

SheSafe – AI based Emergency Alert System for Women Safety

Mrs. M. Madhavi Latha¹, Banavath Esther Rani², Dayam Kantha Ramola³, Komirisetti Jagruthi⁴ and Sakethapuram Raaga Lahari⁵

¹⁻⁵G. Narayanamma Institute of Technology and Science/Dept. of CSE (Data Science), Hyderabad, India
Email: madhavipanem@gnits.ac.in, 22251A6704@gnits.ac.in, 22251A6712@gnits.ac.in, 22251A6747@gnits.ac.in, 22251A6757@gnits.ac.in

Abstract— In the modern world, women's safety has become a top priority, particularly in circumstances where a manual response might not be possible right away. To solve this problem, the proposed project offers an AI-Based Emergency Alert System for Women Safety, which automatically detects distress by using machine learning and artificial intelligence techniques.

The system uses a Voice Emotion Recognition (VER) model built into a mobile application to analyze voice emotions like fear, panic, or anger. A user's live location, audio recording, and emergency message are shared with pre-designated contacts or local authorities when the system recognizes distress in the user's voice. The model is optimized to reliably distinguish between normal and distressing conditions using a variety of emotional speech datasets. Even in situations where the user is unable to manually operate the device, real-time detection and response are ensured by the Android application, which acts as the user interface. A smart, proactive, and life-saving safety mechanism for women is intended to be provided by this integration of AI-driven emotion recognition with mobile alert systems, improving individual security and fostering a safer community.

I. INTRODUCTION

Women's safety has emerged as a critical issue, especially in situations where immediate assistance is required. Most existing mobile safety applications rely on manual activation such as pressing an SOS button, shaking the device, or opening an app interface. In real emergencies, however, victims may be unable to perform these actions due to panic, shock, or physical constraints, making such systems unreliable.

To overcome these limitations, the proposed system introduces an intelligent emergency alert mechanism driven by voice emotion recognition.

The application continuously monitors speech in real time and identifies distress emotions such as fear or panic using a deep learning model. Once distress is detected, the system automatically sends an emergency alert to pre-registered contacts along with the user's GPS location and timestamp. A voice-based SOS command is also incorporated to enable hands-free activation. Furthermore, GSM fallback ensures alert delivery even in low-network environments.

This work aims to enhance women's safety by providing an automated, real-time, and scalable solution that minimizes human intervention and reduces response time during emergencies.

II. PROPOSED SYSTEM

In the proposed AI-Based Emergency Alert System for Women Safety, two main modules are suggested, namely Voice Emotion Recognition and Alert Processing, which are aimed at guaranteeing the real-time distress monitoring and automatic transmission of the alerts.

A. System Architecture

Flow of the Proposed System

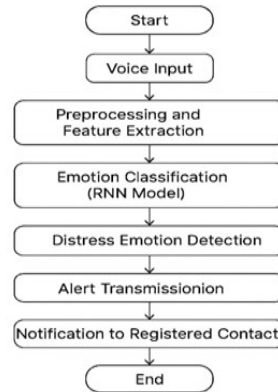


Fig 1

The Voice Emotion Recognition module records the speech of the user, preprocesses and identifies the emotion which could be either fear or panic using a Recurrent Neural Network (RNN)-based emotion classifier. The system is real-time-based and analyzes voice signals to identify distress or prearranged SOS signals (e.g., help me).

When activated, the Alert Processing module automatically captures GPS related information, the time, and the audio evidence of the user and sends the emergency message to the registered contacts over Firebase or GSM SMS fallback when the network coverage is weak.

Such architecture guarantees hands-free and reliable network-independent safety mechanisms that can work even where connectivity is low or in rural areas.

B. Algorithmic Workflow

Voice Emotion Detection (RNN-Based)

The audio signal of the microphone is preprocessed (reduced noise, normalized, and MFCC feature) and simply sent to a ready-trained RNN (LSTM/GRU). The model identifies the emotional tone (fear, anger, panic, and so on) in real time.

Emergency Alert Triggering

In case distress emotion and SOS command is observed, GPS and time data are gathered, and short audio evidence are recorded, and alerts are sent via Firebase or GSM.

Model Training

The RNN model is trained with the help of the preprocessed speech data with the help of the TensorFlow/Keras, which is optimized through the Adam and the categorical cross-entropy. The trained model is transformed into TensorFlow Lite to run inference on Android to provide low latency.

Android Integration

The Android app that was written in Kotlin processes the constant voice input and emotion recognition with the help of the TensorFlow Lite and alerts dispatching with the help of the Firebase/GSM APIs.

III. IMPLEMENTATION AND SYSTEM ARCHITECTURE

The proposed AI-based emergency alarm system, SheSafe, is implemented as an independent Kotlin (Android) mobile app, which is hardware-light, affordable, and real-time safety technology. The system design is modular and relies upon low latency through processing on the device. The central intelligence relies upon a deep

learning Speech Emotion Recognition (SER) model that triggers a strong, multi-channel alarm protocol. The entire system life cycle, from audio reception through to alert send, is designed for highest reliability and lowest intervention time.

A. Technology Stack and On-Device Deployment

The system depends on a proprietary tech stack, best suited for high-performance edge computing:

- Mobile Platform: Native development in Kotlin for Android, leveraging its heightened performance profile and native access to key hardware APIs (microphone, GPS).
- AI Framework: TensorFlow Lite converts and optimizes the trained deep learning model. This will support on-device inference, eliminating the dependence on the network for the essential distress detection functionality, a key requirement of an emergency response system with low latency.
- Core Model: Recurrent Neural Network (RNN) is used in SER due to its improved capacity to model temporal relationships between speech sequences, as well as retaining the dynamic nature of emotional vocal expression.
- Communication Protocols: Firebase is being used by a hybrid alert system for real-time data-driven alerts and a GSM SMS Fallback for assured delivery in low-connectivity areas.

TABLE I: TECHNOLOGY STACK USED

Tool / Library	Purpose
Librosa	Audio loading and feature extraction
NumPy	Numerical operations
TensorFlow / Keras	Model building and training
Scikit-learn	Data preprocessing and encoding
TensorFlow Lite (TFLite)	Model conversion for mobile deployment
Python 3.10 +	Programming environment

B. Real-Time Feature Engineering and Voice Emotion Recognition

The system's intelligence results from strong mapping of raw audio signals to human-understandable features for the RNN classifier.

Audio Preprocessing and MFCC Extraction

The Input Capture Module buffers all microphone audio. Raw signal is heavily preprocessed in the Preprocessing Module:

- Noise Reduction: The background noise is removed using real-time digital filtering, essentially increasing the input voice signal-to-noise ratio.
- Framing and Windowing: The filtered signal is segmented into short overlapping frames (around 20-40 ms) and is subjected to a Hamming window to minimize spectral leakage.
- Feature Extraction: Mel-Frequency Cepstral Coefficients (MFCCs) are extracted per frame. These involve the following transformation: DFT, application of Mel filterbank (modeling the non-linear human hearing system), calculation of logarithmic energy, and a Discrete Cosine Transform (DCT) to de-correlate features eventually, giving a compact consistent feature vector (typically 12 or 13 coefficients with delta and double-delta features).

TABLE II: DATASET DETAILS (RAVDESS DATASET)

Field	Meaning	Example Value
01	Modality (Speech)	03
02	Vocal Channel	01
03	Emotion	06 → Fear
04	Emotional Intensity	01 (Normal) / 02 (Strong)
05	Statement	01 / 02
06	Repetition	01 / 02
07	Actor	01-24

RNN-Based Distress Classification

MFCC feature vector streams are presented as input to the employed RNN model. The Distress Detection Module repeatedly infers to forecast the emotional state of the user in discrete classes (e.g., Neutral, Fear, Panic).

- Automatic Trigger: The alert is automatically triggered when the measure of confidence in a high-arousal negative emotion (Panic or Fear) is above an empirically determined threshold, thereby serving as a hands-free, proactive alert.

- Manual Trigger: Redundant mechanism provides trigger by pre-specified voice command or button press, with immediate feedback in the event of compromised audio quality for the AI model.

TABLE III: TEST CASES FOR SPEECH EMOTION RECOGNITION MODEL

S. No	Test Case	Description	Expected Output	Actual Output	Status
1	Voice sample expressing fear	Emotion = Fear	Emotion = Fear	Emotion = Fear	Pass
2	Voice sample expressing anger	Emotion = Angry	Emotion = Angry	Emotion = Angry	Pass
3	Voice sample expressing calm tone	Emotion = Calm	Emotion = Calm	Emotion = Calm	Pass
4	Voice sample with background noise	Emotion = Neutral	Emotion = Neutral	Emotion = Neutral	Pass

C. Alert Protocol and Redundant Communication

The final step of the system facilitates redundant and consistent presentation of the emergency data to trusted friends and family members.

Emergency Data Compilation

Upon triggering, the Emergency Alert Module does three things at once:

- Location Services: It occupies the device's current GPS location, which is highly precise in localization.
- Evidence Capture: Evidence Recording Module captures a 5-10 second audio recording of what happened. The captured audio is stored and transmitted as irrefutable audio evidence.
- Data Structuring: Packaged in one encrypted bundle of an alert are the captured location, time, and audio evidence.

TABLE IV: MODULE OVERVIEW (IMPLEMENTATION WORKFLOW)

Module No.	Module Name	Key Function
1	Dataset Download and Extraction	Download and extract audio dataset for training
2	Feature Extraction (MFCC)	Extract 40 MFCC features from each audio sample
3	Data Preparation	Encode emotion labels and pair with MFCC features
4	Model Building (LSTM – RNN)	Train deep learning model for emotion recognition
5	Model Training	Train for 10 epochs, batch size 32
6	Model Conversion (TFLite)	Optimize for mobile deployment on Android

Multi-Channel Alert Delivery

Alert Delivery Module utilizes a trusted, two-channel alert delivery system:

- Primary Channel (Firebase): The entire alert packet containing the sound evidence and GPS location is delivered to emergency contacts in real-time using a speed-up internet-based service (Firebase Cloud Messaging).
- Secondary Channel (GSM SMS Fallback): For added redundancy security, the system also provides an ordinary sequence of SMS messages in parallel over the GSM network. The sequence of SMS contains a safety warning in brief and reference to the victim's instant GPS location such that contacts can receive the alarm and coordinates even where mobile data is low or zero. The GSM fallback dramatically improves the reliability of the safety system.

IV. CONCLUSIONS

The application worked in both settings during testing.

It was able to record sound from background, apply the ML models and analyze, and react in very less time. Alerts with location will be sent to emergency contacts and local police stations with the help of GPS integration and APIs.

The system is now capable of:

- Identify distressing feelings of the user.
- Sends location alerts automatically without user input.
- To send messages to contacts in less time.
- Able to do well in low network conditions and areas.

Integration with Wearable Technology: Connecting the system to fitness bands or smartwatches allows it to automatically identify motion, falls, or abnormal heart rates that could be signs of danger so that whenever such situations occur it will directly send location. Multi-Language Emotion Recognition: To improve accessibility across various communities, the emotion detection model is being expanded to support a number of regional languages and dialects.

Using a secure cloud platform to store user information, alert history, and voice recordings for analysis and evidence gathering is known as cloud-based data storage, which is very useful. Integration with Police Networks: For a quicker reaction in an emergency, the alert system can be directly connected to nearby police control rooms or emergency hotlines. Offline Functionality: By utilizing SMS-based alerts and cached location data, the app can function even in places with poor or nonexistent internet connectivity that is known as GSM fallback. Improved AI Models: Using deep learning architectures like CNN-LSTM or transformers to detect even minute changes in tone and stress, the emotion recognition model's accuracy is increased.

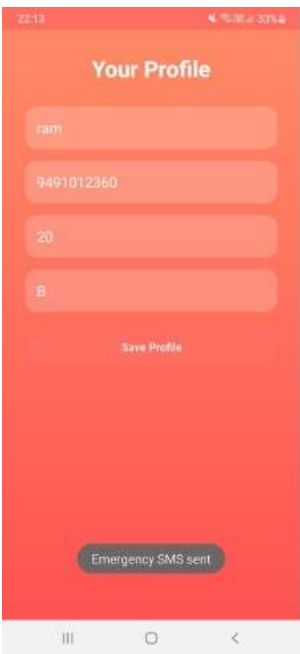


Fig 1

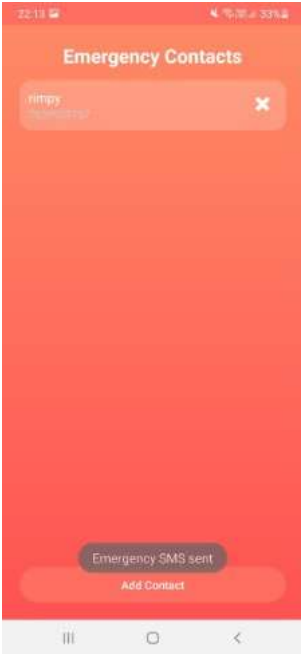


Fig 2

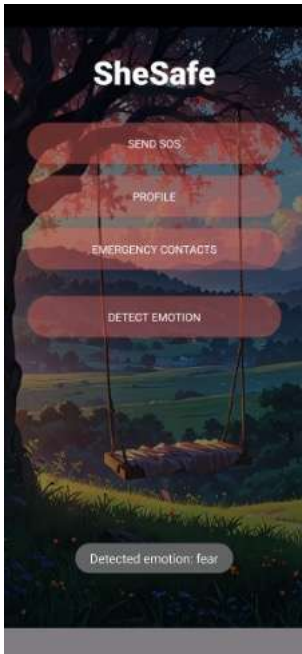


Fig 3

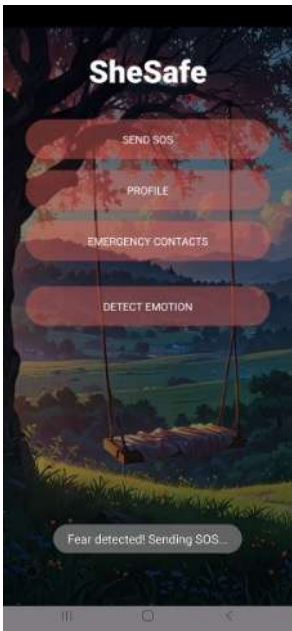


Fig 4

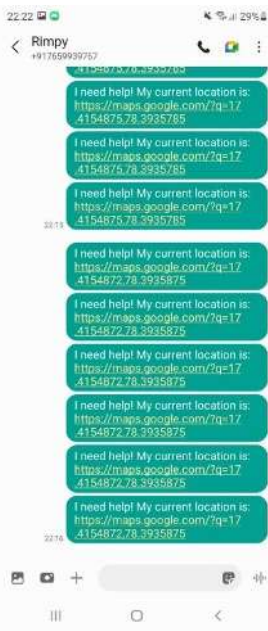


Fig 5

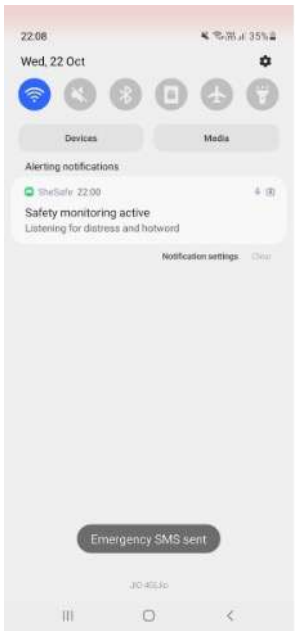


Fig 6

APPENDIX

Voice Recognition: Ability of a program to receive and interpret dictation or to understand and perform spoken commands with the help of Artificial Intelligence.

Language Model: A language model is a statistical model that represents the likelihood of word sequences or sentences in a given language. It is used in natural language processing tasks, such as speech recognition and machine translation, to predict and generate coherent and contextually relevant text.

Sentiment Analysis: Sentiment analysis, also known as opinion mining, is a natural language processing technique that involves determining and extracting subjective information from text, often to understand the sentiment or emotion expressed. It can be used to analyze opinions, reviews, or social media content, categorizing them as positive, negative, or neutral.

ACKNOWLEDGMENT

The authors would like to express their sincere gratitude to their project supervisor and the Department of Computer Science and Engineering for their continuous guidance, encouragement, and valuable feedback throughout the development of this project. Their insightful suggestions greatly contributed to improving the quality and effectiveness of this work.

We would also like to thank our institution, G. Narayanamma Institute of Technology and Science (for Women), for providing the necessary facilities and support to carry out this research successfully. Finally, we extend our appreciation to our peers and mentors whose constructive discussions and cooperation helped in the completion of this project, “AI-Based Emergency Alert System for Women Safety.”

REFERENCES

- [1] B. Sathyasri, U.J. Vidhya, G.J. Sree, T. Pratheeba, K. Ragapriya, Design and implementation of women safety system based on Iot technology, *Int. J. Recent Technol. Eng.* 7 (6S3) (2019)
- [2] ie Cao, Li Zhu, He Han, and Xiaodong Zhu. 2018. Emergency Management System. In *Modern Emergency Management*. Springer Singapore, Singapore, 13–101.
- [3] K.S.SagarReddy, N.Tarakram Sai Kalyan, Sk.Masthansharif, T.Dhanush, V.Venkatesh.: Trigonous Shielding System For Women. In: *International Journal of Research in Engineering, IT and Social Sciences*, Vol. 13 Issue 05, pp. 131-137. (2023)
- [4] Palani Thanaraj Krishnan, Alex Noel Joseph Raj, and Vijayarajan Rajangam. 2021. Emotion classification from speech signal based on empirical mode decomposition and nonlinear features: Speech emotion recognition. *Complex & Intelligent Systems* 7 (2021), 1919– 1934.
- [5] A. Mitchell and M. Minami, *The ESRI Guide to GIS Analysis: Geographic Patterns Relationships*, vol. 1. Redlands, CA, USA : ESRI, 1999.
- [6] K. Chitrivekar, V. Mishra, S. Meshram, and L. Sharma, “Raksha - A Women Safety Application,” *Int. J. Novel Research and Development (IJNRD)*, vol. 8, no. 5, pp. c186–c194, May 2023.
- [7] S. K. Kar, F. S. Prity, N. Fahad, M. M. Hasan, M. K. Morol, and M. A. A. Jubair, “Rokkha: Enabling Offline Emergency Location Sharing via SMS”, *International Conference on Computing Advancements (ICCA 2024)*
- [8] P. S. Rao, “AI-ENabled Wearable Safety Devices for Women Protection,” *IEEE Access*, vol. 11, pp.65432-65441, 2023.