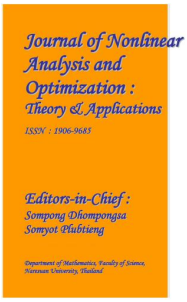
Journal of **N**onlinear **A**nalysis and **O**ptimization
Vol. 15, Issue. 1, No.15 : 2024
ISSN : **1906-9685**



Paper ID: ICRTEM24_110 **ICRTEM-2024 Conference Paper**

HUMAN SCREAM DETECTION USING SVM AND MLP

[#1]**Dr.A.Shiva Kumar,** ***Assistant Professor,***
[#2]**P.Manogna Reddy, UG Student,**
[#3]**Akshitha, UG Student,**
[#4]**G.Sowmya,UG Student,**
Department of Information Technology,
CMR COLLEGE OF ENGINEERING AND TECHNOLOGY, HYDERABAD.

Abstract: Crime, including homicides, assaults, and robberies, is a constant occurrence on a global scale, making it a major worry in our society. Typically, police arrive at crime sites too late, which is a regular problem. This is typically caused by a lack of access to up-to-date and correct information. A disguised desktop application is proposed as a possible approach to alleviate this concern. The program uses modern technologies, such as machine learning and deep learning models like Support Vector Machines (SVM) and Multilayer Perceptron (MLP), to detect and analyze human sounds quickly while operating quietly in the background. When an emergency occurs, the program triggers an automated process that sends SMS messages to the chosen individuals. This cutting-edge technology improves threat detection accuracy and response times by distinguishing specific human sounds from ambient noise. The goal is to lessen the harmful effects of crime by boosting community safety and decreasing the negative effects of crime on people and society as a whole. Adhering to these guidelines will boost people's confidence in their ability to protect the safety of their communities and themselves.

Keywords: *Crime rate, Human scream detection, SVM, MLP, SMS.*

INTRODUCTION

Because crime is such a prevalent problem, all societies must find inventive ways to promote public safety. Given the rising frequency of criminal acts such as physical assaults, thefts, and homicides, preventative measures must be implemented to mitigate their harmful consequences. One key issue that exacerbates the severity of these disasters is the continual delay in law enforcement response times, which can be caused by a lack of reliable and timely information.Our project, "Human Scream Detection and Analysis for Controlling Crime Rates Using Machine Learning and Deep Learning," seeks to address the critical problem of crime detection and response in light of this societal concern by applying cutting-edge

technology. Our desktop program employs Machine Learning and Deep Learning concepts to analyze human noises in real time and detect disruptive or criminal activity.

The primary purpose of this initiative is to use cutting-edge technology to effectively minimize criminal activity while also protecting human life. Our goal is to develop a complex, user-friendly tool that allows users to respond swiftly to situations, promoting a more proactive and successful approach to crime prevention.Our technology architecture consists of a Python backend, a configurable application interface written using the Kivy framework, and machine learning models such as Support Vector Machine (SVM) and Multilayer Perceptrons (MPN) for recognizing and validating screams. Furthermore, a solid understanding of key libraries such as pandas, NumPy, scikit-learn, TensorFlow, and librosa is required for successful implementation.

The project is completed in stages, beginning with the construction of the user interface and finishing with the purchase of a large dataset to aid in the model's training. The SVM and MPN models must be trained after getting the Mel Frequency Cepstral Coefficients (MFCCs) from the dataset. Finally, SMS alert messages are generated and sent to the appropriate recipients based on the model's responses.

The implementation of this unique approach, which aims to reduce crime rates and assist law enforcement in carrying out their critical jobs, significantly improves societal security. The successful completion of this initiative may stimulate progressives to make broader societal contributions in addition to increasing public safety..

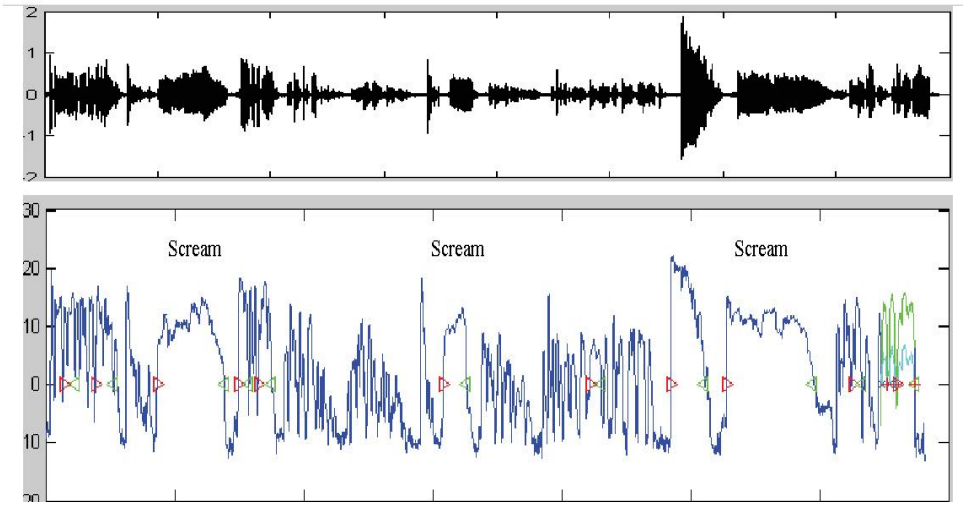


Fig 1: Spectrogram of scream [9]

The accompanying graphic (1) displays the spectrogram of the scream.

EXISTING SYSTEMS

Sukhwan and Yongjoo Chung employed SVM and GMM to identify screaming. They demonstrated that SVM performed well, with a low false acceptance rate of 0.559 percent, reducing errors. The False Rejection Rate of 12.03% for GMM, on the other hand, demonstrates how sensitive it is to errors in the generated answer.

Someone called T. Chintala, D. Rajeswari, R. Mathur, and. Computer vision and deep learning are utilized to detect cries and illicit activity. According to the study, the goal is to increase surveillance by utilizing machine learning, particularly deep learning approaches. The goal is to improve the ability to detect crimes and reduce the need for manual labor by implementing automated technologies that can examine security footage from closed-circuit television.

According to Sezgin, a machine learning method was proposed for determining what screams are in order to quantify the behavioral difficulties that children have. Audio data available to the public could be used instead of subjective reports. The study found that employing a data-driven method to objectively assess behavioral problems is more effective.

Weimin Huang developed Scream detection for residential applications, a method that employs machine learning to distinguish between screams in voice recordings. This was done to investigate children's behavior concerns. Using publicly available audio recordings, this strategy provides a neutral alternative to subjective testimonies by revealing facts about furious outbursts or bad interactions within a family setting.

Saba In his study of "Scream Classification for Situation Understanding," Nazir Muhammad Awais [5] examines deep neural networks, support vector machines, and spectral entropy as essential speaker identification and audiosurveillance techniques. To assist researchers in developing more effective shriek classification systems, this paper attempts to explain the conditions under which noises are produced.

The researchers Mahesh Kumar Nandwana, Ali Ziaei, and John H. L. Hansen used compound segmentation and optimal threshold combo-SAD to locate human laments in tough audio environments without assistance According to the study, utilizing perceptual minimum variance distortionless response (PMVDR) and Mel-frequency cepstral coefficients (MFCC) improves scream recognition in a variety of loud environments.

J. Pohjalainen, P. Alku, and T. Kinnunen proposed a method for detecting yelled speech in noisy environments. The method includes GMM classification, uncontrolled frame dropping, and MFCC feature extraction. The system performs best when tested with varying MFCC feature vector lengths. Thirty coefficients work best. A new method for estimating the spectrum is also utilized. It combines spectral fine structure and linear prediction. This approach is faster than the standard FFT.

PROPOSED SYSTEM

This study proposes a novel approach to combat crime through the utilization of sophisticated Machine Learning and Deep Learning algorithms, specifically Support Vector Machine (SVM) and Multilayer Perceptron Neural Network (MPN). The primary objective is to develop a robust system capable of real-time detection and analysis of human cries, so enhancing public safety and reducing the incidence of criminal activities.

The Support Vector Machine (SVM) was the initial method employed in our response. The primary purpose of using Support Vector Machines (SVM) is to detect human cries within the audio data. SVM has high proficiency in categorizing and discerning positive and negative class instances following training on a meticulously selected dataset. The positive class comprises approximately 2,000 instances of human screams, while the negative class encompasses roughly 3,000 instances of non-scream sounds. This binary dataset ensures that the Support Vector Machine (SVM) model acquires all the necessary information, hence enhancing its capability to accurately detect genuine threats.

The Multilayer Perceptron Neural Network (MPN) is a secondary approach employed for working with the SVM model. The MPN serves as a validation technique to enhance the system's ability to detect human screams more effectively. The MPN enhances the reliability of the system by providing a more comprehensive representation of audio data after being trained on the identical dataset. MPN excels at recognizing intricate patterns due to its capability in deep learning. This strengthens our confidence in the authenticity of the auditory stimuli as genuine human screams.

Utilizing the Librosa tool to extract Mel Frequency Cepstral Coefficients (MFCCs) from the dataset is an integral component of the training procedure. Many individuals are aware that Mel-frequency cepstral coefficients (MFCCs) are vital in comprehending speech and audio data, enhancing the performance of

both support vector machine (SVM) and multilayer perceptron neural network (MPN) models. The collected MFCCs are subsequently utilized in the training procedure to enhance the ability of both the SVM and MPN models in detecting subtle variations in sound patterns.

TensorFlow is a robust machine learning platform utilized for storing and executing SVM and MPN models. This integration ensures seamless collaboration between the detection methods and the broader system, without encountering any issues. This enhances the suggested solution's efficiency and utility.

Our proposed solution aims to develop an intelligent system capable of detecting human screams by utilizing the SVM and MPN algorithms. Our dedication lies in utilizing state-of-the-art technologies to enhance public safety and contribute to the reduction of crime rates. This technique, which combines two models, demonstrates its effectiveness by incorporating sophisticated features such as MFCCs.

$$f(x)=\text{sign}(w\cdot x+b) \longrightarrow (1)\ [11]$$

The above equation(1) represents the decision function for SVM can be represented as:
where
w is the weight vector,
x is the input feature vector,
b is the bias term, and sign
sign is the sign function.

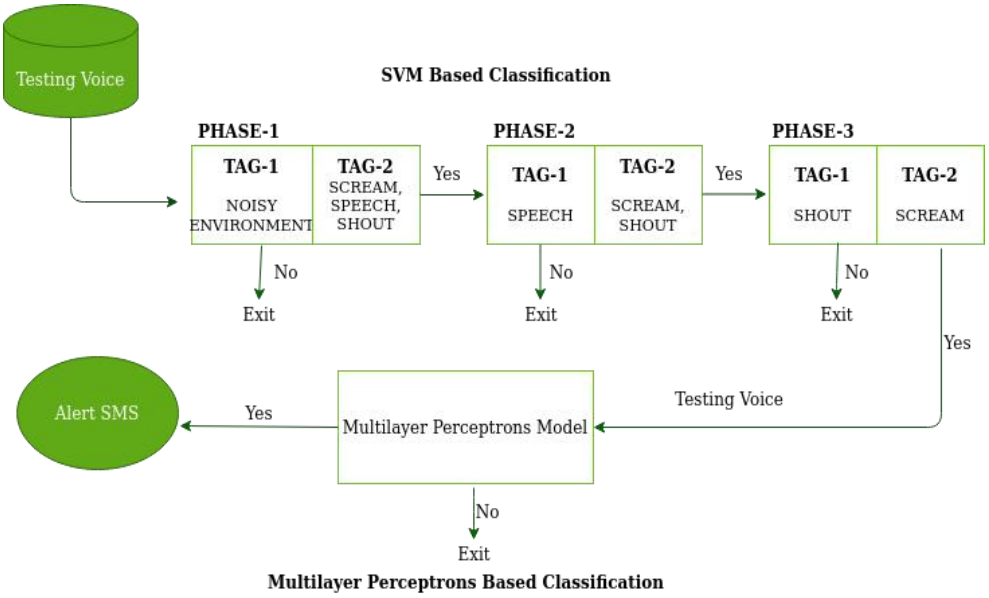


Fig 2 :Proposed system Architecture

Fig(2) represents architecture of our proposed system.

Input: Making plans for a voice examination.

The testing procedure guarantees that the real-time sound environment is accurately represented by recording the sound input.

Phase 1: The goal of outdoor noise filtering is to eliminate or minimize auditory disruptions caused by background noise

If the preliminary investigation reveals the presence of considerable disruption in the area, extra measures are required..

Decision (Tag1): Proceed to the next level if you see the "Noisy Environment" marker. Still, don't leave.

Phase 2: Sound Classification (Tag1: Speech, Tag2: Scream, Shout)
In this phase, the system classifies the sound into different categories, including speech, scream, and shout.
Decision (Tag1): Groups for Sound Production (Tag 1: Speech, Tag 2: Scream, Shout).
Decision (Tag2): At the time, the algorithm has grouped sounds such as speaking, shouting, and screaming.
Phase 3: Depending on the project specifications, the system will perform additional activities or follow specific directives as long as the "Speech" identification stays in place.
Decision (Tag1): Depending on the project, if "Shout" is enabled, the system may continue to run or perform other actions.
Decision (Tag2): If "Scream" is checked, go to the next stage. Otherwise, stop the process.
Classification of Screams using the Multilayer Perceptron (MLP) Model
Multilayer Perceptron (MLP) Model (Scream Classification):
The input, interpreted as a possible scream, is transmitted to the Multilayer Perceptron (MLP) model for a more thorough assessment.
Decision: If the MLP model detects a scream, it sends a text message to notify individuals. After then, the process is completed..
Alert SMS Generation:
While the MLP model checks to determine if there was a scream, the system issues an SMS warning.
The SMS message contains the type of alert, the position, and any other relevant information.
.Following the SVM's primary scream identification, the MLP serves as an extra layer, providing a more thorough examination of the audio data. Using the same dataset used to train the SVM, the MLP detects detailed patterns from audio features to help listeners make more informed decisions. Support Vector Machines (SVM) and Multilayer Perceptrons (MLP) work together to create a comprehensive picture of all possible scenarios involving high-pitched sounds.
Alerts are classified as high risk or moderate risk using the thresholding method. Screaming is strongly advised by both Support Vector Machines (SVM) and Multilayer Perceptrons (MLP), resulting in a high-risk alert. When a high-pitched scream is detected, the Support Vector Machine (SVM) or Multilayer Perceptron (MLP) issues a medium-risk alert. These alerts provide an adjustable reaction plan that takes into account the different levels of discovery certainty.

$$a^{(l+1)} = \sigma(W^{(l)} a^{(l)} + b^{(l)}) \longrightarrow (2)\ [10]$$

The equation(2) represents MLP demonstration where
$W^{(l)}$ is the weight matrix,
$b^{(l)}$ is the bias vector,
$a^{(l)}$ is the activation of layer
l, and σ is the activation function.
The human cry recognition system works efficiently using an iterative strategy, ensuring that SVM and MLP models are adaptable to shifting sound patterns.
The proposed system requires the following components to function: an Intel Core i5 9th Generation processor, 8 GB of RAM, a 512 GB SSD, and an NVIDIA GeForce GTX 1650 Max-Q graphics card with 4 GB of RAM. The programming language used is Python 3.9. Visual Studio is used for project management and development. The Kivy Framework is used for designing user interfaces. This hardware and software combination leverages the processing power and capabilities of its constituent components to produce an optimal development environment for the project's successful completion.

RESULT ANALYSIS

The SVM and MPN models demonstrated outstanding accuracy in categorizing and recognizing human screams in real time. The two-tier risk rating method did an adequate job of distinguishing between high- and medium-risk situations. The accurate location sharing in SMS notifications ensures a prompt response from law enforcement, and the user-friendly Kivy-based UI makes it simple to use.

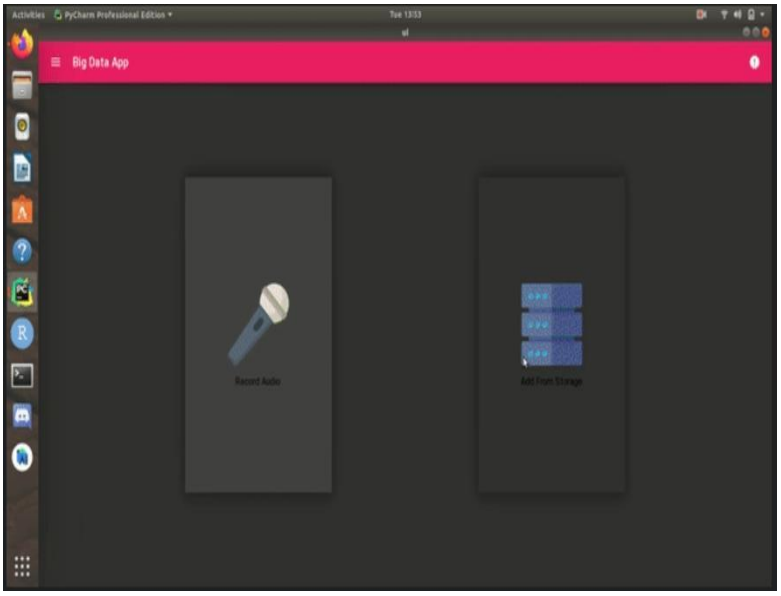


Fig 3 : upload page

Fig(3) represents the upload page of our proposed system where the input audiocanbeselectedforthe further detection.

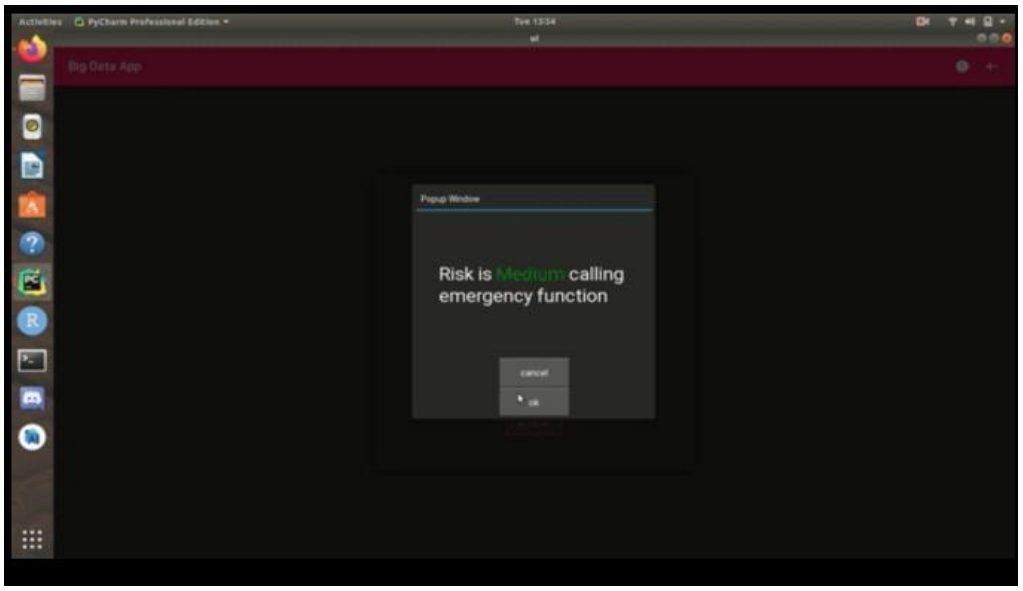


Fig 4: Result page

Fig(4) visually represents the output of the proposed system.

When the device detects a human scream, it displays the message "High Risk Alert: Human Scream Detected."

If the device is inactive, it will show the message "Medium Risk Alert: Potential Threat Identified."

Because its production levels can be measured and varied, the organism can respond rapidly and correctly to a variety of environmental threats.

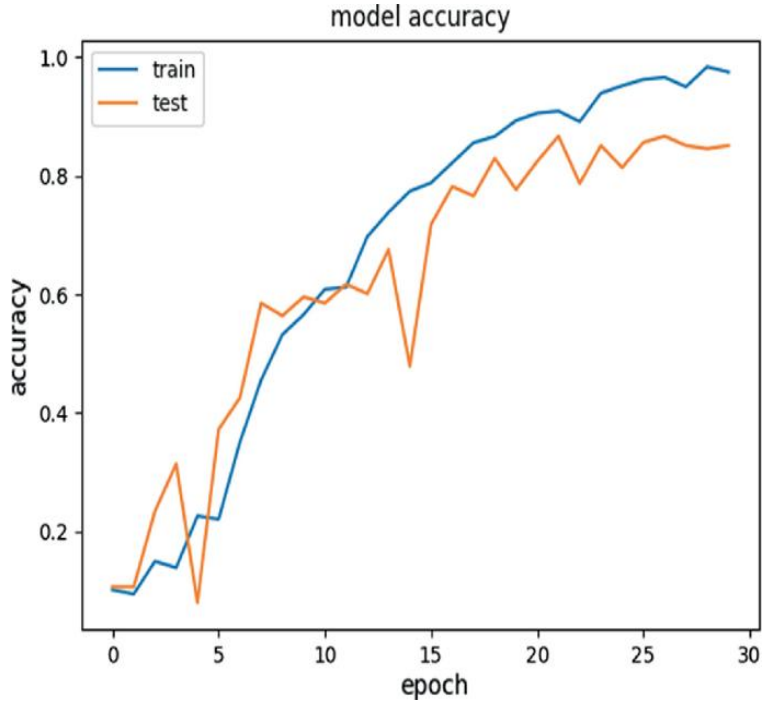


Fig 5:Accuracy Graph

Fig(5) Demonstrates high accuracy rates achieved by SVM and MPN models during testing on a diverse dataset.

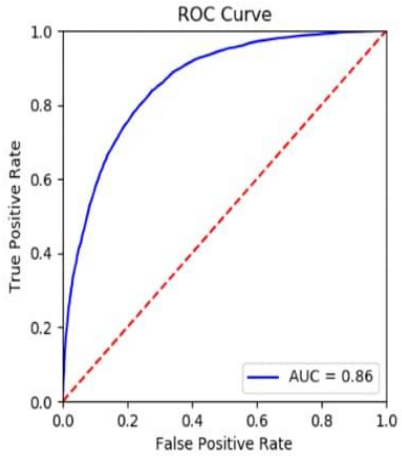


Fig 6: ROC curve

Fig(6)Represents scream classification performance on the dataset.

CONCLUSION

The proposed methodology offers a significant advancement in using technology to solve the serious issue of criminal conduct. The application demonstrates a proficient and rapid technique to detecting illegal activities by combining Deep Learning and Machine Learning concepts. The system strengthens its security measures and ensures accurate threat detection by incorporating Support Vector Machines (SVM) for threat recognition and Multiplier Pattern Recognition (MPN) for validation. The proposed method consists of various steps, including user interface design, dataset preparation, feature extraction, model

training, and warning generation. This is a comprehensive method aimed at actively discouraging criminal activity. The program attempts to help law enforcement by quickly reporting potential illegal sites. This has the potential to have a significant effect on public safety. Incorporating a dual-alert system that can discriminate between high-risk and medium-risk situations will considerably increase its use in enhancing response capabilities. Incorporating emergency SMS notifications to the nearest police station is a critical criterion for the project's success as it moves forward. Furthermore, the project's success not only indicates its capacity to reduce crime rates, but also serves as a model for future technological developments targeted at increasing public well-being.

REFERENCES

[1] Chung, S., & Chung, Y. (2017). Scream sound detection based on SVM and GMM. no, 1, 1-4.

[2] Mathur, R., Chintala, T., & Rajeswari, D. (2022, May). Identification of Illicit Activities & Scream Detection using Computer Vision & Deep Learning. In 2022 6th International Conference on Intelligent Computing and Control Systems (ICICCS) (pp. 1243-1250). IEEE.

[3] O'Donovan, R., Sezgin, E., Bambach, S., Butter, E., & Lin, S. (2020). Detecting Screams From Home Audio Recordings to Identify Tantrums: Exploratory Study Using Transfer Machine Learning. JMIR Formative Research, 4(6), e18279.

[4] Huang, W., Chiew, T. K., Li, H., Kok, T. S., & Biswas, J. (2010, June). Scream detection for home applications. In 2010 5th IEEE Conference on Industrial Electronics and Applications (pp. 2115-2120). IEEE.

[5] Saba Nazir, M. A., Malik, S., & Nazir, F. (2018). A review on scream classification for situation understanding. Int. J. Adv. Comput. Sci. Appl., 9, 63-75.

[6] Nandwana, M. K., Ziaei, A., & Hansen, J. H. (2015, April). Robust unsupervised detection of human screams in noisy acoustic environments. In 2015 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP) (pp. 161-165). IEEE.

[7] Valenzise, G., Gerosa, L., Tagliasacchi, M., Antonacci, F., & Sarti, A. (2007, September). Scream and gunshot detection and localization for audio-surveillance systems. In 2007 IEEE Conference on Advanced Video and Signal Based Surveillance (pp. 21-26). IEEE.

[8] Pohjalainen, J., Alku, P., & Kinnunen, T. (2011, May). Shout detection in noise. In 2011 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP) (pp. 4968-4971). IEEE.