AI-Powered Predictive Traffic Signal Control for Emergency Response Optimization

Dr N Vini Antony Grace [a], Bhuvanashree T [b], Najmus Sahar M I [c], Oduru Darshitha [d]

[a] *Associate Professor/ Department of ECE,* R.M.D. Engineering College, TamilNadu - 601206

[b,c,d,e] *UG Student/ Department of ECE,* R.M.D. Engineering College, TamilNadu - 601206

E-mail: [a] nv.ece@rmd.ac.in,[b]22104024@rmd.ac.in,[c]22104100rmd.ac.in,[d]22104108@rmd.ac.in

***Abstract*— The rapid movement of emergency vehicles through intersections is very important for efficient response time and saving lives. The current control systems rely mainly on GPS technology for preemption and sirens for manual detection, which might fail because of connectivity problems and lack of autonomous detection systems. The paper proposes an efficient and cost-effective solution for an edge-based solution for emergency vehicle detection with an ESP32-CAM and a microphone module for the analysis of sirens in real time. The solution analyzes the frequency range of sirens (700-1600 Hz) using Fast Fourier Transform (FFT) analysis and then uses a direction estimation module for the source of the siren. The camera module is directed towards the source of the siren. The optional solution for better prediction is a GPS app for drivers. The results show that the solution has an accuracy of 95% for real-time siren detection with a response time of 2 seconds from the system.**

***Keywords*— Emergency vehicle prioritization, Intelligent traffic system, IoT, AI, ESP32-CAM, FFT, Edge Computing**

I. INTRODUCTION

Rapid emergency response plays an important role in reducing the number of fatalities and damage to properties in the event of medical emergency, fire, and law enforcement operations. In the case of densely populated cities, traffic congestion at intersections causes considerable delay for emergency vehicles, such as ambulances and fire trucks. This can be critical for the survival rates of patients.

Existing traffic control systems mainly employ time control, manual monitoring, and GPS signal preempt systems. Though the GPS signal preempt system can provide predictive traffic clearance, it is based on network connectivity. In contrast, the sound-based emergency detection system can provide an independent and autonomous method for identifying emergency vehicles approaching the intersection. However, the existing systems are mostly non-directional and lack embedded processing.

This paper aims to provide a low-cost emergency detection and directional camera activation system based on the ESP32-CAM module with a microphone sensor. The system carries out real-time Fast Fourier Transform (FFT) analysis for the incoming audio signals, and upon identifying the specific siren frequency bands, which fall in the range of 700 Hz and 1600 Hz, a direction estimation process identifies the location of the sound source, and a servo motor-driven camera is directed in the specific direction for monitoring and verification.

The proposed architecture differs from other cloud-dependent architectures in that it carries out signal processing and decision-making directly on the hardware, making it an edge computing architecture. The proposed architecture has shown 95% siren detection accuracy with a response time of 2 seconds when experimentally implemented under real-time traffic noise conditions.

The proposed architecture provides a scalable and cost-effective solution for smart intersections and IoT-based urban traffic systems.

II. RELATED WORK

Jayakumar et al. [1] have proposed a detection method that utilizes a combination of time and frequency domain features, including STFT and MFCC, for the classification of emergency sirens with robustness in noisy scenarios, showing the effectiveness of dual analysis for the detection. Convolutional neural networks (CNNs) have also been utilized for the detection of emergency vehicles, showing more than 90% accuracy with the use of a spectrogram for distinguishing siren sounds from other environmental sounds [2].

Marchegiani and Newman [3] have proposed a multi-tasking detection method for the localization and classification of siren sounds with the use of a spectrogram for semantic segmentation, showing directional estimation capabilities for the detection. Extreme learning machines (ELMs) have also been utilized for the detection and classification of siren sounds with high detection accuracy and reduced computational complexity [4].

Integrated sound-based traffic control systems for ambulance prioritization have been studied by implementing algorithms for ambulance sirens' detection and identification of directions to achieve real-time green light priority at intersections [5]. Multi-modal frameworks for traffic control systems based on audio, vision, and other sensors have been proposed to address the issue of traffic congestion while prioritizing ambulances [6]. Hybrid control systems have been designed to address traffic optimization and ambulance prioritization within smart city frameworks [7, 15].

GPS and V2I sensor preemption systems enable the pre-emptive adjustment of traffic signals for emergency vehicles and can utilize additional infrastructure [7]. Visual sensing systems coupled with acoustic detection improve the accuracy of traffic control systems [8]. Large data sets of sirens and traffic noise have been developed for acoustic model training for emergency detection systems [9]. Additionally, wearable edge computing devices have been proposed for siren detection to assist visually and hearing-impaired individuals [10]. State-of-the-art deep learning models, such as self-attention CNNs, demonstrate improved accuracy and efficiency in classifying emergency sirens and traffic noise from large data sets [11].

The frameworks for reinforcement learning have been implemented for the optimization of traffic signals, which considers the routing and signal preemption for emergency vehicles, resulting in the minimization of travel time for emergency and regular vehicles alike [12]. Fuzzy logic-based systems have also been modeled for the priority control of emergency vehicles at intersections [13]. Automated traffic signal systems, which respond to ambulance sirens, and audio detection control, which considers emergency vehicle priority, also show the application of real-time detection systems.

Existing methods either lack real-time emergency prioritization, require high infrastructure cost, or depend on centralized processing. There is a critical need for a low-cost, autonomous, and edge-based solution. The proposed system addresses these issues using FFT-based siren detection, directional camera activation using ESP32-CAM, and optional GPS alerts, enabling reliable emergency vehicle detection and monitoring directly at intersections without heavy infrastructure or cloud dependency.

III. EXISTING WORK

CAMERA BASED DETECTION MODEL:

Camera-based models use image processing or vision-based deep learning to detect ambulances and emergency vehicles to override signals. YOLO and CNN are models used to classify vehicles in images of traffic.

The problems associated with camera-based detection models are:

- Affected by light and weather conditions.
- Vehicular occlusion is a major problem in camera-based detection models.
- Computation and processing power are required in camera-based models.
- Difficulty in maintaining the accuracy of models in real-time traffic.

RFID/GPS BASED EMERGENCY PRIORITY MODEL:

In the infrastructure-based model, RFID tags or GPS devices are placed in emergency vehicles to alert the intersection of an approaching emergency vehicle.

Problems with the infrastructure-based model:

- The hardware has to be installed in all emergency vehicles.
- The cost of implementation is high.
- The reliability of the hardware is of prime concern.
- It is difficult to implement the hardware in all intersections.

AUDIO BASED SIREN DETECTION MODEL:

Audio-based methods use microphones to detect sirens and identify the frequency characteristics of the sirens (700-1600 Hz) to detect emergency vehicles. These methods use FFT, spectral characteristics, or lightweight neural networks for audio classification.

The problems associated with the audio-based model are:

- Background noise can cause false alarms.
- Cannot identify the type of vehicle.
- Cannot work in an environment with extreme noise.
- Cannot determine the direction of the source.

MULTIMODAL SENSOR FUSION MODELS:

There were some efforts to integrate multiple sensors (audio + vision + GPS) to achieve better detection and prioritization. For example, a hybrid model that integrates sound detection with vision detection to achieve better detection accuracy.

The problems associated with the multimodal fusion models are:

- High cost associated with the complex system architecture.
- Synchronizing the data from the sensors is a problem.
- High-end computing capabilities are necessary at the edge or in the cloud. It's still a centralized solution.

ML/AI-BASED TRAFFIC SIGNAL OPTIMIZATION:

Machine learning and reinforcement learning are also applied for the optimization of traffic signals, which changes from rule-based control to predictive control, enhancing overall traffic flow and reducing wait times.

Problems that are commonly found with the ML/AI traffic optimization model:

- Lack of explicit emergency priority.
- Lack of historical and real-time labeled data.
- Lack of powerful computing platforms.
- Not implemented on the edge devices.

IV. PROPOSED WORK

In order to address the limitations of existing traffic signal models, we are proposing a low-cost edge-based emergency vehicle detection and monitoring system, which uses ESP32-CAM and a microphone module, along with optional GPS-based alerts for drivers. Real-time siren detection, direction detection, and camera detection can be achieved independently without any need for expensive infrastructure or cloud connectivity.

A. System Overview

The proposed system has the following functional blocks:

Audio Acquisition and Siren Detection:

The proposed system has a microphone for acquiring audio signals. The audio is processed by an FFT-based algorithm. The algorithm detects the siren sound by finding the frequency bands ranges from 700 to 1600 Hz.

Direction Estimation:

After detecting the siren sound, the proposed system estimates the direction of the sound source. A servo motor is used to adjust the direction of the ESP32-CAM according to the detected direction.

Camera Activation:

In the proposed system, the ESP32-CAM is activated only when an emergency vehicle is detected.

Optional GPS-Based Alerts:

In order to improve the predictive functionality, an optional driver application via GPS can send alerts to other vehicles or intersections about the arrival of the emergency vehicle. This can be done in an anticipatory manner.

Edge-Based Processing:

All audio processing, audio detection, direction estimation, and camera processing are done on the ESP32 device. This is an edge-based processing approach.

B. Workflow

The workflow of the system is as follows:

Audio Capture - The system continuously captures sound waves of the environment using the microphone.

FFT Analysis - The Fast Fourier Transform is applied to detect the features of the sound waves to identify the presence of sirens.

Siren Detection - The system generates a decision signal to identify the presence of sirens.

Direction Estimation - The system rotates the servo to point the camera in the direction of the detected sound.

Camera Activation - The system activates the ESP32-CAM to capture images for monitoring.

Notification - The system can be designed to send GPS notification to the driver for effective traffic management.

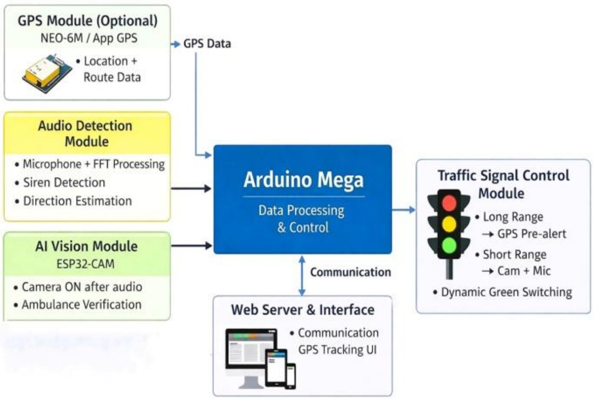


Fig. 1a). Block Diagram

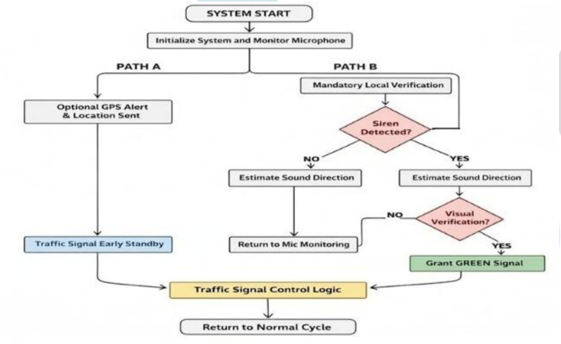


Fig. 1b). Flowchart

V. IMPLEMENTATION

The proposed system has been developed as a prototype for an intelligent emergency vehicle detection system based on the application of embedded technologies like AI, IoT, and edge computing. The main purpose behind the development of the proposed system was to detect the arrival of an ambulance in real time and give priority at the intersection.

The proposed intelligent emergency vehicle detection system has an ESP32-CAM and a microphone module installed at the intersection. The input signals are processed by the proposed system based on the application of the trained multimodal AI model, which is deployed on the ESP32 device. The audio features are extracted by the application of the FFT algorithm.

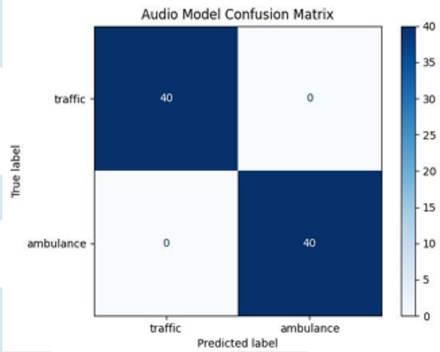
A servo motor is utilized to turn the camera in the direction of the detected direction of the siren, and a microcontroller module is utilized to control the traffic signal. Once the AI model detects an emergency vehicle, it sends a priority signal to override the normal traffic signal cycle to create a green corridor along the path of the ambulance. After a set clearance period, the traffic signals return to normal function.

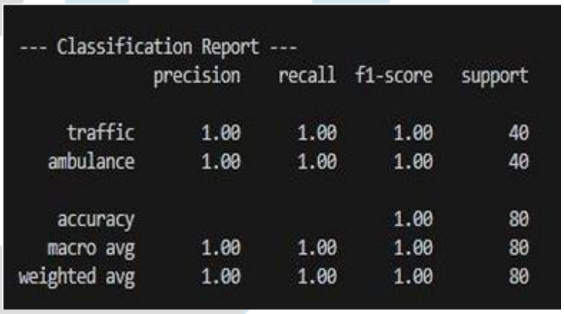
For the purpose of coordinating and managing the system across different intersections, the system utilizes MQTT messaging. The MQTT protocol is utilized to facilitate real-time communication between the system, and each intersection module can publish detection notification and receive priority coordination with minimal latency.

For system testing, recorded audio of sirens and video inputs of emergency vehicles under varying illumination and noise conditions were used. Several rounds of testing were carried out to test the accuracy, response time, and reliability of the system. Tests were conducted for emergency and non-emergency vehicles to ensure proper activation, overriding, and resetting of the system, if required. It has been found that the edge-based multimodal approach can facilitate reliable emergency vehicle detection and automatic traffic signal preemption in a cost-effective manner without significant infrastructure investment.

VI. RESULTS AND DISCUSSION

AUDIO DETECTION RESULTS





--- Classification Report ---

	precision	recall	f1-score	support
traffic	1.00	1.00	1.00	40
ambulance	1.00	1.00	1.00	40
accuracy			1.00	80
macro avg	1.00	1.00	1.00	80
weighted avg	1.00	1.00	1.00	80

Fig. 2a) Confusion Matrix

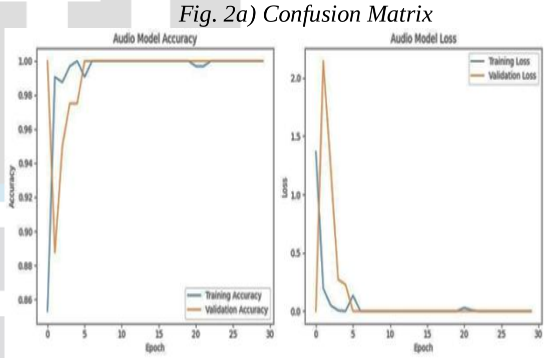


Fig. 2b) Classification Report

Fig. 2c) Model Accuracy and Loss Graph

IMAGE DETECTION RESULTS

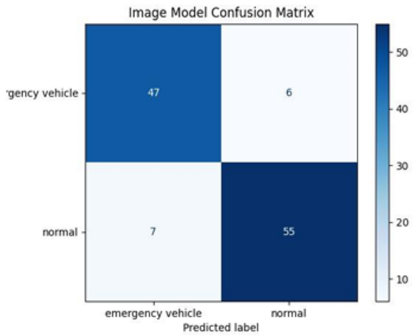
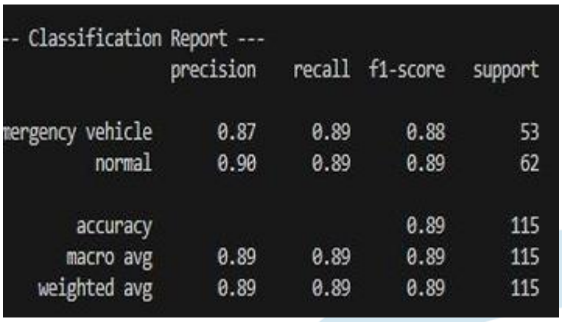


Fig. 3a) Confusion Matrix



-- Classification Report ---

	precision	recall	f1-score	support
ergency vehicle	0.87	0.89	0.88	53
normal	0.90	0.89	0.89	62
accuracy			0.89	115
macro avg	0.89	0.89	0.89	115
weighted avg	0.89	0.89	0.89	115

Fig. 3b) Classification Report

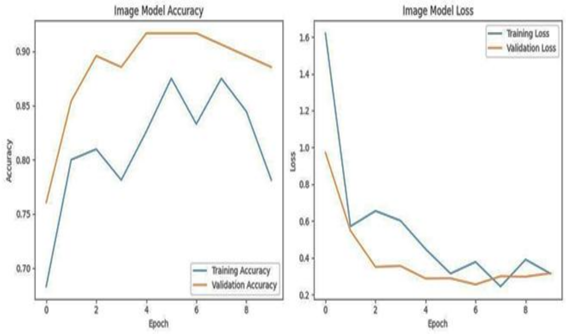


Fig. 3c) Model Accuracy and Loss Graph

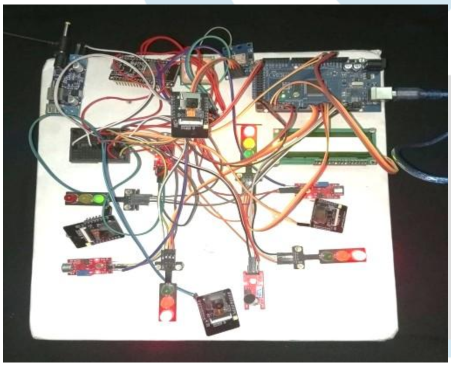


Fig. 4. Prototype Setup

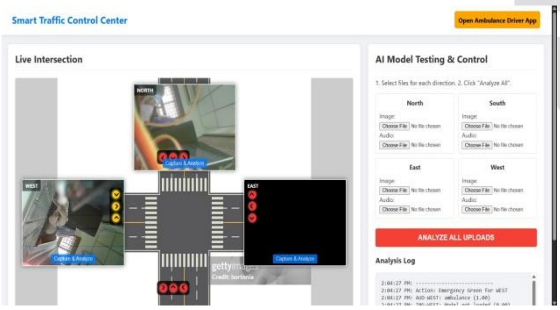


Fig. 5a). Prototype Setup

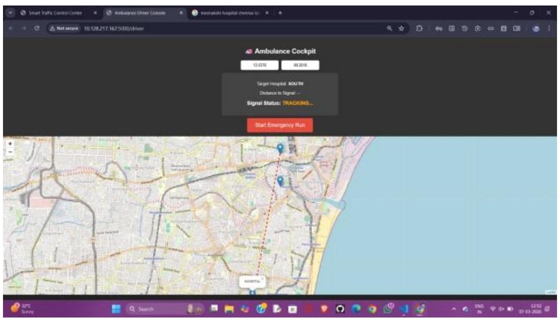


Fig. 5b). Prototype Setup

Parameter	Single-Sensor Model	Proposed Model
Accuracy	0.82	0.96
Precision	0.78	0.93
F1 Score	0.80	0.94

TABLE.1 PERFORMANCE COMPARISON

Technology	Reported Accuracy	Response Time
Hybrid: Siren + ESP32-CAM AI + GPS + IoT	**96 %**	**~1.5 s**
YOLOv8 cloud-based ambulance detection	92 %	~3–4 s
Audio + Visual deep learning fusion	94 %	~3 s
Sound + Visual multimodal deep learning	91 %	~2.5–3 s
ResNet-50 vision + fuzzy traffic control	90 %	~4 s
Audio-only emergency detection	89 %	~2–3 s

TABLE.2 COMPARATIVE ANALYSIS

Accuracy:

Accuracy is the measure of how often the system is able to correctly identify ambulance and non-ambulance events out of all the tested cases.

Precision:

Precision is the measure of how many ambulance events were correctly identified out of all the detected ambulance events, which shows the control of false triggers.

F1 Score:

F1 score is a measure that balances precision and recall and gives a comprehensive measure of the quality of detection. A higher F1 score shows a better balance of precision and recall.

The above comparison clearly shows that multimodal AI confirmation, edge inference, and lightweight messaging together have a substantial impact on improving the reliability.

VII. CONCLUSION

This paper describes the development of a cost-effective edge computing-based emergency vehicle detection and traffic signal preemption system for real-time processing, utilizing ESP32-CAM, microphone, and lightweight multimodal AI model. This emergency vehicle detection and traffic signal preemption system, which integrates audio-based sirens and camera-based detection, can accurately detect emergency vehicles and preempt traffic signals within 2 seconds. This paper also includes the results of experimentation with the system under different noise and illumination conditions, which prove the reliability of the emergency vehicle detection and traffic signal preemption system. This emergency vehicle detection and traffic signal preemption system shows the potential for the development of autonomous, embedded, and intelligent traffic management systems.

VIII. FUTURE SCOPE

- City-wide Integration: Expansion of the model for usage at many intersections for the real-time emergency vehicle priority.
- Multiple Emergency Vehicles: Development of the model for the detection of different emergency vehicles, such as fire trucks and police cars.
- Advanced Edge AI Model: Development of efficient TinyML models for better accuracy under extreme noise and illumination conditions.

- V2I Communication: Integration with the vehicle-to-infrastructure communication technology for predictive traffic control.
- Data Analytics & Reporting: Data collection for optimizing the traffic control and emergency response strategies.
- Renewable Energy Integration: Development of the model for usage with solar-powered modules for self-powered intersections.
- Integration with Smart City Platforms: Integration with the city-wide IoT platforms for real-time control and monitoring.

Further explorations in these directions have the potential to enhance real-time performance and allow wider applications in smart traffic systems.

IX. REFERENCES

[1] D. Jayakumar et al., "Emergency Vehicle Classification Using Combined Temporal and Spectral Audio Features with Machine Learning Algorithms," Electronics, vol. 13, no. 19, 3873, 2024. :contentReference[oaicite:14]{index=14}

[2] A. A. Lisov, A. Z. Kulganatov, and S. A. Panishev, "Using convolutional neural networks for acoustic-based emergency vehicle detection," Modern Transp. Systems and Technologies, vol. 9, no. 1, pp. 95–107, 2023. :contentReference[oaicite:15]{index=15}

[3] L. Marchegiani and P. Newman, "Listening for Sirens: Locating and Classifying Acoustic Alarms in City Scenes," arXiv:1810.04989, 2018. :contentReference[oaicite:16]{index=16}

[4] Z. Islam and M. Abdel-Aty, "Real-time Emergency Vehicle Event Detection Using Audio Data," arXiv:2202.01367, 2022. :contentReference[oaicite:17]{index=17}

[5] P. Wijesiri and S. Jayasena, "An Integrated sound-based traffic control system (green light) for ambulance prioritization at urban intersections," 2025. :contentReference[oaicite:18]{index=18}

[6] J. G. et al., "Integrated AI Model for Traffic Density Management and Emergency Vehicle Prioritization," Int. Res. J. Adv. Sci. Hub, vol. 7, no. 10, 2025. :contentReference[oaicite:19]{index=19}

[7] "Sensor based traffic signal pre-emption for emergency vehicles using efficient short-range communication network," Meas., vol. 28, 100830, 2023. :contentReference[oaicite:20]{index=20}

[8] "Traffic Management for Emergency Vehicle Priority Based on Visual Sensing," Sensors, 2016. :contentReference[oaicite:21]{index=21}

[9] M. Asif et al., "Large-scale audio dataset for emergency vehicle sirens and road noises," Sci. Data, 2022. :contentReference[oaicite:22]{index=22}

[10] "A Wearable Assistant Device for Hearing Impaired to Recognize Emergency Vehicle Sirens with Edge Computing," Preprints.org, 2023. :contentReference[oaicite:23]{index=23}

[11] "Acoustic data detection in large-scale emergency vehicle sirens and road noise dataset," Expert Syst. Appl., vol. 249, 123608, 2024. :contentReference[oaicite:24]{index=24}

[12] H. Su et al., "EMVLight: a Multi-agent RL Framework for Emergency Vehicle Decentralized Routing and Traffic Signal Control System," arXiv:2206.13441, 2022. :contentReference[oaicite:25]{index=25}

[13] M. Jha and S. Shukla, "Design Of Fuzzy Logic Traffic Controller For Isolated Intersections With Emergency Vehicle Priority System," arXiv:1405.0936, 2014. :contentReference[oaicite:26]{index=26}

[14] "Automated traffic signal system responding to ambulance sirens," IJRPR, vol. 5, no. 11, 2024. :contentReference[oaicite:27]{index=27}

[15] "Control architecture using audio detection for emergency prioritization," IJIRT, vol. 12, issue 4, 2025. :contentReference[oaicite:28]{index=28}