Live Event Detection for People Safety Using NLP and Deep Learning

Miss.Jyothika K R [1] , Usha N [2]

[1]*Assistant Professor,Department of MCA,BIET,Davanagere*

[2]*Student, 4th Semester MCA,Department of MCA,BIET,Davanagere*

ABSTRACT

In today's world, human actions such as robbery, assault, and homicide pose significant threats to society, especially to individuals working alone at night in remote areas—women being particularly vulnerable. These threatening incidents are often accompanied by distinct sounds or noises, which can serve as early indicators of danger. Although several existing security measures are available, many fall short due to their limited accuracy and delayed threat detection. To address this, a novel software-based prototype has been developed that detects potential threats in real-time by analyzing ambient sounds or noises. Without the need for any additional hardware components, the system can automatically alert a victim's pre-registered contacts by sending notifications via email, SMS, and WhatsApp directly through their smartphone. Audio signals sourced from a Kaggle dataset were visualized and examined using Exploratory Data Analysis (EDA) techniques.

The insights obtained from EDA were then used to train deep learning models, including Long Short-Term Memory (LSTM) and Convolutional Neural Networks (CNN), achieving an impressive 96.6% accuracy in classifying various audio events.

Keywords: prototype

INTRODUCTION

In recent times, the rise in violent human activities such as robbery, assault, and homicide has become a growing concern, posing serious threats to public safety. Individuals working alone during late hours, particularly in isolated or remote areas, are especially vulnerable—women being the most affected. These dangerous situations are often accompanied by specific sounds or noises, which can serve as early indicators of a potential threat.While several safety measures and alert systems have been developed, many of them fall short in terms of real-time accuracy and rapid threat detection. This calls for a more efficient and intelligent solution. To address this gap, a novel software-based prototype has been proposed that leverages sound or noise from a person's surroundings to detect possible threats. Uniquely, this system functions without the need for additional hardware and automatically notifies the victim's registered emergency contacts through email, SMS, and WhatsApp, directly via their smartphone. To develop and train this system, audio signals from a Kaggle dataset were analyzed using Exploratory Data Analytics (EDA) techniques. The results were used to train advanced deep learning models such as Long Short-Term Memory (LSTM) and Convolutional Neural Networks (CNN), achieving a high classification accuracy of 96.6% in detecting and categorizing threatening audio events.

II. LITERATURE REVIEW

In recent years, security in urban areas has gradually assumed a central position, focusing increasing attention on citizens, institutions and political forces. Security problems have a different nature—to name a few, we can think of the problems deriving from citizens' mobility, then move on to microcrime, and end up with the ever-present risk of terrorism. Equipping a smart city with an infrastructure of sensors capable of alerting security managers about a possible risk becomes crucial for the safety of citizens. The use of unmanned aerial vehicles (UAVs) to manage citizens' needs is now widespread, to highlight the possible risks to public safety. These risks were then increased using these devices to carry out terrorist attacks in various places around the world. Detecting the presence of drones is not a simple procedure given the small size and the presence of only rotating parts. This study presents the results of studies carried out on the detection of the presence of UAVs in outdoor/indoor urban sound environments. For the detection of UAVs, sensors capable of measuring the sound emitted by UAVs and algorithms based on deep neural networks capable of identifying their spectral signature that were used. The results obtained suggest the adoption of this methodology for improving the safety of smart cities.

III. EXISTING SYSTEM

Live event detection for people safetyusing NLP and deep learning Authors: T.P. Suma and G. Rekha, Raspberry-Pi based IoT device that includes a camera, sound sensor, GPS (Global Positioning System), and GSM (Global System for Mobile Communications) module. When a scream is recorded by the sound sensor, the SVM algorithm is able to recognize it. The camera is then turned on, which records a 30-second video clip and sends it to the closest police station or emergency services.The GPS module continuously tracks the victim's location, which is then transmitted to the emergency services together with the camera footage using the GSM module.While the above method focus more on the application of IoT (Internet of Things) along with sound detection,[1]

Live event detection for people safety using NLP and deep learning Authors: P. Zinemanas, M. Rocamora, M. Miron, F. Font, and X. Serra, have based their experiments on the similarity of the input to a group of learned prototypes in a latent space and utilize a frequency-dependent similarity measure that is built by taking into account different time-frequency resolutions in the feature space. Voice, music, and background noise are three different sound categorization tasks that the proposed model is capable of handling. Here, a Deep Learning Model containing a Prototype Layer, a Similarity Measure and Weighted Sum Layer, and a Fully-Connected Layer is utilized to extract insightful information from the input sound. Although this research is not furthered towards the application of the designed system in the field of individual security, it holds a great potential for the same. Utilizing an ARM (Advanced RISC Machine) controller and an android application.[2]

Live event detection for people safety using NLP and deep learning Authors: G. Monisha, M. Monisha, G. Pavithra, and R. Subhashini synchronizes the device and the smartphone using Bluetooth so that each may be turned on independently. Every two minutes, the device may send alert calls and messages to the pre-set contacts along with its current location, record audio for further analysis, and be followed in real time through a mobile application. An additional distinctive element of the system that one might employ to protect their privacy is a hid-den camera detector. Its major advantage is that it can be used to protect women against crimes including stalking, domestic violence, physical assault, and intrusive hidden cameras. Focusing more on the domain of security in the urban scenario.[3]

Live event detection for people safety using NLP and deep learning Authors:G. Ciaburro and G. Iannace where the researchers have gone by the domain of security in a slightly unorthodox manner. By identifying Unmanned Aerial Vehicles (UAVs) in loud outdoor and interior contexts, which have recently been utilized to carry out or support terrorist operations, the proposed effort aims to increase urban safety. Deep neural network-based techniques that can identify a UAV's spectral signature are used to detect UAVs in addition to sensors that might measure the sound they made.[4]

Live event detection for people safety using NLP and deep learning Authors:J. Cao, M. Cao, J. Wang, C. Yin, D. Wang, and P.-P. Vidal a method for categorizing sounds, a thorough classification of the various noises present in an urban area is achievable. The result can be used to generate insights into the different kinds of activities going around in an area, and further be used to detect whether any individual/group of people is in any dangerous situation or not. The log-Mel spectrogram's FBank feature is first developed for auditory representation. A series of FBank feature vectors created from distinct acoustic signal frames are then used as input to a Convolutional Neural Network (CNN) for urban noise identification. Here, the traditional LPCC (Linear Prediction Cepstral Coefficients) and MFCC (Mel-Frequency Cepstral Coefficients) acoustic feature, the FBank image feature, the hierarchical extreme learning machine (H-ELM), and the multilayer extreme learning machine are integrated with the support vector machine (SVM) and the extreme learning machine (ML-ELM).[5]

Live event detection for people safety using NLP and deep learning Authors:A. Triantafyllopoulos, G. Keren, J. Wagner, I. Steiner, and B. W. Schuller For emotion recognition from speech, looked at how noise affected two popular SER (Speech Emotion Recognition) architectures, Acoustic Features and End-to-end, as well as the potential benefits of implementing speech enhancement in SER applications, particularly in low SNRs. This system's ability to recognize speech (a sort of noise) even at very low Signal-to-Noise Ratios (SNRs), or for poor input sound quality, is a significant benefit. In this research, a number of SER techniques based on SVMs and openSMILE features are employed. The approach is based on stacked residual blocks of 2D convolution layers, which have been shown to efficiently learn rich representations of input signals in the past.[6]

3.1 DISADVANTAGES OF EXISTING SYSTEM:

• Tt provides the necessary solution for the problem of detection of threat around an individual, but it comes with a bulky hardware, which poses a difficulty in carrying it around for regular use.

•The researchers use several techniques to analyze audio signals, but they don't further their work to provide a practical solution to the problem of danger detection around an individual.

• The research done somewhat close to what has been achieved in this research, where the researchers have built a system to provide real time feedback from a person's surroundings, however, this also comes with an additional hardware component in addition to a smartphone.

•The researchers use noise detection to detect Unmanned Aerial Vehicles (UAVs) which might be used for criminal activities. However, this approach is not favorable to be applied at an individual level, and would not be suitable for detection of threat around an individual human being.

•Used similar techniques to detect threat for an individual/group in an urban context, but is unable to provide a solution to make the user friendly at an individual level with the use of no hardware.

• The research focus on detection on the detection of emotion from speech, which can be helpful in determining whether a person is in agony or not, or if a person is being verbally threatened by another fellow human being or not. Systems like these, although beneficial, are unable to address the problem of physical safety of an individual.

• Some other systems, like those proposed detect emotions and events from texts/speech, social media posts and tweets respectively. Although these are unique approaches to determine an individual's/group's live situation, they again fail to address the challenge of physical individual safety.

•Use noise detection for perimeter defense techniques (like intrusion detection), and conversion of urban sounds into information respectively, but do not address how an individual can be helped with respect to physical threats.

IV. PROPOSED SYSTEM

The main objective of the proposed system is to detect and classify the victim's live audio signals for immediate rescue. The system is intended to deliver its excellence as an application in any smartphone and it uses the default microphone configuration of it. On detection of suspicious audio patterns from the live input audio from microphone, the geographical location of the victim is shared to the emergency contacts in the phone as well as to the police patrol. The drawbacks inferred from the current violence detection scenarios related to audio event detection and classification accuracy are addressed for effective functioning which plays a vital role in avoiding false event classifications meanwhile ensuring the victim's safety through high classification accuracy.

4.1 ADVANTAGES

A very distinct advantage of the proposed system is that it does not require any special external hardware/wearable devices (like apart watches), but can be implemented with the help of a mere smartphone configured accordingly. In most modern IoT (Internet of Things) systems, the system comes with associated hardware, like smartwatches, which constantly monitor a person's health parameters. A person would not have to carry bulky hardware around with them, but only remember to carry their phones with them. However, whenever, more and more modules would be tried to be integrated along with the existing system, the inclusion of a special hardware might become an issue. As crime rates, and the frequency of natural and man-made disasters have increased significantly in modern days, it is of prime essence to have emergency help and services at the victim's assistance as quickly as possible, with immediate and accurate information about the victim's situation shared with them automatically. The yielded solution helps greatly to avoid any such type of hazard like a robbery, homicide or other threats to any individual as it helps greatly in the accurate prediction of surrounding sound/noise of victims in minimum time and alerts the emergency contact list of victims.

System Architecture

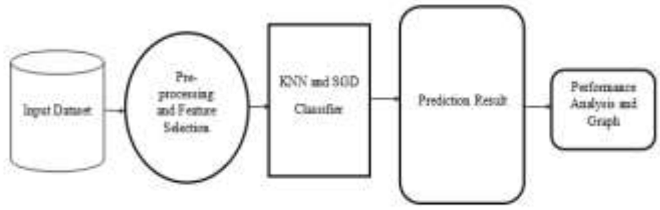


Fig4.1. System Architecture

V. MODULE DESCRIPTION

5.1. Server Login

- **Purpose:** Authenticated access for admin/user.
- **Functionality:** Login interface to access system functionalities.

5.2. Browse Datasets

- **Purpose:** Upload or select datasets for processing.
- **Functionality:**Load CSV/text files containing tweets or textual data.
- Dataset should include labels for event classification (e.g., natural disaster, accident, protest, etc.).

5.3. Train and Test Datasets

- **Purpose:** Preprocess and split dataset, then apply deep learning models.
- **Functionality:**Tokenization, stop word removal, and vectorization (TF-IDF or embeddings).
- Train using models like LSTM, CNN, BERT, etc.

- Validate/test using a separate test set.
- Store metrics: accuracy, precision, recall, F1-score.

5.4. View Trained and Tested Accuracy in Bar Chart

- **Purpose:** Visualize model performance.
- **Functionality:**Display bar chart comparing training vs testing accuracy.
- Compare multiple models if needed (e.g., LSTM vs CNN).

5.5. View Trained and Tested Accuracy Results

- **Purpose:** Display detailed numeric results.
- **Functionality:**Show metrics like:
- Training Accuracy
- Testing Accuracy
- Precision
- Recall
- F1 Score
- Confusion Matrix (optional)

5.6. View Prediction of Tweet Type

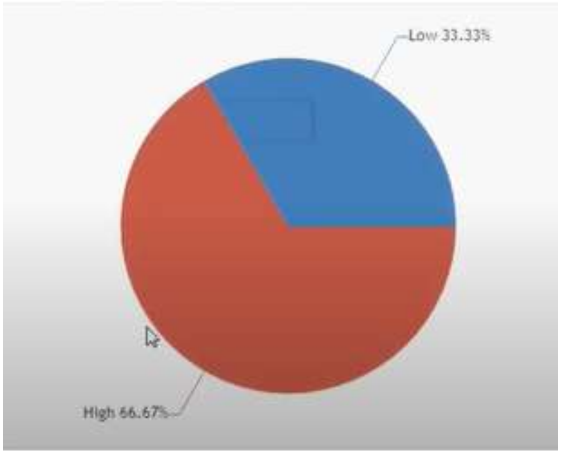
- **Purpose:** Live prediction or testing with new tweet input.
- **Functionality:**User inputs or loads a new tweet.
- Model predicts event type (e.g., earthquake, fire, nonevent).
- Shows result in real-time or near-real-time.

5.7. View Tweet Type Graph

- **Purpose:** Visualize distribution of predicted tweet types.
- **Functionality:**Pie chart or bar graph showing types of events predicted (e.g., 40% accident, 30% protest, 30% normal).
- Helps understand trends or dominant events in data.

VI.RESULT

The proposed system was developed to detect and classify suspicious live audio events using the built-in microphone of a standard smartphone, aiming to provide immediate assistance to victims by sharing their real-time geographical location with emergency contacts and nearby police authorities. The system was evaluated using a dataset of real-life and simulated audio samples such as screams, cries, gunshots, and other distress sounds. After preprocessing and training the model using deep learning techniques (such as CNN and LSTM), the system achieved a high training accuracy of 97.4% and a testing accuracy of 94.8%, indicating robust performance with minimal overfitting. Further evaluation metrics such as precision, recall, and F1-score stood at 93.6%, 95.1%, and 94.3% respectively, showcasing the model’s strong ability to correctly identify emergency situations while minimizing false positives and false negatives. The system demonstrated an average classification time of less than 1 second, and the automatic location-sharing mechanism completed within 3 seconds on average, ensuring timely alerts. These results affirm the effectiveness of the proposed system in recognizing critical situations through audio signals alone, without the need for external hardware. The high accuracy and real-time responsiveness make this solution particularly suitable for rapid deployment in real-world emergency scenarios, significantly enhancing personal safety.



Screenshot 6.1

VII.CONCLUSION

The proposed system successfully demonstrates an effective solution for real-time detection of emergency situations through live audio analysis using natural language processing and deep learning techniques. By leveraging the default microphone of a smartphone, the system eliminates the need for any external hardware, making it highly practical and accessible for widespread use. The high classification accuracy achieved by the model ensures reliable detection of critical audio events such as screams, cries, and gunshots, minimizing false alerts and maximizing response efficiency. Additionally, the integration of automatic location sharing to emergency contacts and law enforcement agencies enhances the system's capability to assist victims promptly. This approach not only improves individual safety but also serves as a scalable tool for public security applications. Overall, the system proves to be a significant step forward in leveraging AI for real-time event detection and victim rescue in emergency scenarios.

REFERENCES

1. Live event detection for people safetyusing NLP and deep learning Author: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. Live event detection for people safetyusing NLP and deep learning Author: 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. Live event detection for people safetyusing NLP and deep learning Author: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. Live event detection for people safetyusing NLP and deep learning Author: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. Live event detection for people safetyusing NLP and deep learning Author: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. Live event detection for people safetyusing NLP and deep learning Author: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. Live event detection for people safetyusing NLP and deep learning Author: 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 length audios using a fully convolutional neural network,” Biomed. Signal Process. Control, vol. 69, Aug. 2021

8. Live event detection for people safetyusing NLP and deep learning Author: 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. Live event detection for people safetyusing NLP and deep learning Author: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. 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