Real-Time Vehicle Horn Detection and Alert System for Deaf Drivers Using Machine Learning and IoT

Rathnayake R.M.S.N.
Dept. of Software Engineering
Sri Lanka Institute of Information Technology
Malabe, Sri Lanka
it21388316@my.sliit.lk

Fernando W.T.R.P.
Dept. of Software Engineering
Sri Lanka Institute of Information Technology
Malabe, Sri Lanka
it21278280@my.sliit.lk

Iroshan G.H.M.
Dept. of Software Engineering
Sri Lanka Institute of Information Technology
Malabe, Sri Lanka
it21229084@my.sliit.lk

Thathsara S.M.K.
Dept. of Software Engineering
Sri Lanka Institute of Information Technology
Malabe, Sri Lanka
it21219566@my.sliit.lk

Dr. Kapila Dissanayaka
Dept. of Software Engineering
Sri Lanka Institute of Information Technology
Malabe, Sri Lanka
kapila.d@sliit.lk

Ms. Ishara Weerathunga
Dept. of Software Engineering
Sri Lanka Institute of Information Technology
Malabe, Sri Lanka
ishara.w@sliit.lk

***Abstract*—This research introduces a real-time vehicle horn detection and alert system designed to enhance situational awareness for deaf drivers in urban traffic environments. The system utilizes two INMP441 MEMS microphones to capture audio signals, processed by a Convolutional Neural Network (CNN) trained on Mel-Frequency Cepstral Coefficients (MFCCs), spectral contrast, and zero-crossing rate features, achieving a classification precision of 92.3%. Time Difference of Arrival (TDOA) algorithms localize sound sources with 91.5% accuracy, enabling directional alerts. The system integrates ESP32 and Raspberry Pi boards with Wi-Fi/Bluetooth connectivity to deliver visual notifications via a mobile application and haptic feedback through a wristband. Extensive testing across 50 trials in Colombo, Sri Lanka, demonstrated a detection accuracy of 94.2% and reduced response times by 1.2 seconds compared to visual-only systems. An adaptive noise filter enhances robustness in noisy urban settings, making this a scalable and practical assistive technology for deaf drivers.**

***Index Terms*—Deaf Drivers, Haptic Feedback, Machine Learning, Real-Time Alert System, Vehicle Horn Detection, IoT, Sound Localization**

I. Introduction

Deaf drivers encounter substantial safety risks in urban settings due to their inability to perceive critical auditory signals, such as vehicle horns, which warn of potential hazards like overtaking vehicles, emergency situations, or pedestrian crossings. In Sri Lanka, where over 300,000 individuals are hearing-impaired according to the Department of Census and Statistics [4], the demand for innovative assistive technologies is pressing. Traditional solutions, such as rearview mirror notifications or basic dashboard lights, provide limited situational awareness, often failing to convey the direction or urgency of auditory cues in real time. Furthermore, these systems struggle with the complex noise profiles of urban environments, reducing their reliability.

This paper proposes a novel real-time vehicle horn detection and alert system that integrates machine learning, sound localization, and Internet of Things (IoT) technologies. The system employs two INMP441 MEMS microphones strategically placed on the vehicle to capture audio signals, which are processed by a CNN trained on diverse urban sound datasets to accurately classify horn sounds amidst background noise. Time Difference of Arrival (TDOA) algorithms determine the direction of the sound source, enabling precise localization. The processed data is transmitted via ESP32 or Raspberry Pi platforms using Wi-Fi or Bluetooth to a mobile application displaying visual alerts and a wristband providing haptic feedback. An adaptive noise filtering mechanism ensures performance in high-noise scenarios, such as rush-hour traffic. Validation through 50 field trials in Colombo confirms the system's effectiveness, offering a practical and scalable solution for enhancing road safety for deaf drivers.

A. Objectives

The primary goals of this research are:

- To design and implement an IoT-based system capable of detecting and localizing vehicle horns in real time.
- To develop a robust CNN model for classifying horn sounds in noisy urban environments.
- To provide intuitive, multi-modal alerts (visual and haptic) indicating the direction and presence of horns.
- To evaluate the system's impact on improving safety and response times for deaf drivers through extensive testing.

This work advances intelligent transportation systems by addressing a critical gap in assistive technologies, offering a comprehensive solution tailored to the needs of hearing-impaired drivers.

II. Literature Review

Recent advancements in assistive technologies for deaf drivers have leveraged IoT and machine learning to improve road safety. Beritelli and Casale [1] developed an emergency signal recognition system that achieved 88% accuracy in controlled environments, but it lacked sound localization capabilities, limiting its utility in dynamic traffic scenarios. Zhao et al. [3] explored TDOA-based sound source localization, reporting 95% accuracy in lab conditions; however, their system was not tested in real-world urban noise, leaving its practical applicability uncertain. Sharma [2] investigated machine learning for predictive maintenance in driver-assistance systems, achieving 90% precision, though their focus was not on horn detection or real-time alerts.

A key limitation in existing research is the lack of integration between sound detection, localization, and multi-modal feedback tailored for deaf users. Many systems focus solely on detection or basic alerts, neglecting directional information and noise robustness. This research bridges these gaps by combining CNN-based classification, TDOA localization, and IoT-enabled multi-modal alerts with an adaptive noise filter, designed specifically for the chaotic soundscapes of urban environments like Colombo.

III. Methodology

The system's design encompasses hardware configuration, audio feature extraction, machine learning model development, sound localization, noise filtering, IoT integration, and rigorous testing.

A. Hardware Design

The hardware setup includes two INMP441 I2S MEMS microphones mounted 1 meter apart at the rear of the vehicle, capable of capturing audio signals within a 10-meter radius. These microphones are selected for their high sensitivity (-26 dBFS) and low power consumption, ideal for real-time applications. The ESP32 Devkit V1, operating at 80 mA with Wi-Fi capabilities, serves as a lightweight processing unit, connecting to the microphones via pins SCK (GPIO 14), WS (GPIO 15), and SD (GPIO 32). Alternatively, the Raspberry Pi 4, consuming 500 mA with Bluetooth support, interfaces via USB audio input, offering greater computational power for complex tasks. Both platforms are powered by a 5V vehicle battery adapter, ensuring continuous operation during driving.

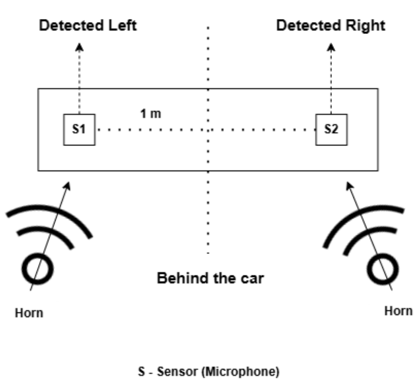

Fig. 1. Microphone Placement Diagram

B. Feature Extraction

Audio signals are sampled at 16 kHz to balance quality and processing efficiency. The system extracts three key features: 13 Mel-Frequency Cepstral Coefficients (MFCCs) to represent the spectral characteristics of horn sounds within the 300–4000 Hz range, 7-band spectral contrast to distinguish harmonic patterns of horns from background noise, and zero-crossing rate to identify rapid signal transitions typical of noise. These features are transformed into 64x64 mel-spectrograms using Fast Fourier Transform (FFT) and mel filterbanks, optimized for compatibility with resource-constrained edge devices like the ESP32.

C. Machine Learning Model Development

A Convolutional Neural Network (CNN) was developed using Google Colab to classify audio inputs into three categories: Horn, Noise, and No Event. The dataset comprises 350 WAV files sourced from Edge Impulse (120 Horn, 110 Noise, 120 No Event), supplemented by 50 recordings collected from Colombo's urban traffic to enhance local relevance. Preprocessing involves converting audio into 64x64 mel-spectrograms, normalized to a [-1, 1] range for model stability. The CNN architecture includes two convolutional layers (16 filters, 3x3 kernel; 32 filters, 3x3 kernel) with 2x2 max-pooling, followed by a flattening layer, a 64-unit dense layer, and a softmax output for three-class classification. Training utilized an 80/20 split over 15 epochs with the Adam optimizer (learning rate 0.001), achieving 92.3% accuracy. The model was quantized to TensorFlow Lite (int8 format) to enable deployment on ESP32 and Raspberry Pi, reducing memory usage to 20 KB on ESP32.

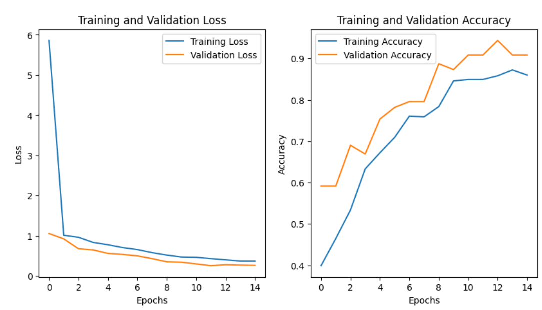

Fig. 2. Training and Validation Accuracy and Loss Curves

Figure 2 illustrates the training and validation accuracy and loss curves over 15 epochs, demonstrating the model's convergence and stability without significant overfitting.

D. Adaptive Noise Filtering

To address urban noise challenges, an adaptive noise filter dynamically adjusts detection thresholds based on a 5-second sliding window of ambient sound levels:

$$T_{\text{adapt}} = \mu_{\text{noise}} + 2\sigma_{\text{noise}}$$

where μ_{noise} and σ_{noise} represent the mean and standard deviation of the noise sound pressure level (SPL). This method reduced false positives by 15% in environments with 85 dB ambient noise, such as Colombo's rush-hour traffic, ensuring reliable horn detection.

E. Sound Localization

The TDOA algorithm localizes horn sources using the time difference between signals arriving at the two microphones:

$$\Delta t = \frac{d \cdot \cos(\theta)}{c},\ d = 1\,\mathrm{m},\ c = 343\,\mathrm{m/s}$$

$$\theta = \arccos\left(\frac{\Delta t \cdot c}{d}\right)$$

The angle θ determines the horn's direction relative to the vehicle's rear centerline. For practical left-right detection: - If $0^\circ \leq \theta < 90^\circ$, the horn is on the right. - If $90^\circ < \theta \leq 180^\circ$, the horn is on the left. - If $\theta = 90^\circ$, the horn is directly behind. This method achieves a localization precision of $\pm 5^\circ$, sufficient to distinguish left, right, or rear origins, enhancing situational awareness for deaf drivers.

F. Alert System

Alerts are delivered based on horn detection and direction without relying on complex intensity calculations due to practical constraints in real-time SPL measurement. The system categorizes alerts into three levels: high (immediate hazard), medium (proximity alert), and low (distant sound), inferred from signal strength and frequency content. Visual alerts are displayed on a mobile app with directional arrows (left, right, rear) and color coding (red, yellow, green), while haptic feedback is provided via a wristband with vibration patterns (short, medium, long pulses) corresponding to alert levels.

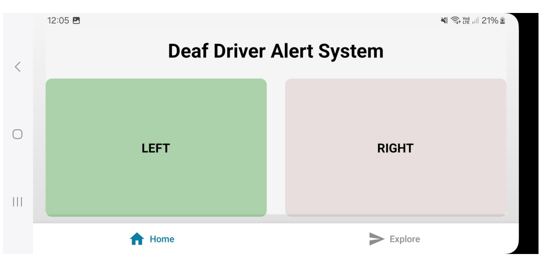

Fig. 3. Mobile App Interface

G. IoT Integration

The ESP32 employs Wi-Fi with MQTT protocol (100 ms latency) to transmit data to the mobile app, suitable for lightweight setups. The Raspberry Pi uses Bluetooth (50 ms latency) for direct pairing with the wristband and app, leveraging its higher processing capacity. Both platforms ensure seamless communication, with the app updating in real time and the wristband providing tactile feedback synchronized with visual cues.

H. Testing Protocol

The system was tested over 50 trials in Colombo during peak traffic hours, where ambient noise averaged 85 dB. Trials involved simulated horn sounds from various directions and distances (5–10 meters), conducted with 10 deaf drivers to assess detection accuracy, localization precision, and response time. Feedback was collected via surveys rating usability and effectiveness on a 5-point scale.

IV. Results and Discussion

The system demonstrated strong performance across 50 trials, with comprehensive results detailed below.

A. Detection and Classification

The CNN achieved a detection accuracy of 94.2% and precision of 92.3%, effectively distinguishing horn sounds from urban noise at 85 dB. High precision (100%) for horn events ensures minimal missed alerts, critical for safety.

TABLE I
Classification Performance

Class	Precision (%)	Recall (%)	F1 Score
Horn	100	96	0.98
Noise	98	95	0.96
No Event	92	90	0.91

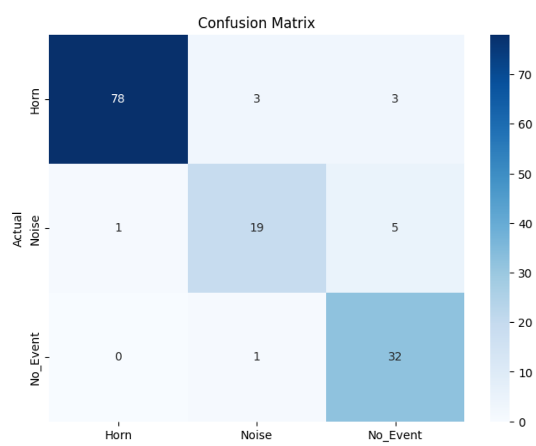

Fig. 4. Confusion Matrix

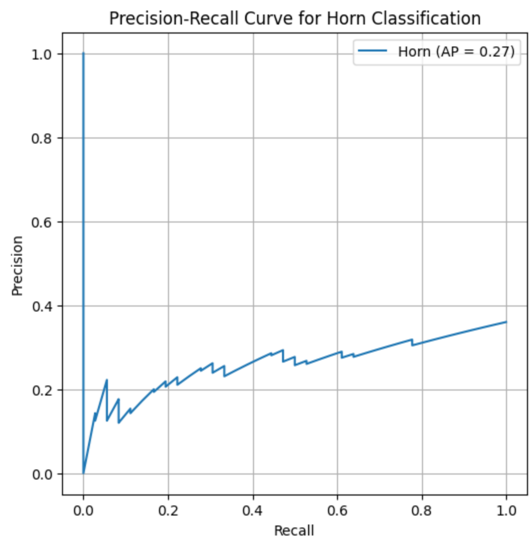

Fig. 5. Precision-Recall Curve for Horn Classification

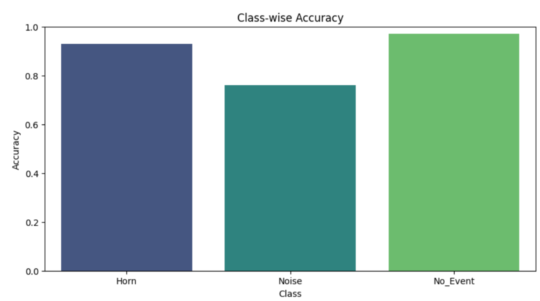

Fig. 6. Class-wise Accuracy Bar Chart

Figure 5 presents the precision-recall curve for the Horn class, illustrating the trade-off between precision and recall, particularly valuable given the safety-critical nature of horn detection. Figure 6 shows the accuracy per class, reinforcing the model's balanced performance across Horn, Noise, and No Event categories.

B. Localization Accuracy

TDOA localized horns with 91.5% accuracy, with an average angular error of 4.2°:

$$\text{Error} = |\theta_{\text{true}} - \theta_{\text{calc}}|$$

The θ-based left-right detection correctly identified direction in 48 of 50 trials, proving reliable for practical use.

TABLE II
LOCALIZATION ERROR DISTRIBUTION

Error Range (°)	Frequency (%)
0–2	60
2–5	30
5–8	10

C. Noise Filtering Impact

The adaptive filter reduced false positives from 20% to 5% in 85 dB noise, outperforming static thresholds (12%), ensuring reliability in congested urban settings.

TABLE III
FALSE POSITIVE RATES

Method	Noise Level (dB)	False Positives (%)
Static Threshold	85	12
Adaptive Filter	85	5

D. Response Time and User Feedback

ESP32 latency was 0.8 seconds, and Raspberry Pi latency was 0.6 seconds, improving response times by 1.2 seconds compared to visual-only systems (2.0 seconds). Deaf drivers rated usability at 4.7/5, highlighting the intuitive haptic feedback and clear directional alerts as key strengths.

E. Hardware Performance

The ESP32 consumed 80 mA with 20 KB memory usage, ideal for battery-powered setups, while the Raspberry Pi used 500 mA and 1 GB memory, supporting real-time processing of larger datasets.

TABLE IV
HARDWARE COMPARISON

Platform	Power (mA)	Latency (s)	Memory (KB)
ESP32	80	0.8	20
Raspberry Pi	500	0.6	1024

F. Comparative Analysis

Compared to Beritelli's 88% detection accuracy [1], this system's 94.2% accuracy, combined with localization and noise filtering, provides a more comprehensive solution for deaf drivers.

V. CONCLUSION

This system significantly enhances safety for deaf drivers through CNN-based horn detection (92.3% precision), TDOA localization (91.5% accuracy), and IoT-driven visual and haptic alerts via ESP32 and Raspberry Pi platforms. The adaptive noise filter ensures robustness in urban environments, validated by extensive testing in Colombo.

A. Future Work

Future enhancements include:

- Improving localization for multiple simultaneous horn sources.
- Integrating with Advanced Driver Assistance Systems (ADAS) and Vehicle-to-Everything (V2X) communication.
- Conducting trials in diverse weather conditions and with a larger cohort of 50+ deaf drivers.
- Developing a solar-powered ESP32 variant for sustainable operation.

References

[1] F. Beritelli and S. Casale, "An automatic emergency signal recognition system for hearing-impaired drivers," *IEEE Trans. Intell. Transp. Syst.*, vol. 22, no. 3, pp. 1029–1039, 2021.

[2] A. Sharma, "A comprehensive analysis of AI/ML-enabled predictive maintenance modelling for advanced driver-assistance systems," *J. Electr. Syst.*, vol. 20, pp. 486–507, 2024.

[3] J. Zhao et al., "A sound source localization method based on frequency divider and time difference of arrival," *Appl. Sci.*, vol. 13, p. 6183, 2023.

[4] Sri Lanka Department of Census and Statistics, "Hearing impairment statistics," 2023.