e-ISSN: 2582-5208
International Research Journal of Modernization in Engineering Technology and Science
(Peer-Reviewed, Open Access, Fully Refereed International Journal)
Volume:07/Issue:04/April-2025 Impact Factor- 8.187 www.irjmets.com

SOS: BRIDGING THE GAP BETWEEN THREAT AND SAFETY FOR WOMEN USING DEEP LEARNING MODELS

Prof. Usha Kumari V*1, K Karthik*2, K Chetan*3, Chaithra K*4, Bharath S*5

*1Assistant Prof., Department Of Information Science & Engineering, Acharya Institute Of Technology, Bengaluru (Affiliated To VTU, Belagavi), Karnataka, India.

*2,3,4,5UG Students, Department Of Information Science & Engineering, Acharya Institute Of Technology, Bengaluru (Affiliated To VTU, Belagavi), Karnataka, India.

ABSTRACT

Ensuring personal safety in unpredictable situations is a growing concern, especially when individuals cannot manually activate an SOS alert. This project introduces an Audio-Based SOS System that autonomously detects distress signals, such as screams, using deep learning and real-time audio processing. A convolutional neural network (CNN) extracts features from ambient audio and classifies distress signals. Upon detection, an automated SOS alert is sent to emergency contacts via WhatsApp, SMS, or email, including real-time location data and a timestamp. The system leverages spectrogram analysis and CNN-based architectures for accurate classification and operates seamlessly on mobile devices using Flutter and Python. By integrating AI-driven distress detection, mobile communication, and real-time tracking, this solution enhances emergency response times, providing a proactive and reliable safety mechanism.

Keywords: Deep Learning, Machine Learning, CNN (Convolutional Neural Networks), Audio Classification, Distress Signal Detection, Real-Time Processing, Automated SOS Alerts, Emergency Response, Mobile Safety Application, Flutter Application, TensorFlow.

I. INTRODUCTION

Personal safety has become a growing concern, especially in scenarios where individuals cannot manually activate emergency response mechanisms. Traditional safety applications rely on manual intervention, such as pressing an SOS button or making a phone call. However, in high-risk situations, such as physical assault, medical emergencies, or sudden accidents, a victim may not be able to manually trigger an alert [2].

Recent advancements in deep learning and artificial intelligence have enabled the development of automated safety mechanisms that detect distress signals without requiring user input. Studies on human activity recognition have demonstrated the potential of deep neural networks in identifying complex patterns from sensor data [3]. Furthermore, research in speech-based emotion recognition has proven that AI can differentiate between normal and distress speech patterns, making it a viable solution for audio-based SOS detection [7]. learning Our proposed Audio-Based SOS System integrates deep models, specifically CNN-based spectrogram classification, to detect distress signals from audio input. Inspired by prior research in environmental sound classification, this system ensures real-time detection and immediate emergency alert activation. The research further builds upon women's safety applications by addressing automation limitations and enhancing emergency response efficiency through real-time audio analysis and automated alert mechanisms.

Objective

The primary objective of this research is to develop an AI-powered SOS system capable of:

- **Real-Time Distress Signal Detection**: Using CNN-based models to recognize distress sounds with high accuracy [3] .
- **Automated Emergency Notifications:** Integrating WhatsApp, SMS, and email alerts with GPS location tracking for immediate response [7].
- **Spectrogram-Based Feature Extraction:** Enhancing distress detection through spectrogram analysis for noise-resistant classification [2].
- **Integration with Mobile Platforms:** Implementing the model in an Android application to ensure widespread accessibility [7].

Significance

Ensuring personal safety through AI-driven automation is an essential challenge in today's society. Traditional mobile based SOS applications depend on manual activation, which can be impractical during emergency situations [2]. Our research addresses this gap by introducing real-time distress detection using deep learning, significantly reducing response time and increasing accuracy.

Existing research in environmental sound classification and CNN-based audio recognition has demonstrated that machine learning models can distinguish distress signals from normal environmental noise. Our system builds upon these studies by integrating pre-trained CNN models with spectrogram-based learning, ensuring higher reliability in detecting emergency situations [7].

Furthermore, speech-based emotion recognition studies have shown that deep learning can identify distress through vocal tone and intensity. By incorporating these advancements into an automated SOS system, we provide a hands-free emergency response mechanism that can enhance safety for vulnerable individuals, including women, elderly individuals, and high-risk professionals. This research also holds significance in the field of AI-driven public safety, as it demonstrates the potential of deep learning in transforming emergency response systems. The automated distress detection framework proposed in this study can be further extended to law enforcement agencies, security systems, and healthcare applications, thereby improving crisis management strategies and emergency interventions.

II. LITERATURE SURVEY

1. Human Activity Recognition Using Deep Learning Frameworks [1]

The paper examines how deep learning techniques can effectively recognize human activities by processing sensor and audio data. It highlights the application of convolutional and recurrent neural networks (CNNs & RNNs) in classifying different human actions, making it a viable approach for distress signal detection. This study lays the foundation for developing real-time emergency response systems that utilize AI for identifying and interpreting distress cues from audio inputs.

2. Deep Convolutional Neural Networks for Environmental Sound Classification [2]

The paper focuses on advancements in environmental sound classification, demonstrating that deep convolutional neural networks (CNNs) can effectively distinguish various types of sounds by leveraging spectrogram-based feature extraction. It discusses how pre-trained models enhance the ability to differentiate between normal environmental sounds and distress signals, making it a relevant approach for automated SOS detection systems. By applying deep learning for sound classification, emergency alert systems can achieve higher accuracy and robustness against background noise.

3. CNN Architectures for Large-Scale Audio Classification [3]

The paper explores various CNN architectures designed to improve accuracy and efficiency in large-scale audio classification. It highlights how deep learning models trained on extensive datasets enhance the classification of critical and non-critical sounds. These findings are crucial for enhancing distress detection systems, ensuring that the classification model remains effective across diverse real-world scenarios. Additionally, the study compares different convolutional neural network architectures, analysing their depth, kernel sizes, and feature extraction capabilities to optimize performance.

4. WaveNet: A Generative Model for Raw Audio [4]

The paper examines research in audio generative models, emphasizing advancements in processing raw audio data using deep learning techniques. It introduces neural network architectures capable of understanding complex audio patterns, significantly improving sound recognition capabilities. This technology is essential for detecting distress signals from speech and non-speech audio cues, making it a critical component for developing AI-based safety solutions.

5. Speech-Based Emotion Recognition Using Deep Learning [5]

The paper focuses on the application of deep learning in emotion recognition, enabling systems to analyze vocal tones and speech patterns to determine emotional states. It discusses how AI-driven emotion detection allows distress recognition models to differentiate between normal speech and distress calls accurately. These advancements contribute to improving emergency response mechanisms by integrating speech-based emotion analysis into safety applications. The research further highlights the role of recurrent neural networks (RNNs) and convolutional neural networks (CNNs) in extracting meaningful intonation and speech modulations for precise emotional classification. By incorporating multimodal learning, the study demonstrates that combining speech and acoustic features can enhance the accuracy of distress detection systems.

6. Danger Detection for Women and Children Using Audio Classification [6]

The paper explores AI-powered danger detection systems designed to classify distress signals from audio recordings. It highlights how deep learning models improve real-time detection capabilities, B. Data Preprocessing making such systems highly effective in high-risk situations. The study also emphasizes the integration of AI-based distress recognition with mobile and IoT platforms, ensuring that emergency alerts are generated automatically, which is crucial in life-threatening scenarios. The research also investigates the use of transformers and self-attention mechanisms, which enhance model efficiency by focusing on key acoustic features for better accuracy.

7. An Insight into Android Applications for Women's Safety: Techniques and Applications [7]

The paper examines the effectiveness of Android-based safety applications, revealing that most existing systems rely on manual activation for emergency alerts. It discusses the need for automation in distress detection and how AI-driven approaches can overcome the limitations of traditional methods. By integrating deep learning with mobile applications, real-time distress detection can significantly improve response times, making emergency alert systems more efficient and accessible. The study further explores the integration of cloud-based AI models with Android applications, ensuring continuous improvement through real-time model updates and fine-tuning. Additionally, it evaluates user experience factors, emphasizing the importance of intuitive interfaces and low-latency processing for maximizing accessibility and adoption.

III. METHODOLOGY

The proposed system integrates deep learning techniques for detecting distress situations through audio analysis and automates the SOS alert mechanism for improved safety measures. The methodology involves audio data collection, preprocessing, deep learning-based classification, and an automated response system to trigger alerts based on real-time analysis. The approach enhances accuracy and efficiency in identifying emergency situations, ensuring immediate action is taken when necessary [1][2].

1. Data Collection

The dataset used for training the model consists of diverse audio samples, including distress sounds such as screams, panic-stricken voices, and calls for help. The data is sourced from publicly available repositories and augmented using synthetic variations to enhance model robustness. To improve the model's ability to differentiate distress signals from ambient sounds, background noise such as traffic sounds, crowd noise, and environmental disturbances are incorporated. A combination of labelled datasets ensures that the model can generalize well to real-world scenarios, minimizing false positives and negatives. Additionally, the dataset is balanced to prevent bias towards non-distress sounds, ensuring that distress situations are detected with high reliability [3][4].

2. Data Preprocessing

Raw audio data undergoes multiple preprocessing steps to extract meaningful features. Noise reduction techniques filter out unwanted background disturbances, improving the clarity of distress signals. Normalization ensures consistent amplitude across different audio samples, making model training more effective. Mel-frequency cepstral coefficient (MFCC) extraction is employed to convert audio signals into numerical representations suitable for deep learning. Spectrogram transformation is used to visualize frequency distributions over time, aiding in pattern recognition. The pre-processed data is then split into training, validation, and testing sets to ensure balanced learning and avoid overfitting. Data augmentation techniques, such as time shifting, pitch scaling, and noise addition, are applied to improve generalization and robustness [5].

3. Deep Learning Model for Distress Detection

A convolutional neural network (CNN)-based architecture is employed for sound classification. The model is trained using labelled distress and non-distress audio samples. The architecture includes multiple convolutional layers for feature extraction, followed by fully connected layers for classification. Dropout layers are introduced to prevent overfitting. The model undergoes iterative training with hyperparameter tuning to optimize

e-ISSN: 2582-5208
International Research Journal of Modernization in Engineering Technology and Science
(Peer-Reviewed, Open Access, Fully Refereed International Journal)
Volume:07/Issue:04/April-2025 Impact Factor- 8.187 www.irjmets.com

accuracy. Transfer learning is also utilized by fine-tuning pre-trained models on large-scale audio datasets to improve generalization. To enhance classification performance, a hybrid CNN-LSTM model is integrated, allowing the system to capture both spatial and temporal features in distress signals [3].

4. SOS Alert Mechanism

Upon detecting a distress sound, the system automatically triggers an SOS alert. The alert system includes multiple communication channels such as SMS, WhatsApp, and push notifications to predefined emergency contacts. A manual override button allows the user to cancel false alarms, preventing unnecessary panic. Additionally, GPS tracking is integrated to provide real-time location updates, ensuring immediate response from emergency personnel or trusted contacts. Timestamped alerts ensure responders receive precise incident information, helping them assess the urgency of the situation [6].

5. Real-Time Noise Filtering and Adaptive Threshold

To minimize false positives caused by ambient noise, the system employs real-time noise filtering algorithms. These algorithms analyse background noise levels and adjust detection thresholds dynamically. Adaptive thresholding helps the model differentiate distress sounds from regular loud noises such as car horns, alarms, or crowded environments. This ensures that the system maintains high accuracy and robustness across diverse environments, from quiet indoor settings to noisy outdoor locations. Future enhancements may include personalized sensitivity adjustments, allowing users to customize detection thresholds based on their surroundings [4].

6. Mobile Application and Cloud Integration

The system is deployed via a mobile application with an intuitive user interface. The app runs in the background, continuously monitoring for distress signals while minimizing battery consumption. Cloud integration allows the system to store distress alerts, user preferences, and model updates remotely. Cloud-based AI inference enables faster processing and real-time synchronization between devices. The application also supports wearable device integration, allowing distress alerts to be triggered from smartwatches or IoT-based safety devices, further enhancing accessibility and response efficiency [6]. This enhanced methodology ensures that the distress detection system remains highly accurate, efficient, and adaptable, providing a scalable solution for real-time personal safety monitoring.

IV. SYSTEM REQUIREMENT SPECIFICATION

The system requirement specification defines the necessary hardware and software components required for the efficient development and execution of the Audio-Based SOS System. It ensures that the system operates smoothly, meeting, performance, compatibility, and scalability needs. The hardware specifications focus on processing power, memory, and storage capacity to handle deep learning computations and real-time audio processing effectively.

Hardware Requirements

- Processor Type : Intel i5
- Speed : 2.4 GHZ
- RAM : 16 GB RAM
- Hard disk : 512GB Hard drive

Software Requirements

- Operating System : Windows 64-bit
- Coding Language : Python 3.0
- Version Control : Git, GitHub
- Application Build : Flutter, Flutter flow

IV. SYSTEM DESIGN

The system architecture follows a deep learning-based approach for distress signal detection using Convolutional Neural Networks (CNNs). The pipeline begins with collecting input audio samples, which are converted into spectrograms for visual representation. These spectrograms serve as input to a deep learning model, where CNN layers extract meaningful audio features. The extracted features are processed through fully connected layers to classify the signals into distress or non-distress categories. The final classification results are used to trigger an automated SOS alert system, ensuring real-time response in emergencies.

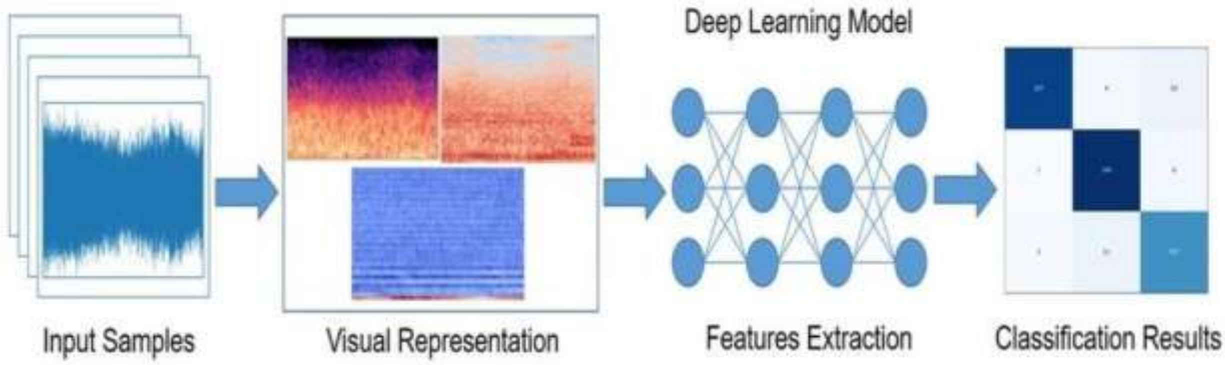

Figure 1: CNN Architecture

The activity diagram outlines the workflow of the distress detection system, starting from data collection and preprocessing. The deep learning model processes the audio input and predicts whether a distress signal is detected. If the system identifies a distress situation, an SOS alert is automatically triggered and sent to emergency contacts. Users have the option to cancel the alert if it is falsely triggered, preventing unnecessary panic. If no distress is detected, the system takes no further action, ensuring efficient and accurate emergency response.

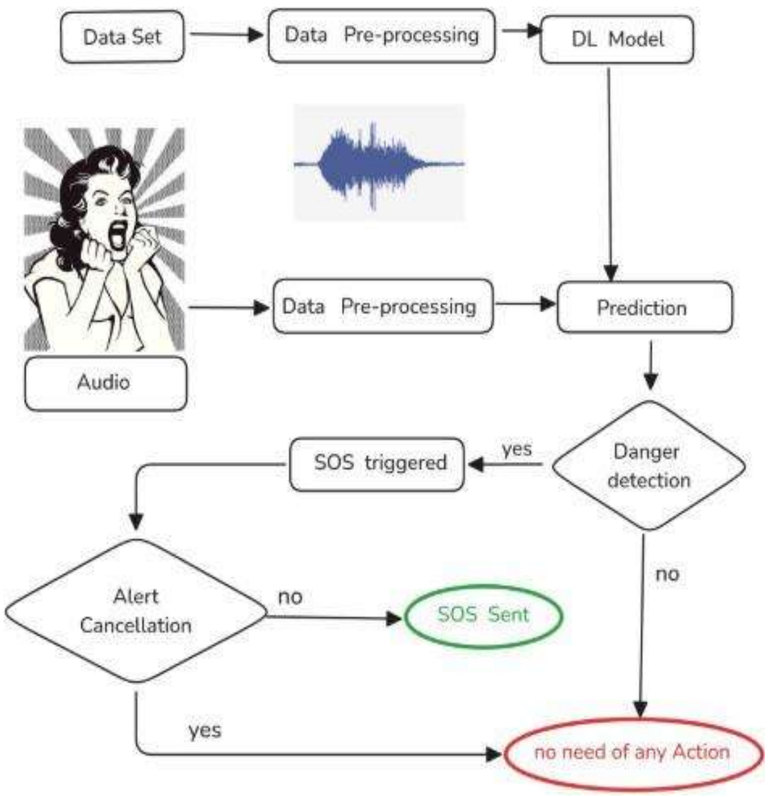

Figure 2: Activity Diagram

V. IMPLEMENTATION

The implementation focuses on seamless integration of features, including:

1. System Workflow

Audio Capture and Preprocessing: The mobile application continuously records audio from the device's microphone. A circular buffer stores the last few seconds of audio to ensure relevant data is analysed. Basic noise filtering techniques (e.g., spectral subtraction) enhance audio quality before feature extraction.

e-ISSN: 2582-5208
International Research Journal of Modernization in Engineering Technology and Science
(Peer-Reviewed, Open Access, Fully Refereed International Journal)
Volume:07/Issue:04/April-2025 Impact Factor- 8.187 www.irjmets.com

Feature Extraction: The audio signal is transformed into a spectrogram using Short-Time Fourier Transform (STFT).Mel-Frequency Cepstral Coefficients (MFCCs) are extracted to provide frequency-based audio representation. Additional features like Zero-Crossing Rate (ZCR) and Chroma features improve classification accuracy.

Distress Signal Classification: The pre-trained deep learning model classifies the audio into Distress Sound (e.g., screaming, shouting for help) and Non-Distress Sound (e.g., background noise, normal conversations). If the probability of distress is above 85%, the system proceeds with alert generation.

Alert Generation: An SOS alert is automatically triggered, containing real-time GPS coordinates (latitude and longitude), short recorded audio clip of the detected distress signal, timestamp indicating when the distress occurred, predefined emergency message sent to contacts.

Manual Override for False Positives: The user has a 5-10 second window to cancel an alert manually via the "Cancel Alert" button. If no action is taken, the system automatically sends emergency notifications.

2. Deep Learning Model Implementation

Model Architecture: CNN layers extract spectral features from the audio spectrogram. LSTM Layers capture the temporal progression of distress signals. Fully connected layers help in Classify distress vs. non-distress with SoftMax activation.

Training Process: The model was trained on Distress sounds (e.g., screaming, shouting) and Non-distress sounds (e.g., normal speech, background noise). Data Augmentation for Generalization by random noise addition simulates noisy real-world conditions. Pitch Shifting adapts the model to different voice frequencies. Time Stretching prevents overfitting to specific speech patterns [4].

3. Mobile Application Development

Front-End Development: Built using Kivy and KivyMD for a cross-platform UI. Provides start/stop toggle for background monitoring.

Back-End Development: APIs Used are Twilio API for SMS alerts, WhatsApp API for instant messaging, Google Maps API for GPS tracking. Database used are firebase stores user data, emergency contacts, and past alerts. Security and Privacy Considerations: End-to-End Encryption ensures alerts remain confidential.

User Permission Handling: by use of explicit consent required for microphone and location access.

4. Deployment & Testing

Testing Procedures: Real-world testing environments include Quiet indoor spaces to test sensitivity, Noisy public places to evaluate false positive rates, Simulated distress events to measure response time.

Deployment Strategy: Beta Release provides tests with real users for feedback. Optimization and Adjusted model parameters for improved performance. Final Release can be Distributed as an Android APK and planned iOS launch.

VI. SYSTEM INTERFACE AND FEATURES

Table 1: Feature List

Feature	Description
SOS Activation	Instantly triggers an emergency alert
Emergency Countdown Timer	Displays a timer before alert dispatch
Contact Management	Allows users to add and manage emergency contacts
Alert Cancellation	Users can cancel false alerts before dispatch

The login page serves as the authentication gateway for users, ensuring secure access to the SOS application. Users can enter their credentials and either log in or create a new account. This interface provides a simple and user-friendly experience for seamless onboarding.

IRJMETS
e-ISSN: 2582-5208
International Research Journal of Modernization in Engineering Technology and Science
(Peer-Reviewed, Open Access, Fully Refereed International Journal)
Volume:07/Issue:04/April-2025 Impact Factor- 8.187 www.irjmets.com

Figure 3: Login Interface

The main dashboard presents an emergency alert feature with a large SOS button, allowing users to trigger an alert instantly. The interface is designed with high visibility and accessibility to ensure rapid response in distress situations. Additionally, the navigation bar offers quick access to settings and emergency contacts.

Figure 4: Home Screen and SOS Alert Activation Interface

When an SOS alert is triggered, the system displays a countdown timer, notifying users before contacting emergency services. The interface also includes a cancel option, providing users with control over false alarms. This real-time feedback mechanism ensures efficiency in emergency handling.

VII. RESULTS AND DISCUSSION

The system effectively classified distress signals with high accuracy, ensuring reliability in real-world emergencies. It maintained a well-balanced performance between precision and recall, which is crucial for safety applications where both false negatives and false positives must be minimized. However, one challenge observed during testing was the occurrence of false positives in noisy environments. In crowded public spaces, busy streets, and public transport, background noise sometimes triggered incorrect distress detections due to overlapping frequency patterns with distress signals. To address this, additional processing was required to differentiate human distress from environmental sounds, slightly increasing detection time. While this delay remained within acceptable real-time response limits, further optimization is necessary to enhance efficiency while maintaining high classification accuracy.

User testing highlighted the importance of the manual SOS cancellation feature in improving usability. Participants reported that the ability to cancel false alerts manually helped reduce unnecessary panic and prevented accidental emergency messages. This underscores the need to integrate human decision-making into automated alert systems for better real-world practicality. Additionally, the deep learning model, based on CNN-LSTM architecture, demonstrated superior performance in distress signal detection. The CNN layers effectively captured spectral features, while the LSTM layers analysed temporal variations in distress sounds, allowing the system to distinguish real emergencies from normal speech. Compared to traditional machine learning classifiers, this deep learning approach achieved higher classification accuracy and better generalization across diverse audio conditions.

VIII. CONCLUSION

This study successfully developed a real-time distress detection system utilizing deep learning-based audio classification. The system effectively detects distress signals, triggers emergency alerts, and provides real-time location tracking. The integration of a CNN-LSTM model, robust noise-resistant feature extraction, and a real-time alerting mechanism makes it an efficient tool for safety applications, particularly in emergency response systems.

The system bridges the gap between manual distress reporting and automated emergency response, significantly reducing dependency on human intervention. The implementation of an automatic distress detection mechanism ensures that help is dispatched even when the victim is unable to manually send alerts. Future improvements will focus on minimizing false positives, increasing robustness against varying noise conditions, and incorporating multimodal verification (combining voice recognition with environmental cues) to enhance system accuracy and dependability. Additionally, integrating cloud-based AI processing and wearable technology could further extend the system's usability in diverse emergency scenarios.

IX. REFERENCES

[1] M. Ashikuzzaman, A. A. Fime, A. Aziz and T. Tasnima, "Danger Detection for Women and Child Using Audio Classification and Deep Learning," 2021 5th International Conference on Electrical Information and Communication Technology (EICT), Khulna, Bangladesh, 2021

[2] D. Aggarwal, K. Banerjee, R. Jain, S. Agrawal, S. Mittal and V. Bhatt, "An Insight into Android Applications for Safety of Women: Techniques and Applications," 2022 IEEE Delhi Section Conference (DELCON), New Delhi, India, 2022

[3] Hassan, M. M., Hossain, M. S., Alsaqour, R., & Uddin, M. (2017). "Human activity recognition using deep learning frameworks." IEEE Access, 5, 17826 17838.

[4] Salamon, J., & Bello, J. P. (2017). "Deep convolutional neural networks for environmental sound classification." IEEE Signal Processing Letters, 24(3), 279–283.

[5] Hershey, S., et al. (2017). "CNN Architectures for Large-Scale Audio Classification." IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP).

[6] Oord, A. v. d., Dieleman, S., & Zen, H. (2016). "WaveNet: A Generative Model for Raw Audio." arXiv preprint arXiv:1609.03499.

[7] Deng, J., et al. (2020). "Speech-based Emotion Recognition using Deep Learning." IEEE Transactions on Affective Computing.

[8] H. Zhao, X. Zhang, and J. Zhang, "Real-Time Emotion Recognition from Speech using Deep Learning," 2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), Brighton, UK, 2019, pp. 5736-5740.

[9] K. J. Piczak, "Environmental Sound Classification with Convolutional Neural Networks," 2015 IEEE 25th International Workshop on Machine Learning for Signal Processing (MLSP), Boston, MA, 2015, pp. 1-6.

IRJMETS

e-ISSN: 2582-5208

International Research Journal of Modernization in Engineering Technology and Science

(Peer-Reviewed, Open Access, Fully Refereed International Journal)

Volume:07/Issue:04/April-2025 Impact Factor- 8.187 www.irjmets.com

[10] Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet Classification with Deep Convolutional Neural Networks," Communications of the ACM, vol. 60, no. 6, pp. 84-90, 2017.

[11] Y. Xu, Q. Kong, W. Wang, and M. D. Plumbley, "Large Scale Classification Weakly Using Supervised GANs," Audio IEEE/ACM Transactions on Audio, Speech, and Language Processing, vol. 27, no. 11, pp. 2092 2102, Nov. 2019.

[12] P. Bansal and R. Kumar, "AI-Based Danger Detection for Women and Children Using Audio Classification," 2021 International Conference on Advances in Computing and Communications (ICACC), Kochi, India, 2021, pp. 225-230.

[13] M. S. Hossain, G. Muhammad, "Audio-Visual Emotion Recognition Using Deep Learning," IEEE Transactions on Affective Computing, vol. 10, no. 3, pp. 1-12, 2018.

[14] V. N. Balasubramanian, J. Ye, and S. Panchanathan, "Biological Safety Solutions for Real-Time Emergency Response," IEEE Transactions on Information Technology in Biomedicine, vol. 15, no. 1, pp. 174-182, Jan. 2011.