Human Scream Detection and Analysis for Controlling Crime Rate

Dr. Neha Gupta1, Dr. Meenakshi Gupta2, Dr. Niharika Thakur3, Dr. Sarita Yadav4, Mr. Nitin Kumar

1 Princeton University, Princeton NJ 08544, USA

2 Springer Heidelberg, Tiergartenstr. 17, 69121 Heidelberg, Germany

[lncs@springer.com](mailto:lncs@springer.com)

**Abstract:** Crime in itself is the biggest and most spread social enemy. In India and across the globe more and more crimes are reported every second and some of them may even go unnoticed. Crime has assumed several forms in the face of murders , robberies , rapes , violence , homicide etc. and lack of on-time resolution to such serious situations by the police or government catalyses the imbalance in the society and thus, encouraging criminal mindset. Therefore, our project aims to address such a situation by detecting a crime and spontaneously informing nearby officials through some alert messages and location sharing . One of the major challenges faced in the project is detecting accurate features used to differentiate between human scream from various different natural environmental noises encountered in developed metropolitan cities and villages around. Therefore the aim of the project revolves around development of an efficient system that could accurately isolate and identify scream-like patterns amidst this noise.

The project uses Machine Learning and Deep Learning concepts , to perform real-time analysis of human screams and report the situation likewise . Not only this but our sophisticated system is capable of detecting clear human sound from the background noise.

**Keywords:** Human scream detection, Crime rate analysis, Machine Learning and Deep Learning.

# Introduction

* 1. **Background**

Screams are an important audio signal produced by the human voice that indicates distress and potential threats. As cities and urban populations grow, public safety protection becomes even harder to ensure. Security systems that utilize human operators to monitor video feeds will often miss valuable audio signals, such as screams, which play an important role in detecting assaults and thefts along with other violent crimes. The increasing prevalence of audio surveillance systems as a supplementary measure for improved situational awareness in public safety and security is a result of their ability to address these limitations.

The development of scream detection started with work in audio event detection examining how to identify certain sounds among a lot of noise. The introduction of machine learning and deep learning algorithms enabled the development of automated detection methods, which are now more effective and efficient. Current algorithms show a greater degree of accuracy when distinguishing amongst human cries in addition to many other sounds such as car horns and environmental sounds. Utilizing scream detection technology in the analysis of crime, police departments are able to recognize the emergence of crime patterns in hot-spots, improving their ability to prevent these types of crimes and deploy resources appropriately and effectively.

* 1. **Problem Statement**

Automatic Human Scream Detection: This project will develop an automated system that detects human screams in public and in surveillance camera networks, to assist in the prevention of crime.

Improved Response Times: Police agencies will utilize this scream detection technology to improve how fast they respond to emergencies and how safe communities are. Advanced Sound Processing: The detection process will utilize machine learning models and sound processing efforts to discriminate between human screams and other unrelated background sounds.

This technology combines scream detection with geospatial data and criminal analysis by associating a reported scream with a location and time frame to identify places and times when crime is happening.

This ability to pinpoint crime patterns allows law enforcement agencies to develop crime prevention initiatives to facilitate safety in certain locations.

* 1. **Research Objectives**

• Machine learning approaches are needed to develop a scream detection system to differentiate screams (human sounds) from background sounds.

• The system must provide improved detection accuracy through noise reduction techniques and data augmentation methods to deal with a noisy environment.

• Real-Time Scream Recognition and Crime Reporting:

Establish a scream detection system which identifies distress signals while simultaneously sending notifications to police departments.

• Analyze crime data by correlating scream reports with criminal incidents in order to detect emerging crime hotspots.

• Impact on Public Safety: Evaluate system’s effectiveness in reducing response times and crime rates.

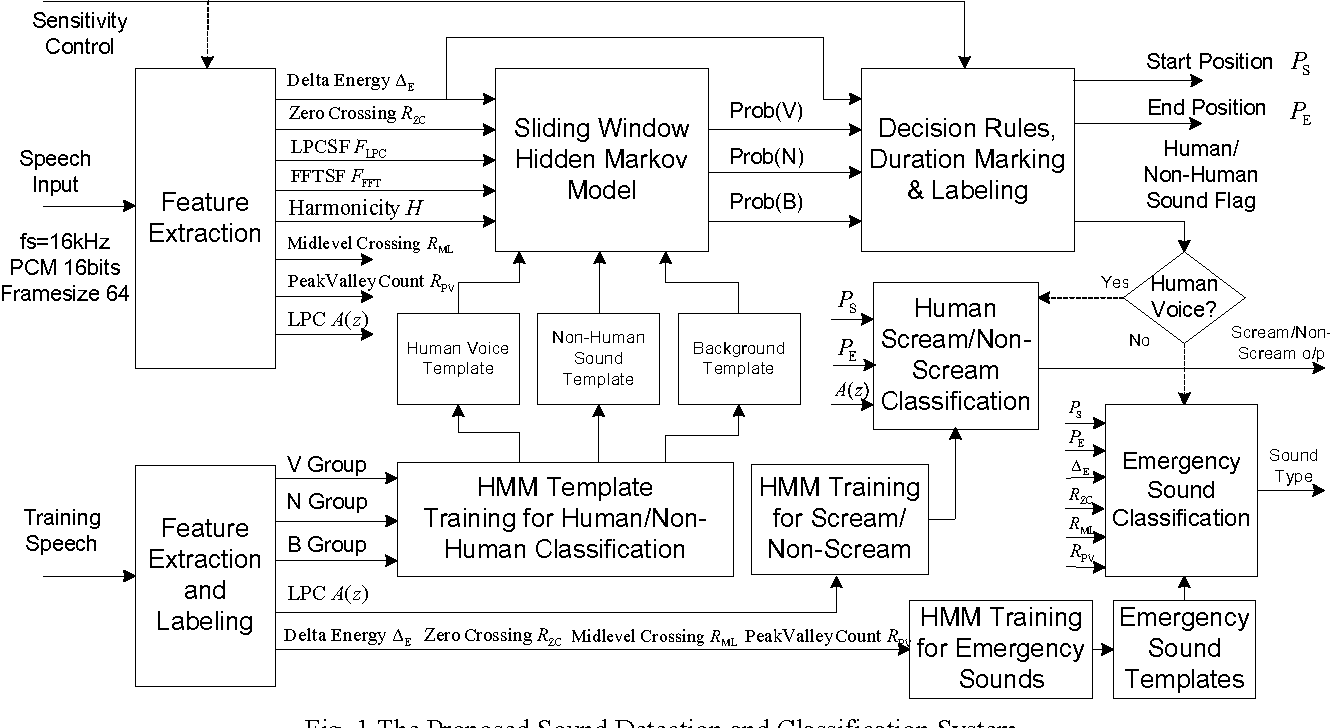
• Ethical Considerations and Privacy Protection: Address ethical and privacy concerns .

# Literature Review

[1] A two-stage supervised learning method is proposed here for scattered-city scream-and-cry detection, with a genuine distress detection rate (DR) of 93.16% and false acceptance rate (FAR) of 4.76% at an SNR of 20 dB. This technique, in essence, considers the SNR of the distress signal. [2] Another approach employs ML to detect screams in voice recordings so as to obviate the necessity to have a large lot of clinical data for model training. The model was trained on publicly available audio datasets and compared against the TV show Supernanny. [3] The first phase of the victim's screams detection in burning sites is an experimental investigation using support vector machine (SVM) and long short-term memory (LSTM) for victim detection in fire crises. The detection and analysis of human sounds using modern technologies such as machine learning and deep learning aim to enhance the accuracy of threat detection and response times. [4] A new three-stage scream detection system based on the K-Nearest Neighbors classifier and Multilayer Perceptron model is able to separate human distress sounds from background noises and send emergency notifications via the Twilio library. [5] This research outlines a scream detection process using analytic and statistical features of sound, namely measures of energy (energy, log energy), autocorrelation, high pitch detection, and a support vector machine (SVM) based classifier. The full algorithm has been tested in a Linux-based set top box with a microphone array, whereby the system could use live data to detect screams in a real time format and the results were reasonably successful. However a solely scream detection heuristics perhaps limits contextual features, the mix of features model, and is open to issues regarding real time.

We are pleased to present a new sound detection and classification system for surveillance applications. The system includes a human/non-human voice classifier and a sliding window Hidden Markov Model (HMM) for scream and gunshot detection. The HMM can detect screams and gunshots in a variety of non-stationary signal-to-noise ratio (SNR) situations. The classification accuracy may be further impacted by overlapping sounds or acoustical environments, or the inability of the adaptation to the particular features used.

[7]This study investigates the sound properties of screams, to identify the weaknesses that traditional speaker recognition methods have when it comes to matching a speaker to a scream. Based on the UT-NonSpeech corpus, the authors were attempting to verify if scream samples could support a speaker verification task, and the findings show that conventional recognition techniques struggle when screams are the entry point.



**Fig 2.1** The Proposed Sound Detection and Classification System

1. **Methodology**

**3.1 Data Collection**

Gather a sizeable dataset that contained both screams and non screams audio. We chose various publicly available datasets as well as recordings we made of scream sounds.

The dataset needed to be classified into scream and non scream folders and labelled. The positive folder had a mix of publicly available and custom screams, while the negative folder contained sounds like traffic sound, conversations, etc.

**3.2 Feature Extraction**

Audio features such as MFCC (Mel-Frequency Cepstral Coefficients), chroma, and spectral features are commonly employed in audio classification.

**3.2.1 MFCC (Mel-Frequency Cepstral Coefficients)**

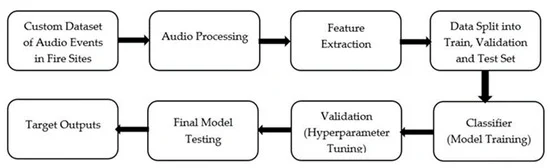
MFCC characterizes the short-term power spectrum of an audio signal and is based on human auditory perception. MFCC has been beneficial in speech recognition, sound classification, and emotion recognition because it captures the most significant frequency components of the signal. MFCC extracts the features by segmenting the audio signal into small frames and applying Fourier transforms.

**3.2.2 Chroma**

Chroma features are all about encoding the pitch content of an audio signal. They are based on twelve different pitch classes. In order to get these features, the audio signal is transformed to the frequency domain and the power of each pitch class is summed together across octave spaces. There are various applications for these features, like music classification, chord recognition, and harmony analysis.

**3.2.3 Spectral Features**

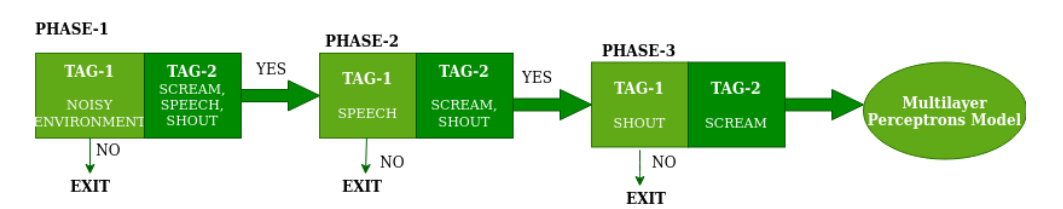
Spectral features analyze a signal's frequency content, from which the shape, energy and distribution of the signal can be inferred. Spectral features must convey meaning, or inform us something about perceptual constructs, such as shape, sound, with a proper range of parameters. Some spectral features include spectral centroid - denoting the center of mass; spectral bandwidth - which tells us the width of the spectral feature; spectral roll-off - this metric tells us how one can tell harmonic sounds from non-harmonic or noisy sounds; and zero-crossing rate (ZCR) - counts the number of times the audio signal crosses the zero-amplitude axis.



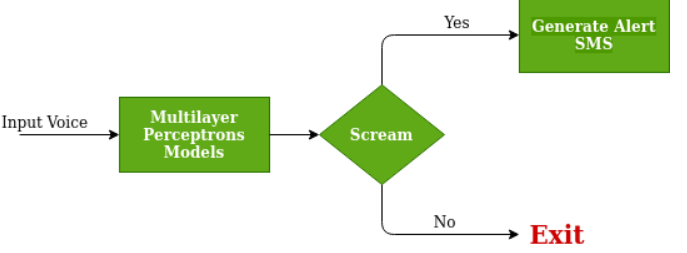
**Fig 3.1** Architectural building blocks for the proposed ML model.

**3***.***3 Model Development**

The model was developed in two stages. The first stage was building an SVM model. In turn, it received three phases of training. The first phase involves separating screams, shouts, and speech from environmental noises. The second phase involves separating speech from shouts and screams to distinguish any normal conversation. The third phase involves separating shouts from screams using their audio features. Once the entire training process is wrapped up, we save our work using the TensorFlow library. After going through these three phases, we isolate any positive audio that contains screams and pass it on to the next stage of the model. In this second stage, we train a multilayer perceptron model on the sounds from our dataset to achieve optimal accuracy. Once the training is complete, we save it again using TensorFlow, just like we did before.



**Fig 3.2** Detection of noise, speech, shout and scream using SVM classifier

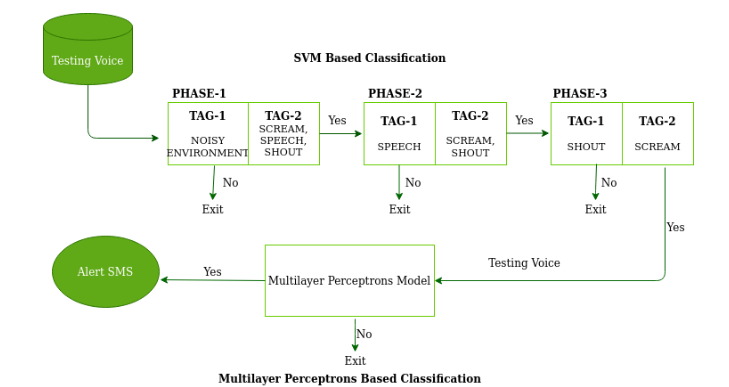


**Fig 3.3** Working of Multilayer Perceptrons Model

**3.3 Generating Alert Messages**

After training and saving the models, they will be tested and their response will determine the level of risk. Alert messages will be generated based on the risk level.

High-risk alert messages will be generated if both models detect human screams in the surroundings, while medium- risk alert messages will be generated automatically if one model detects a human scream in the surrounding area.



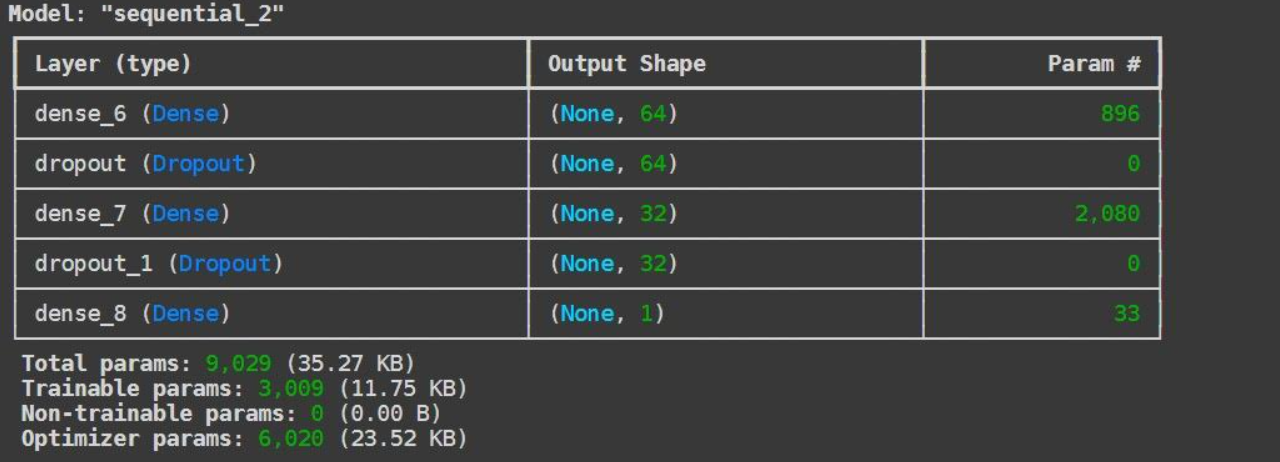
**Fig 3.4** Working Diagram

1. **Result and Discussion**

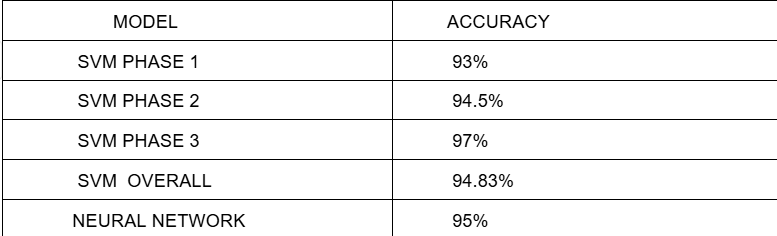
The data for this project came from many websites and audio sources, which were all split into two main classes. The positive class featured approximately 2000 human screams, while the negative class comprised roughly around 3000 non-scream sounds. This dataset became a foundation for training the Support Vector Machine (SVM) model. The SVM model’s evaluation was done on three consecutive classifying stages: Noise vs. Speech: The first stage involved separating human speech from any background noise. Speech vs. High-Pitch: The second stage involved the detection of high-pitched sounds during the speech that had been detected. Scream vs. Shout: In the last phase, screams and shouts were defined as the model authored high-pitched sounds.

The SVM performed extremely well and achieved an overall accuracy score of 0.93. Accuracy this high provides a strong baseline for future comparisons with the neural network model which may be developed in parallel. The evaluation of the performance of these models will be useful in establishing which of them is ideal for instantaneous high-priority alert raising based on the output classification.

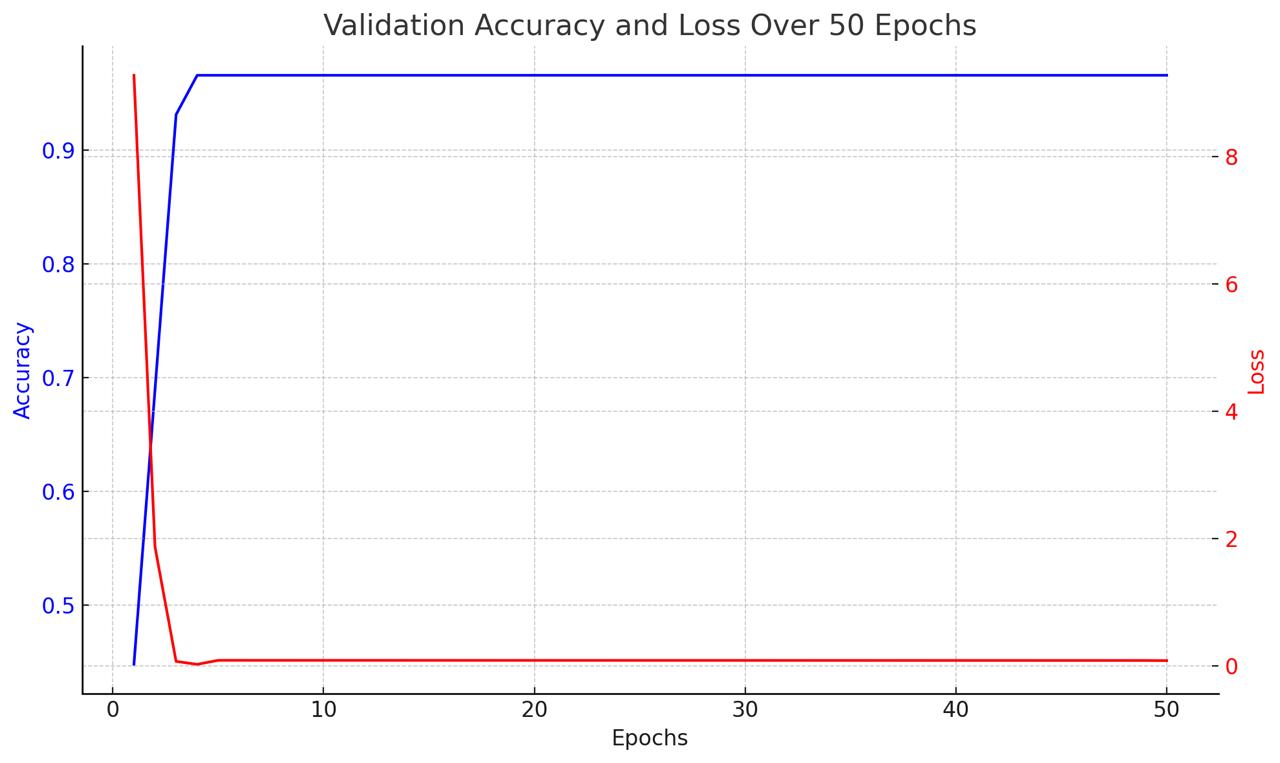
The outcomes demonstrate that the SVM may be an excellent choice for a more advanced scream detection and high-priority alert systems, laying out a benchmark for further improvements.



**Table 4.1** Neural Network Parameters



**Table 4.2** Model Accuracy Results



**Fig 4.1** Validation and Loss over 50 Epochs

1. **Future Scope**

The system has potential for future developments and practical applications, including implementations in a variety of critical domains. For example, by monitoring public spaces such as parks, streets, and markets, the system might be integrated into smart city infrastructure to improve urban safety. This would enable the real-time detection of distress signals, allowing for faster reactions to possible threats or emergencies.

By detecting incidents such as harassment or violent altercations, the technology could be critical in ensuring passenger safety in public transportation hubs such as bus terminals, rail stations, and airports. Such implementations could provide layered security features in addition to existing surveillance systems*.*

1. **References**
2. Saeed, F. S., Bashit, A. A., Viswanathan, V., & Valles, D. (2021). An Initial Machine Learning-Based scream Detection Analysis for Burning Sites.
3. Valenzise, G., Gerosa, L., Tagliasacchi, M., Antonacci, F., & Sarti, A.(2007). Scream and gunshot detection and localization for audio-surveillance systems. IEEE Conference on Advanced Video and Signal Based Surveillance. IEEE.
4. Shankhdhar, A., Rachit, Kumar, V., Mathur, Y. (2021). Human Scream DetectionThrough Three-Stage Supervised Learning and Deep Learning. In: Suma, V., Chen,
5. J.IZ., Baig, Z., Wang, H. (eds) Inventive Systems and Control. Lecture Notes in Networks and Systems, vol 204. Springer, Singapore. [https://doi.org/10.1007/978-981-16-1395-1 28](https://doi.org/10.1007/978-981-16-1395-1%2028).
6. M.K. Nandwana, A. Ziaei, J.H. Hansen, Robust unsupervised detection of humanscreams in noisy acoustic environments, in 2015 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP) (IEEE, 2015), pp. 161–165
7. Khana, N.; Leea, D.; Alia, A.K.; Parka, C. Artificial Intelligence and Blockchain-based Inspection Data Recording System for Portable Firefighting Equipment. In Proceedings of the International Symposium on Automation and Robotics in Construction, Kitakyushu, Japan, 27–28 October 2020; pp. 941–947. [[**Google Scholar**](https://scholar.google.com/scholar_lookup?title=Artificial+Intelligence+and+Blockchain-based+Inspection+Data+Recording+System+for+Portable+Firefighting+Equipment&conference=Proceedings+of+the+International+Symposium+on+Automation+and+Robotics+in+Construction&author=Khana,+N.&author=Leea,+D.&author=Alia,+A.K.&author=Parka,+C.&publication_year=2020&pages=941%E2%80%93947)]
8. Liu, J. Case-based reasoning intelligent decision approach for firefighting tactics. In Proceedings of the 2009 Second International Conference on Intelligent Networks and Intelligent Systems, Tianjian, China, 1–3 November 2009; pp. 437–440. [[**Google Scholar**](https://scholar.google.com/scholar_lookup?title=Case-based+reasoning+intelligent+decision+approach+for+firefighting+tactics&conference=Proceedings+of+the+2009+Second+International+Conference+on+Intelligent+Networks+and+Intelligent+Systems&author=Liu,+J.&publication_year=2009&pages=437%E2%80%93440)]
9. AlHaza, T.; Alsadoon, A.; Alhusinan, Z.; Jarwali, M.; Alsaif, K. New concept for indoor fire fighting robot. *Procedia-Soc. Behav. Sci.* **2015**, *195*, 2343–2352. [[**Google Scholar**](https://scholar.google.com/scholar_lookup?title=New+concept+for+indoor+fire+fighting+robot&author=AlHaza,+T.&author=Alsadoon,+A.&author=Alhusinan,+Z.&author=Jarwali,+M.&author=Alsaif,+K.&publication_year=2015&journal=Procedia-Soc.+Behav.+Sci.&volume=195&pages=2343%E2%80%932352&doi=10.1016/j.sbspro.2015.06.191)] [[**CrossRef**](https://doi.org/10.1016/j.sbspro.2015.06.191)] [[**Green Version**](https://www.mdpi.com/2076-3417/11/18/8425)]
10. Pospelov, B.; Andronov, V.; Rybka, E.; Skliarov, S. Design of fire detectors capable of self-adjusting by ignition. *Вoстoчнo-Еврoпейский Журнал Передoвых Технoлoгий* **2017**, *4*, 53–59. [[**Google Scholar**](https://scholar.google.com/scholar_lookup?title=Design+of+fire+detectors+capable+of+self-adjusting+by+ignition&author=Pospelov,+B.&author=Andronov,+V.&author=Rybka,+E.&author=Skliarov,+S.&publication_year=2017&journal=%D0%92o%D1%81%D1%82o%D1%87%D0%BDo-%D0%95%D0%B2%D1%80o%D0%BF%D0%B5%D0%B9%D1%81%D0%BA%D0%B8%D0%B9+%D0%96%D1%83%D1%80%D0%BD%D0%B0%D0%BB+%D0%9F%D0%B5%D1%80%D0%B5%D0%B4o%D0%B2%D1%8B%D1%85+%D0%A2%D0%B5%D1%85%D0%BDo%D0%BBo%D0%B3%D0%B8%D0%B9&volume=4&pages=53%E2%80%9359&doi=10.15587/1729-4061.2017.108448)] [[**CrossRef**](https://doi.org/10.15587/1729-4061.2017.108448)] [[**Green Version**](http://journals.uran.ua/eejet/article/download/108448/104186)]
11. Mandal, S.; Song, G. Thermal sensors for performance evaluation of protective clothing against heat and fire: A review. *Text. Res. J.* **2015**, *85*, 101–112. [[**Google Scholar**](https://scholar.google.com/scholar_lookup?title=Thermal+sensors+for+performance+evaluation+of+protective+clothing+against+heat+and+fire:+A+review&author=Mandal,+S.&author=Song,+G.&publication_year=2015&journal=Text.+Res.+J.&volume=85&pages=101%E2%80%93112&doi=10.1177/0040517514542864)] [[**CrossRef**](https://doi.org/10.1177/0040517514542864)]
12. Pinales, A.; Valles, D. Autonomous embedded system vehicle design on environmental, mapping and human detection data acquisition for firefighting situations. In Proceedings of the 2018 IEEE 9th Annual Information Technology, Electronics and Mobile Communication Conference, Vancouver, BC, Canada, 1–3 November 2018; pp. 194–198. [[**Google Scholar**](https://scholar.google.com/scholar_lookup?title=Autonomous+embedded+system+vehicle+design+on+environmental,+mapping+and+human+detection+data+acquisition+for+firefighting+situations&conference=Proceedings+of+the+2018+IEEE+9th+Annual+Information+Technology,+Electronics+and+Mobile+Communication+Conference&author=Pinales,+A.&author=Valles,+D.&publication_year=2018&pages=194%E2%80%93198)]
13. Jaradat, F.B.; Valles, D. A Victims Detection Approach for Burning Building Sites Using Convolutional Neural Networks. In Proceedings of the 2020 10th Annual Computing and Communication Workshop and Conference, Las Vegas, NV, USA, 6–8 January 2020; pp. 280–286. [[**Google Scholar**](https://scholar.google.com/scholar_lookup?title=A+Victims+Detection+Approach+for+Burning+Building+Sites+Using+Convolutional+Neural+Networks&conference=Proceedings+of+the+2020+10th+Annual+Computing+and+Communication+Workshop+and+Conference&author=Jaradat,+F.B.&author=Valles,+D.&publication_year=2020&pages=280%E2%80%93286)]
14. Baum, E.; Harper, M.; Alicea, R.; Ordonez, C. Sound identification for fire-fighting mobile robots. In Proceedings of the 2018 Second IEEE International Conference on Robotic Computing, Laguna Hills, CA, USA, 31 January–2 February 2018; pp. 79–86. [[**Google Scholar**](https://scholar.google.com/scholar_lookup?title=Sound+identification+for+fire-fighting+mobile+robots&conference=Proceedings+of+the+2018+Second+IEEE+International+Conference+on+Robotic+Computing&author=Baum,+E.&author=Harper,+M.&author=Alicea,+R.&author=Ordonez,+C.&publication_year=2018&pages=79%E2%80%9386)]
15. [**https://developer.nvidia.com/embedded/learn/get-started-jetson-nano-devkit**](https://developer.nvidia.com/embedded/learn/get-started-jetson-nano-devkit) (accessed on 9 August 2021).
16. Liu, P.; Yu, H.; Cang, S.; Vladareanu, L. Robot-assisted smart firefighting and interdisciplinary perspectives. In Proceedings of the 2016 22nd International Conference on Automation and Computing (ICAC), Colchester, UK, 7–8 September 2016; pp. 395–401. [[**Google Scholar**](https://scholar.google.com/scholar_lookup?title=Robot-assisted+smart+firefighting+and+interdisciplinary+perspectives&conference=Proceedings+of+the+2016+22nd+International+Conference+on+Automation+and+Computing+(ICAC)&author=Liu,+P.&author=Yu,+H.&author=Cang,+S.&author=Vladareanu,+L.&publication_year=2016&pages=395%E2%80%93401)]
17. Hamins, A.P.; Bryner, N.P.; Jones, A.W.; Koepke, G.H. *Research Roadmap for Smart Fire Fighting*; National Institute of Standards and Technology: Gaithersburg, MD, USA, 2015. [[**Google Scholar**](https://scholar.google.com/scholar_lookup?title=Research+Roadmap+for+Smart+Fire+Fighting&author=Hamins,+A.P.&author=Bryner,+N.P.&author=Jones,+A.W.&author=Koepke,+G.H.&publication_year=2015)]
18. Carbon Monoxide Poisoning and FIRE Fighters. Available online: [**https://www.iaff.org/carbon-monoxide/**](https://www.iaff.org/carbon-monoxide/) (accessed on 9 August 2021).ROS. Available online: [**https://www.ros.org/**](https://www.ros.org/) (accessed on 9 August 2021).
19. Nandwana, M.K.; Hansen, J.H.L. Analysis and identification of human scream: Implications for speaker recognition. In Proceedings of the 15th Annual Conference of the International Speech Communication Association, Singapore, 14–18 September 2014. [[**Google Scholar**](https://scholar.google.com/scholar_lookup?title=Analysis+and+identification+of+human+scream:+Implications+for+speaker+recognition&conference=Proceedings+of+the+15th+Annual+Conference+of+the+International+Speech+Communication+Association&author=Nandwana,+M.K.&author=Hansen,+J.H.L.&publication_year=2014)]
20. Chan, C.-F.; Eric, W.M. An abnormal sound detection and classification system for surveillance applications. In Proceedings of the 2010 18th European Signal Processing Conference, Aalborg, Denmark, 23–27 August 2010; pp. 1851–1855. [[**Google Scholar**](https://scholar.google.com/scholar_lookup?title=An+abnormal+sound+detection+and+classification+system+for+surveillance+applications&conference=Proceedings+of+the+2010+18th+European+Signal+Processing+Conference&author=Chan,+C.-F.&author=Eric,+W.M.&publication_year=2010&pages=1851%E2%80%931855)]
21. Pillai, A.; Kaushik, P. AC: An Audio Classifier to Classify Violent Extensive Audios. In *Speech and Language Processing for Human-Machine Communications*; Springer: Berlin/Heidelberg, Germany, 2018; pp. 1–13. [[**Google Scholar**](https://scholar.google.com/scholar_lookup?title=AC:+An+Audio+Classifier+to+Classify+Violent+Extensive+Audios&author=Pillai,+A.&author=Kaushik,+P.&publication_year=2018&pages=1%E2%80%9313)]