IJARETY

International Journal of Advanced Research in Education and TechnologY (IJARETY)

Volume 12, Issue 3, May-June 2025

Impact Factor: 8.152

www.ijarety.in editor.ijarety@gmail.com

International Journal of Advanced Research in Education and TechnologY(IJARETY)

| ISSN: 2394-2975 | www.ijarety.in| | Impact Factor: 8.152| A Bi-Monthly, Double-Blind Peer Reviewed & Refereed Journal |

|| Volume 12, Issue 3, May-June 2025 ||

DOI:10.15680/IJARETY.2025.1203047

Real-Time Audio Event Detection for Threat Identification using 1D-CNN

Pathri Dilip

Department of CSE, Guru Nanak Institutions Technical Campus, Hyderabad, India

M. Vamshi Krishna

Department of CSE, Guru Nanak Institutions Technical Campus, Hyderabad, India

P.V Araprakash

Department of CSE, Guru Nanak Institutions Technical Campus, Hyderabad, India

N. Mahipal Yadav

Department of CSE, Guru Nanak Institutions Technical Campus, Hyderabad, India

Ms. A. Vishalakshi

Department of CSE, Guru Nanak Institutions Technical Campus, Hyderabad, India

ABSTRACT: Today, humans pose the greatest threat to society by getting involved in robbery, assault, or homicide activities. Such circumstances threaten the people working alone at night in remote areas especially women. Any such kind of threat in real time is always associated with a sound/noise which may be used for an early detection. Numerous existing measures are available but none of them sounds efficient due to lack of accuracy, delays in exact prediction of threat. Hence a novel software-based prototype is developed to detect threats from a person's surrounding sound/noise and automatically alert the registered contacts of victims by sending email, SMS, WhatsApp messages through their smartphones without any other hardware components. Audio signals from Kaggle dataset are visualized, analyzed using Exploratory Data Analytics (EDA) techniques. By feeding EDA outcomes into various Deep Learning models: Long short-term memory (LSTM), Convolutional Neural Networks (CNN) yields accuracy of 96.6% in classifying the audio-events.

I. INTRODUCTION

1. GENERAL

In the physical world, the occurrence of any physical event would bear with it a sound, particular to that event only. Be it the sound of a stone falling to the ground, a river flowing, a bird chirping, a person walking, a road being constructed etc. It is thus fair to take into account, just like an event with which positivity can be associated (crowds cheering, friends greeting each other, celebratory fireworks etc.) has a sound paired with it, negative events (a road accident, a landslide, a gunshot etc.) also have specific sounds associated with them as well. It is worth pondering how great would be a system which would be able to detect ambient noise and judge whether itis related to something positive or something negative. Even if direct applications as described here are not available to mankind's everyday use, the technology which is required to make it a reality already exists in one The associate editor coordinating the review of this manuscript and approving it for publication was Felix Albu form or the other. Before directly discussing the prospects of natural language processing in the domain of security, one needs to understand the current scenario, and where and how sound/text detection/analysis is done by NLP and used by mankind in modern times.

Modern smartphones equipped with Artificial Intelligence based voice recognition systems are classic examples of the application of speech-based natural language processing in our day-to-day lives. With the likes of Google's Voice Assistant available on every android powered smartphone which searches the web and brings the necessary information in front of us without us typing a single search keyword, or the evolution of Apple's Siri through the various versions of the iconic iPhone series which primarily works as a personal assistant for its users, or the availability of Amazon's Alexa as a smart home voice assistant which can be used to control virtually any smart electrical device in our home, or the development of Samsung's own artificial intelligence-based personal assistant named Bixby natural language

| ISSN: 2394-2975 | www.ijarety.in| | Impact Factor: 8.152| A Bi-Monthly, Double-Blind Peer Reviewed & Refereed Journal |

|| Volume 12, Issue 3, May-June 2025 ||

DOI:10.15680/IJARETY.2025.1203047

processing has already come a long way in the domain of voice/text-based language processing for the betterment of human lives, all while keeping at par with the compatibility with the latest technologies and hardware

SCOPE OF THE PROJECT

The developed system has demonstrated strong potential for real-world applications, but there is ample room for enhancement and expansion. One major area for future work involves increasing the size and diversity of the audio dataset to include more sound classes, which would make the model more robust in identifying a wider range of threat scenarios. Additionally, integrating advanced deep learning architectures such as GRU, Bi-GRU, Bi-LSTM, CTRNN, and HRNN could further improve prediction accuracy and system reliability. Feature extraction methods like amplitude, phase, and harmonic distortion analysis may also enhance model performance by providing deeper insights into the audio signals.

OBJECTIVE

The primary objective of this project is to develop a real-time audio event detection system using deep learning techniques, specifically 1D-CNN, to identify potentially dangerous or threatful environments based on surrounding sound data. The system aims to accurately classify audio signals into multiple categories and, upon detecting a threat, automatically send alert notifications to registered emergency contacts through email, SMS, and WhatsApp. This includes sharing the recorded audio of the incident, thereby enabling faster and more informed responses to emergencies. The broader goal is to enhance personal safety using AI-powered threat detection without the need for any additional hardware beyond a smartphone..

PROBLEM STATEMENT

In an increasingly unsafe world, individuals often find themselves in potentially dangerous situations where they are unable to call for help manually. There is a critical need for an intelligent, real-time system that can automatically detect threatful or harmful environments based on ambient sounds and initiate emergency responses without human intervention. This research addresses the challenge of identifying real-world audio events—such as gunshots, screams, explosions, or breaking glass—using a multi-class classification approach with deep learning models. The objective is to classify these sounds accurately and trigger an automated alert system that immediately informs registered emergency contacts via email, SMS, and WhatsApp, including the audio recording as evidence. The goal is to enable proactive safety measures through efficient, intelligent sound-based threat detection

III. EXISTING SYSTEM

- Most existing emergency alert systems rely on manual input from users (like dialing emergency numbers or pressing a panic button), which may not be possible in critical situations where the victim is incapacitated.
- Some safety systems require specialized wearable devices or hardware (like smartwatches or fitness bands) to monitor distress signals, making them costly and less accessible for widespread use.
- Current systems that use sound recognition are typically limited in detecting a narrow range of sounds and lack advanced real-time classification using deep learning techniques
- Traditional systems may suffer from delays in processing and alerting emergency contacts, as they often require intermediate human validation or lack real-time automation.

EXISTING SYSTEM DISADVANTAGES

- Existing systems often require manual intervention, which delays emergency response and is ineffective if the victim cannot operate the device
- Many systems cannot accurately classify or detect a wide variety of real-world threat-related sounds like gunshots, screams, or explosions.
- These systems usually rely on additional hardware (e.g., smartwatches or specialized sensors), increasing costs and reducing portability.

III. LITERATURE SURVEY

Several researchers have explored the domain of audio-based threat detection and classification using machine learning and deep learning techniques. In [1], the authors utilized Mel-frequency cepstral coefficients (MFCC) and support vector machines (SVM) for environmental sound classification. While the method provided decent accuracy, it lacked real-time responsiveness and scalability to multiple sound classes. Another study in [2] applied 2D Convolutional

| ISSN: 2394-2975 | www.ijarety.in| | Impact Factor: 8.152| A Bi-Monthly, Double-Blind Peer Reviewed & Refereed Journal |

|| Volume 12, Issue 3, May-June 2025 ||

DOI:10.15680/IJARETY.2025.1203047

Neural Networks (CNNs) on spectrogram images for urban sound classification, demonstrating improved performance over traditional classifiers. However, this model required significant computational resources and struggled with overlapping audio events.

In [3], a hybrid approach involving Long Short-Term Memory (LSTM) networks was used for sequential audio data analysis, which yielded promising results in detecting anomalies in audio streams. Yet, the system did not integrate real-time alert mechanisms for emergency scenarios. Meanwhile, [4] implemented an IoT-based wearable solution for women's safety using sound and location tracking but faced challenges with data accuracy and hardware dependencies. These studies collectively highlight the need for a lightweight, real-time, and highly accurate threat detection system that can function without reliance on external hardware and provide instant alerts through multiple communication channels.

PROPOSED SYSTEM

- Utilizes 1D-CNN deep learning model for real-time audio event detection.
- Automatically classifies environmental sounds into predefined threat categories.
- Sends immediate alerts via email, SMS, and WhatsApp with recorded audio attachment.
- Operates solely through smartphones without the need for external hardware.

PROPOSED SYSTEM ADVANTAGES

The system can identify potentially dangerous sounds instantly, enabling faster response and increased personal safety. It functions using standard smartphone hardware, eliminating the need for additional wearable or specialized devices. Automatically sends alerts through email, SMS, and WhatsApp, ensuring the emergency contacts are quickly informed through multiple communication channels.

IV. APPLICATION

GENERAL

The proposed system uses deep learning models to detect real-time audio events that indicate potential threats. It classifies sounds like screams, gunshots, or glass breaking using models such as 1D-CNN, 2D-CNN, and LSTM. Upon detection, it sends immediate alerts to registered contacts through email, SMS, and WhatsApp.

FUTURE ENHANCEMENT

The system can be improved by integrating GPS functionality to share the real-time location of the user along with threat alerts, enabling quicker emergency response. Incorporating additional biometric sensors to monitor vital signs such as heart rate, oxygen levels, and body temperature would provide a more comprehensive picture of the user's condition during emergencies. Expanding the audio dataset to include more diverse and complex sound classes will enhance the model's ability to detect a wider range of threats. Advanced deep learning architectures like Bi-GRU, Bi-LSTM, or Hierarchical RNNs can be explored to improve classification accuracy and handle longer sequences of audio data. Implementing noise reduction and audio enhancement techniques will increase the robustness of the system in noisy environments. Integrating the system with wearable devices or IoT gadgets can provide a multi-modal safety solution. Future work can also focus on reducing computational latency to ensure even faster real-time detection. The alert system could be expanded to include integration with emergency services and law enforcement for direct communication. Adding a user-friendly mobile app interface with customization options for alert preferences would improve usability. Finally, continuous learning mechanisms can be implemented so the system adapts and improves from new audio data over time

V. CONCLUSION

CONCLUSION

This project successfully demonstrates the use of deep learning, specifically 1D-CNN, for real-time audio event detection aimed at enhancing personal safety. The system effectively classifies various environmental sounds that could indicate potential threats, such as footsteps, gunshots, or glass breaking. By automatically sending alerts via multiple channels like email, SMS, and WhatsApp, it ensures timely communication to emergency contacts. The approach does not require additional hardware, making it accessible and convenient for users with just a smartphone. Experimental results show high accuracy in detecting and classifying audio events, proving the model's reliability. While the current system addresses critical safety needs, it leaves room for future enhancements such as location tracking and biometric

| ISSN: 2394-2975 | www.ijarety.in| | Impact Factor: 8.152| A Bi-Monthly, Double-Blind Peer Reviewed & Refereed Journal |

|| Volume 12, Issue 3, May-June 2025 ||

DOI:10.15680/IJARETY.2025.1203047

monitoring. The project paves the way for more advanced, integrated safety solutions using IoT and AI technologies. Overall, this research contributes to improving situational awareness and rapid response in potentially dangerous situations. The simplicity, efficiency, and real-time capabilities make it a practical solution for personal security. Continued development and integration with other systems could greatly expand its impact in public safety domains.

REFERENCES

[1] T. P. Suma and G. Rekha, ‘‘Study on IoT based women safety devices with screaming detection and video capturing,’’ *Int. J. Eng. Appl. Sci. Technol.*, vol. 6, no. 7,pp. 257–262, 2021.
[2] P. Zinemanas, M. Rocamora, M. Miron, F. Font, and X. Serra, ‘‘An interpretable deep learning model for automatic sound classification,’’ *Electronics*, vol. 10, no. 7, p. 850, Apr. 2021.
[3] D. G. Monisha, M. Monisha, G. Pavithra, and R. Subhashini, ‘‘Women safety device and application-FEMME,’’ *Indian J. Sci. Technol.*, vol. 9, no. 10, pp. 1–6,Mar. 2016.
[4] G. Ciaburro and G. Iannace, ‘‘Improving smart cities safety using sound events detection based on deep neural network algorithms,’’ *Informatics*, vol. 7, no. 3, p. 23,Jul. 2020.
[5] J. Cao, M. Cao, J. Wang, C. Yin, D. Wang, and P.-P. Vidal, ‘‘Urban noise recognition with convolutional neural network,’’ *Multimedia Tools Appl.*, vol. 78, no. 20,pp. 29021–29041, Oct. 2019.
[6] A. Triantafyllopoulos, G. Keren, J. Wagner, I. Steiner, and B. W. Schuller, ‘‘Towards robust speech emotion recognition using deep residual networks for speech enhancement,’’ in *Proc. INTERSPEECH*, Graz, Austria, Sep. 2019.
[7] M. T. García-Ordás, H. Alaiz-Moretón, J. A. Benítez-Andrades, I. García-Rodríguez, O. García-Olalla, and C. Benavides, ‘‘Sentiment analysis in non-fixed lengthaudios using a fully convolutional neural network,’’ *Biomed. Signal Process. Control*, vol. 69,Aug. 2021, Art. no. 102946.
[8] A. Bhardwaj, P. Khanna, S. Kumar, and Pragya, ‘‘Generative model for NLP applications based on component extraction,’’ *Proc. Comput. Sci.*, vol. 167, pp. 918–931,Jan. 2020.
[9] I. S. Malova and D. V. Tikhomirova, ‘‘Recognition of emotions in verbal messages based on neural networks,’’ *Proc. Comput. Sci.*, vol. 190, pp. 560–563, Jan. 2021.
[10] Kodi, D. (2024). Automating Software Engineering Workflows: Integrating Scripting and Coding in the Development Lifecycle . Journal of Computational Analysis and Applications (JoCAAA), 33(4), 635–652.
[11] A. Hodorog, I. Petri, and Y. Rezgui, ‘‘Machine learning and natural language processing of social media data for event detection in smart cities,’’ *Sustain. Cities Soc.*, vol. 85, Oct. 2022, Art. no. 104026.
[12] D. D. Nguyen, M. S. Dao, and T. V. T. Nguyen, ‘‘Natural language processing for social event classification,’’ in *Knowledge and Systems Engineering*. Cham, Switzerland: Springer, 2015, pp. 79–91.lves,” Nature, vol. 421, pp. 799–800 Feb.2003.

International Journal of Advanced Research in Education and Technology

ISSN: 2394-2975 Impact Factor: 8.152

www.ijarety.in editor.ijarety@gmail.com