidents.

To overcome these problems, users need an intelligent system that can automatically detect abnormal sounds and provide instant alerts. The AI Powered Audio Surveillance System addresses this issue by continuously monitoring environmental sounds and using artificial intelligence to identify suspicious audio events, thereby improving safety and reducing human effort.

IV. SCOPE OF PROJECT

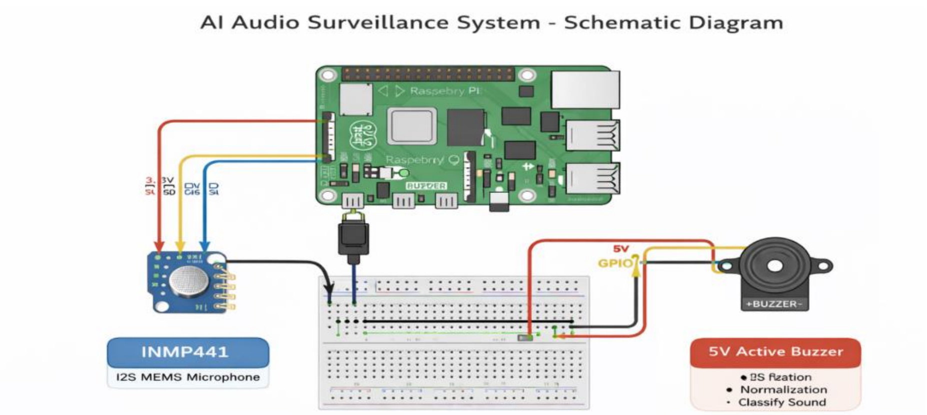
The main aim of this project is to reduce dependence on continuous human surveillance while improving public safety as part of a smart city initiative. Traditional security systems mainly rely on CCTV cameras and manual monitoring, which may fail to always detect critical situations. Many incidents such as screaming, glass breaking, fights, or emergency situations often go unnoticed, especially in crowded areas or during night hours. This creates the need for an intelligent audio-based surveillance solution.

The AI-Powered Audio Surveillance System continuously monitors surrounding sounds and identifies abnormal or suspicious audio events. In many cases, incidents occur outside the camera's field of view or in locations where cameras are not installed. Audio surveillance helps overcome this limitation by detecting unusual sounds such as distress cries, alarms, or aggressive noises and sending immediate alerts to concerned authorities, thereby improving response time.

This project also reduces human effort by automating sound monitoring using artificial intelligence. Instead of manually observing live feeds, the system analyzes audio data in real time and generates alerts when necessary. This enhances efficiency and minimizes delays during emergency situations.

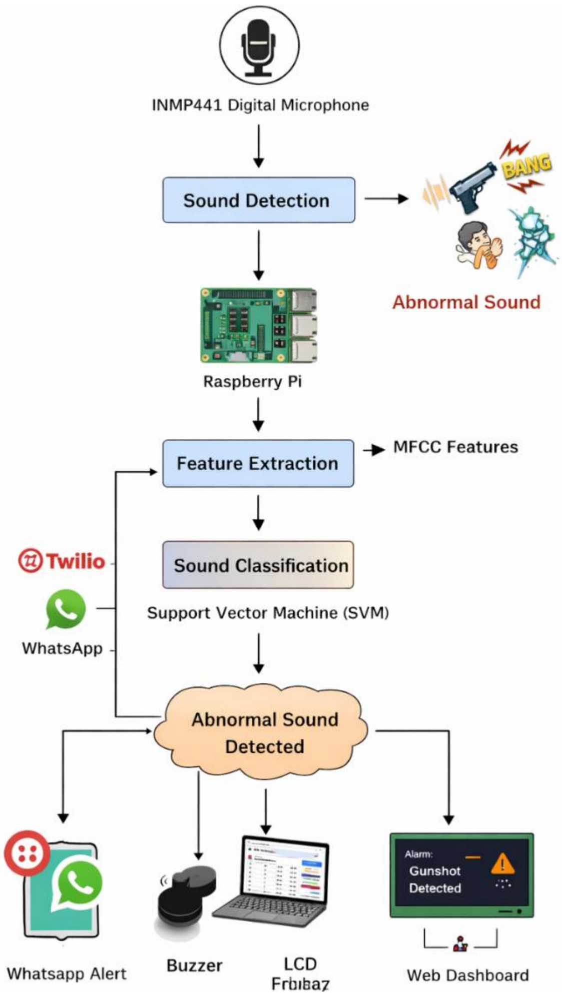
The proposed system supports the Smart City vision by providing a reliable, cost-effective, and privacy-aware surveillance solution. With increasing security challenges in public places like schools, hospitals, railway stations, and residential areas, this system offers an effective approach to improving safety, preventing incidents, and creating a secure environment.

SCHEMATIC DIAGRAM

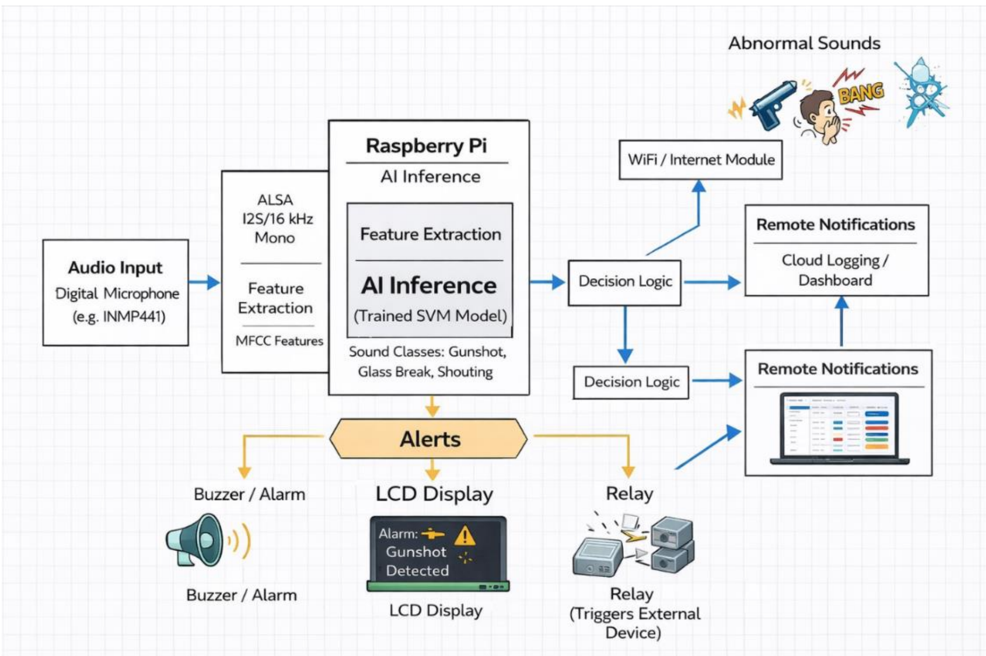


IJIREEICE ISSN (O) 2321-2004, ISSN (P) 2321-5526
International Journal of Innovative Research in Electrical, Electronics, Instrumentation and Control Engineering
Impact Factor 8.414 ⁙ Peer-reviewed & Refereed journal ⁙ Vol. 14, Issue 3, March 2026
DOI: 10.17148/IJIREEICE.2026.14350

FLOW CHART



BLOCK DAIGRAM OF AI BASED AUDIO SURVEILLANCE SYSTEM



International Journal of Innovative Research in Electrical, Electronics, Instrumentation and Control Engineering
Impact Factor 8.414 ⁙ Peer-reviewed & Refereed journal ⁙ Vol. 14, Issue 3, March 2026
DOI: 10.17148/IJIREEICE.2026.14350

V.WORKING

The AI-Powered Audio Surveillance System works by continuously capturing environmental sounds, processing them using artificial intelligence, and generating alerts when abnormal sounds are detected. The complete working process is explained step-by-step below:

A. Hardware Components Description:

1. Connecting Wires / Jumper Wires: Jumper wires are essential for creating flexible, temporary electrical connections between the core components on the breadboard and the Raspberry Pi's GPIO pins.

- Description: A set of multi-colored female-to-female jumper wires, 20 cm in length. The female pins are designed to connect securely onto male headers (like those on the Raspberry Pi GPIO or component pins).
- Function: Facilitates prototyping and allows for rapid changes to the circuit design during development.
- Specifications:

Quantity: Set of 10-40 wires.
Length: 20 cm.
Gauge: 26 AWG.
Current Rating: Up to 1 A.
Insulation Type: PVC (Polyvinyl Chloride).
Reusability: High reusability for prototyping.

2. Primary Controller & AI Processor: Raspberry Pi 4 Model B (8GB RAM):
The Raspberry Pi 4 serves as the central processing unit (CPU) for the entire system. Crucially, the 8GB RAM variant is selected to accommodate the memory-intensive requirements of AI/ML audio classification libraries (e.g., Librosa, TensorFlow Lite) and the Linux operating system. This is a significant upgrade over 2GB/4GB models for AI tasks.

- Microcontroller: Broadcom BCM2711, Quad-core Cortex-A72 (ARM v8) 64-bit SoC.
- Clock Speed: 1.5GHz.
- Operating Voltage: 5.1V via USB-C (requires 3.0A power supply).
- Memory (RAM): 8GB LPDDR4-3200 SDRAM (Essential for running the AI models).
- Connectivity: Gigabit Ethernet Port.
- I/O Pins: 40-pin GPIO (General Purpose Input/Output) header for connecting the microphone and buzzer.

3. Audio Input Sensor: INMP441 I2S MEMS Microphone:
This component is the primary sensor for the audio surveillance system. It is a high-performance, low-power, digital-output MEMS (Micro-Electro-Mechanical System) microphone. It uses the I2S digital audio interface, which provides a cleaner, more robust audio signal to the Raspberry Pi compared to analog microphones which require external analog-to-digital converters (ADCs).

- Interface Type: I2S (Inter-IC Sound) digital interface (no analog conversion needed).
- Signal-to-Noise Ratio (SNR): 61 dBA (Excellent clarity).
- Sensitivity: -26 dBFS.
- Frequency Response: 60 Hz to 15 kHz (Captures the full range of human speech and typical security-relevant sounds like glass breaking).
- Pickup Pattern: Omnidirectional (captures sound from all directions).
- Operating Voltage: 1.8V to 3.3V (Safely interfaces with RPi GPIO logic levels).

4. Prototyping Board (Breadboard GL-12 840 Points):
The breadboard is our primary tool for assembling and testing the circuit without permanent soldering. It allows us to safely connect the microphone and the buzzer to the Raspberry Pi's GPIO pins.

- Description: A high-quality, solderless prototyping board with 830 connection points.
- Features: Connection Points: Arranged into terminal strips (for components) and bus strips (for power and ground lines).

Reusability: Allows components to be easily plugged in and removed.
Design: Features interlocking projectors around the edges, allowing multiple boards to be combined for larger projects.
Adhesive Backing: Optional adhesive tape on the back for permanent mounting within an enclosure.

International Journal of Innovative Research in Electrical, Electronics, Instrumentation and Control Engineering
Impact Factor 8.414 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 14, Issue 3, March 2026
DOI: 10.17148/IJIREEICE.2026.14350

5. Alert Output: 5V Active Buzzer:
When the AI system detects a security event (e.g., a gunshot or glass break), this component provides an immediate, localized audible alert.

- Type: Active buzzer. This means it contains its own internal oscillation circuit and produces a continuous sound when a stable DC voltage (5V) is applied to its terminals via a Raspberry Pi GPIO pin.
- Operating Voltage: 3V to 5V DC.
- Current Draw: Typically < 30mA.
- Connectivity: 2-pin interface (VCC and GND), easily controlled by a single GPIO pin (High/Low signal).

6. Power Supply Unit (PSU): Official Raspberry Pi 15.3W USB-C PSU:
A stable, high-current power supply is vital for the reliable operation of an AI system. The official PSU is specifically designed for the Pi 4 to prevent unexpected restarts or brownouts that can corrupt data or interrupt surveillance.

- Input: 100-240VAC (Standard wall voltage).
- Output: 5.1V DC at 3.0A (15.3 Watts).
- Connector: USB Type-C plug.
- Safety: Includes built-in short circuit, over-current, and over-voltage protection.

B. Software: First, we import the necessary libraries for audio processing, machine learning classification, and communication.

```
import numpy as np            # For mathematical calculations
import librosa                # For audio feature extraction (MFCC)
import subprocess             # To record audio using ALSA
from joblib import load       # To load the trained SVM model
import RPi.GPIO as GPIO       # To control the buzzer
from twilio.rest import Client  # To send WhatsApp alerts
import time
import os
```

After that, we define the audio parameters and the GPIO pin for the buzzer alert.

```
BUZZER_PIN = 23               # GPIO pin for the Active Buzzer
DURATION = 2                  # Audio window size in seconds
ENERGY_THRESHOLD = 0.01       # Minimum energy required for detection
ALERT_SOUNDS = ["gunshot", "shouting", "glass"]  # Abnormal sound classes
```

After that, we configure the Twilio settings to send security alerts to the user's mobile through WhatsApp.

```
ACCOUNT_SID = "Your_Account_SID"
AUTH_TOKEN = "Your_Auth_Token"
FROM_WHATSAPP = "whatsapp:+14155238886"
TO_WHATSAPP = "whatsapp:+91XXXXXXXXXX"
client = Client(ACCOUNT_SID, AUTH_TOKEN)
```

Now, to process the audio, we load the trained SVM model. This loads the pre-trained `audio_model.joblib` file into the Raspberry Pi's RAM for high-speed analysis.

```
model = load("audio_model.joblib")
```

Inside the main execution loop, the system captures a 2-second audio buffer from the INMP441 microphone using the ALSA recording utility.

```
subprocess.run([
    "arecord",
    "-D", "plughw:2,0",
    "-f", "S16_LE",
    "-r", "16000",
    "-c", "1",
    "-d", str(DURATION),
    "live_test.wav"
])
```

IJIREEICE ISSN (O) 2321-2004, ISSN (P) 2321-5526

International Journal of Innovative Research in Electrical, Electronics, Instrumentation and Control Engineering

Impact Factor 8.414 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 14, Issue 3, March 2026

DOI: 10.17148/IJIREEICE.2026.14350

Now, the recorded audio file is loaded and normalized to remove amplitude variations.

```
audio, sr = librosa.load("live_test.wav", sr=22050)
audio = librosa.util.normalize(audio)
```

After that, the system calculates the RMS energy of the sound signal to filter out very low background noise.

```
energy = np.sqrt(np.mean(audio**2))
```

If the energy is below the defined threshold, the system ignores the sound and continues monitoring.

```
if energy < ENERGY_THRESHOLD:
    continue
```

Next, the software extracts MFCC (Mel Frequency Cepstral Coefficients) features from the audio signal. In this system, 13 MFCC coefficients are calculated and averaged to create a feature vector.

```
mfcc = librosa.feature.mfcc(y=audio, sr=sr, n_mfcc=13)
features = np.mean(mfcc.T, axis=0)
```

Now, the system performs the inference process. The trained Support Vector Machine (SVM) classifier compares the extracted features with learned patterns such as "Normal", "Gunshot", "Glass Breaking", or "Shouting".

```
probs = model.predict_proba([features])[0]
confidence = np.max(probs)
prediction = model.classes_[np.argmax(probs)]
```

After that, the system calculates the confidence level of the detected sound. This represents the probability that the detected sound belongs to a particular class.

If the confidence is above 50% and the sound belongs to abnormal categories, the system triggers the buzzer and sends a WhatsApp alert.

```
if prediction in ALERT_SOUNDS and confidence > 0.50:
    GPIO.output(BUZZER_PIN, GPIO.HIGH)
    time.sleep(1)
    GPIO.output(BUZZER_PIN, GPIO.LOW)

    client.messages.create(
        from_=FROM_WHATSAPP,
        body=f"🚨 ALERT! Detected Sound: {prediction}",
        to=TO_WHATSAPP
    )
```

If the detected sound does not meet the abnormal condition, the buzzer remains OFF.

The system then clears the audio buffer and restarts the monitoring loop, providing continuous 24/7 real-time audio surveillance using AI-based classification on the Raspberry Pi.

SPECIFICATIONS OF MALE TO FEMALE CONNECTION WIRES:

1. Quantity: Set of 10-40 wires.
2. Length: 20 cm.
3. Gauge: 26 AWG.
4. Current Rating: Up to 1 A.
5. Insulation Type: PVC (Polyvinyl Chloride).
6. Reusability: High reusability for prototyping.

IJIREEICE ISSN (O) 2321-2004, ISSN (P) 2321-5526

International Journal of Innovative Research in Electrical, Electronics, Instrumentation and Control Engineering

Impact Factor 8.414 ⁙ Peer-reviewed & Refereed journal ⁙ Vol. 14, Issue 3, March 2026

DOI: 10.17148/IJIREEICE.2026.14350

SETUP IS GIVEN BELOW

VI. CONCLUSION

This project work successfully demonstrates the design and implementation of an AI-Powered Audio Surveillance System using a microphone sensor, Raspberry Pi, artificial intelligence techniques, and Wi-Fi connectivity. The system continuously monitors surrounding sounds and processes audio data in real time to identify abnormal or suspicious sound events such as screams, alarms, glass breaking, or aggressive noises. The proposed system reduces the dependency on continuous human monitoring by automatically analyzing audio signals and generating alerts when critical situations are detected. Due to local processing on the Raspberry Pi, the system provides faster response time, improved reliability, and enhanced privacy. The alert mechanism ensures that security personnel or concerned authorities can take timely action, thereby minimizing risks and improving overall safety. Compared to traditional surveillance systems that rely mainly on video monitoring, the AI-Powered Audio Surveillance System works effectively even in low-visibility conditions and camera blind-spot areas. Overall, the system offers a practical, intelligent, and cost-effective solution for enhancing safety and security in public and private environments, supporting modern smart surveillance and smart city applications

REFERENCES

[1]. M. Cobos, J. C. Guirado, and P. F. Alcantara, "AI-Based Audio Surveillance Systems: Techniques and Applications," *Journal of Signal Processing Systems*, vol. 91, pp. 123–140, 2019.

[2]. S. R. N. van de Par et al., "Audio Event Detection for Smart Security Applications," *IEEE Access*, vol. 7, pp. 150234–150245, 2019.

[3]. ISO 22341:2019, "Security and resilience — Guidelines for audio surveillance systems," International Organization for Standardization, 2019.

[4]. D. Stowell and M. D. Plumbley, "Automatic Large-Scale Classification of Environmental Sounds," *IEEE Transactions on Audio, Speech, and Language Processing*, vol. 19, no. 5, pp. 117–127, 2011.

[5]. C. Giannoulis et al., "Detection and Classification of Acoustic Events in Smart Environments," *EURASIP Journal on Audio, Speech, and Music Processing*, 2013.

[6]. A. Krizhevsky, I. Sutskever, and G. Hinton, "ImageNet Classification with Deep Convolutional Neural Networks," in *Advances in Neural Information Processing Systems (NeurIPS)*, 2012.

[7]. "AI in Smart Surveillance: Trends and Future Prospects," *TechTarget*, [Online]. Available: https://www.techtarget.com. Accessed: Dec. 17, 2025.

