INTERNATIONAL JOURNAL OF SCIENTIFIC RESEARCH IN ENGINEERING AND MANAGEMENT (IJSREM)
VOLUME: 07 ISSUE: 12 | DECEMBER - 2023 SJIF RATING: 8.176 ISSN: 2582-3930

A Review on Scream Detection and Classification

Harsh Pandey
Dept. of Information Technology
Maharaja Agrasen Institute of Technology
Delhi, India

Jai Rai Lakhi
Dept. of Information Technology
Maharaja Agrasen Institute of Technology
Delhi, India

Vaishali Goel
Dept. of Information Technology
Maharaja Agrasen Institute of Technology
Delhi, India

Deepika Bansal
Dept. of Information Technolgy
Maharaja Agrasen Institute of Technology
Delhi, India

***Abstract*—In contemporary living environments, the examination of non-speech audio signals is vital for enhancing situational awareness, complementing visual information from video signals. Screams within non-speech audio signals are particularly important for individuals like security personnel, caregivers, and family members due to their association with potential danger. Contrary to traditional beliefs, this review concentrates on automated acoustic systems specialized in non-speech scream classification. The objective is to show that screams can be categorized into emotional classes such as happiness, sadness, fear, and danger. Drawing inspiration from advancements in scream audio detection and classification, the review provides a taxonomy detailing target applications, prominent sound features, classification techniques, and their evolution over recent decades. This comprehensive review aims to assist researchers in choosing the most suitable scream detection and classification techniques, along with relevant acoustic parameters, contributing to a deeper understanding of speaker vocalization conditions.**

Keywords— Scream detection, scream classification, classification techniques, non-speech audio signals.

I. INTRODUCTION

In recent decades, there has been a dedicated exploration into the intricate realm of classifying acoustic data into discernible categories. Audio data, a rich repository of information, serves as a pivotal resource for the nuanced and content-based classification of a diverse array of acoustic signals. At the core of this auditory exploration lies the human vocal tract, a marvel of complexity responsible for producing a plethora of sounds – from the articulate tones of speech to the emotive expressions of crying and laughter. This vocal tract paints a sonic canvas that extends into the realms of both speech and non-speech vocalizations.

Speech, distinguished by intelligible voices structured into coherent sentences, finds decipherability through various Natural Language Processing (NLP) techniques. Conversely, the realm of non-speech vocalizations encompasses a spectrum of expressive sounds, including laughter, sneezing, coughing, snoring, and the distinct vocalizations of screams. It is within the latter, the domain of screams, that our exploration takes center stage. A scream, identified as a non-speech signal, manifests through forceful vocalization, reflecting heightened emotional states and behavioral responses. This emotional tapestry paints screams in diverse hues, from expressions of joy and excitement to signals of danger, manifestations of pain, and echoes of surprise.

The classification and detection of scream sound events extend far beyond academic pursuits, finding extensive applications in the scientific domain. Practical implications resonate in various real-life acoustic systems deployed for purposes such as speaker identification, vigilant oversight facilitated by Audio-Surveillance Systems, and integration into various home applications. Extracting valuable insights from these scream detection systems propels further advancements and applications within this dynamic field.

Recent developments involve fusing time-frequency features with machine learning classifiers, ushering in sophistication in scream detection. Techniques like Support Vector Machines, band-limited spectral entropy, Deep Neural Networks (DNN), Hidden Markov Models (HMM), sound event partitioning, and modulation power spectrum signal a paradigm shift in understanding scream-related acoustics. While many focus on specific acoustic events, this review stands out by offering a comprehensive overview of the state-of-the-art in scream detection and classification. Its unique emphasis, precision, and practical relevance position it as a beacon in exploring scream classification. The goal is to illuminate nuanced concerns and challenges tied to scream classification, scrutinizing and categorizing screams from diverse perspectives. This exploration extends beyond theoretical discussions, featuring a comparative study rooted in problem domains, sound features, and classification techniques within the pages of this review. Presenting a panoramic view, it guides researchers and practitioners in directing their scream-related efforts, navigating optimal sound parameters and classification techniques for enhanced situational understanding through the lens of scream classification.

The structural framework of this review follows a strategic trajectory. Section 2 embarks on a detailed exploration of data collection techniques and the adopted research methodology. Section 3 casts a wide net, offering an insightful overview of different problem domains, sound features, and classification techniques. Section 4 meticulously evaluates diverse data classes, engaging in meaningful discussions surrounding comparisons and accuracy rates. Finally, Section 5 serves as the culminating chapter, synthesizing key insights and concluding this expansive journey through the realms of scream detection and classification.

II. DATA COLLECTION

This review examines 25 distinct research articles related to scream detection and classification across diverse environments. Utilizing highly cited and reputable publications sourced from various digital libraries, the analysis ensures relevance to research interests. A comprehensive evaluation of each article is conducted, focusing on addressing classification challenges that have impeded the advancement and exploration within environments involving scream

TABLE I. SELECTED RESEARCH ARTICLES WITH PROBLEMS DESCRIBED

#	Name/Ref	Year	Problem	Detection/ Classification
1.	Vedant Kalbag et al.[6]	2022	Scream Detection in Heavy Metal Music	Detection
2.	F.S. Saeed et al.[21]	2021	An Initial Machine Learning-Based Victim's Scream Detection Analysis for Burning Sites	Detection
3.	Ashutosh Shankhdhar et al.[22]	2021	Human Scream Detection Through Three-Stage Supervised Learning and Deep Learning	Detection
4.	R. O'Donovan et al.[23]	2020	Detecting Screams From Home Audio Recordings to Identify Tantrums: Exploratory Study Using Transfer Machine Learning	Detection
5.	Xuewen Yao et al.[24]	2020	Classification of Infant Crying in Real-World Home Environments Using Deep Learning	Detection/ Classification
6.	A. Pillai et al.[8]	2018	Classifying violent extensive audios like music, speech, gunshots, and screams	Detection
7.	J. H. L. Hansen et al.[1]	2017	Analyzing human screams for text-independent speaker identification	Detection
8.	N. Hayasaka et al.[4]	2017	Detection of human scream considering noise robustness	Detection
9.	S. Chung et al.[9]	2017	Detecting screams for social problems and violent crimes in public places	Detection
10.	S. Mun et al.[10]	2017	Classification of acoustic scene using screams	Detection
11.	L. Girin et al.[5]	2016	Deep neural networks for automatic detection of screams and shouted speech in subway trains	Detection
12.	Y. Li et al.[11]	2016	Automatically classifying audio events like glass breaking, gunshots, footsteps, and screams for surveillance.	Detection
13.	A. Sharma et al.[12]	2015	Two-Stage Supervised Learning-Based Method to Detect Screams and Cries in Urban Environments	Detection
14.	L. H. Arnal et al.[7]	2015	Using acoustic analysis, psychophysical experiments, and neuroimaging to isolate screaming features, and track their processing in the human brains	Detection
15.	M. Z. Zaheer et al.[14]	2015	A Preliminary Study on Deep-Learning Based Screaming Sound Detection	Detection
16.	J. H. L. Hansen et al.[13]	2015	Robustly detecting screams in noisy areas using unsupervised learning algorithm	Detection
17.	J. Vandermeulen et al. [25]	2015	Discerning Pig Screams in Production Environments	Detection
18.	M. K. Nandwana et al.[15]	2014	Finding out the impact of screaming on the performance of text independent speaker recognition systems	Detection
19.	M. Vacher et al.[16]	2014	Sound classification for patients and elderly people hospitalized at home	Detection
20.	M. Vacher et al.[17]	2014	Detection and classification of acoustic events in a noisy environment	Detection
21.	B. Lei et al.[18]	2014	Power-efficient sound-event detection	Detection
22.	K. Kato[19]	2013	Clarifying audio features of the death growl as well as screaming voice	Detection
23.	B. Uzkent et al.[20]	2012	Classification of non-speech environmental sounds using new feature set	Detection
24.	W. Huang et al.[3]	2010	Detection of human screams using analytic and statistical features as a method of classification	Detection
25.	L. Gerosa et al.[2]	2007	Audio-based surveillance system to detect anomalous acoustic events like screams or gunshots.in public.	Detection

TABLE II. DATA CLASSIFICATION

Sr.	Type	Class
1.	Problem Domain	Surveillance Speaker Identification Feature Enhancement
2.	Classification Techniques	Supervised Learning Unsupervised Learning

This review categorizes selected data into distinct classifications, offering potential alternatives in various directions. Table I showcases literature focusing on scream detection or classification scenarios, outlining the problems addressed in each research article and its efficacy in detecting or classifying screams. The predominant focus of authors centers on integrating scream detection into surveillance systems, given the widely accepted notion that screams signify danger. Some researchers concentrate on augmenting sound features within systems, while others delve into speaker identification through vocal scream samples. Notably, one author indirectly explores scream classification and detection for baby sounds.

These investigations perform an in-depth examination, contrasting crucial elements of various scream detection and classification methods. The primary focus is on ensuring accuracy in detection and classification phases, with the goal of minimizing error rates and choosing the most effective sound features. This assessment places specific importance on assessing the effectiveness and precision of scream classification techniques.

Upon initial inspection of Table I, identifying research gaps can be challenging for newcomers to the field. To address this, each source is categorized based on its problem domain, classification technique, and the results obtained. These categories are further subdivided for a more comprehensive understanding. Additionally, tables and graphs within each category facilitate comparisons between sources, aiding in the identification of the most promising domains, parameters, or techniques for future exploration.

III. DATA CLASSIFICATION

The process of categorizing data into groups and classes to optimize its use is broadly termed data classification, as outlined earlier in this discussion. The gathered data samples from diverse sources undergo analysis based on the parameters detailed in Table II.

A. Problem Domain

Data classification centers around a problem domain, which refers to the specific area of knowledge or application that requires analysis to address a particular issue. Focusing on a problem domain involves narrowing attention exclusively to topics of interest, disregarding extraneous elements. Drawing insights from various research sources, the problem domain is segmented into three primary categories: i) Surveillance, ii) Speaker Identification, and iii) Acoustic Features Enhancement. A detailed exploration of each category is presented below:

1) Surveillance: Surveillance involves overseeing, safeguarding, controlling, or guiding individuals by monitoring unusual activities or evolving information in their environment. Surveillance systems enable distant observation of society for public safety, utilizing electronic devices like audio/video recordings or phone calls. Sound-based surveillance systems enhance remote public protection by examining sound samples from specific locations or individuals, with screams playing a pivotal role in situational analysis for potential danger signs.

2) Speaker Identification: Speaker Identification systems utilize voice biometrics to recognize individuals based on unique vocal features. Distinctive vocal expressions like screams can be effectively employed for text-independent speaker identification.

3) Acoustic Feature Enhancement: A substantial portion of the literature concerning scream analysis focuses on methods dedicated to enhancing acoustic features in scream detection and classification. These techniques aim to bolster the robustness of detection and classification in diverse sound-based systems dependent on screams.

In essence, data classification within the context of scream analysis involves delving into specific problem domains such as surveillance, speaker identification, and acoustic feature enhancement. Each of these domains offers unique insights into the application of scream-related data for diverse purposes, ranging from public safety and speaker recognition to the enhancement of acoustic features for robust scream detection and classification systems.

TABLE III. MACHINE LEARNING TECHNIQUES FOR SCREAM CLASSIFICATION

Category	Classification Technique
Supervised Learning	K-nearest-neighbors (KNN)
	Neural Networks (RBF, MLP)
	Support Vector Machines (SVM)
	Bayesian Networks
	Linear Discriminants
	Rule-based Classifiers
Unsupervised Learning	Hierarchical and Partition Clustering
	Hidden Markov Models (HMM)

	Gaussian Mixture Models (GMM) Clustering
	Neural Networks

B. Techniques for Scream Classification

Classifying screams can be done through various methods, including conventional strategies. Manual classification, carried out by human experts, exemplifies such strategies. Despite its reliability owing to analysts' expertise, this method is time-consuming and demanding. To automate the process and minimize human intervention, two widely used approaches are employed in scream detection and classification. These approaches, supervised and unsupervised, are outlined in Table III, along with their sub-techniques. It's crucial to highlight that semi-supervised learning algorithms are not extensively explored in the realm of scream classification and are hence not addressed in this review.

1) Supervised Learning Algorithms: These algorithms intend to establish a connection between a provided input/vector and the desired output/supervisory signal. Once this association is analyzed and identified, a pattern or inferred function is generated, applicable for mapping new examples. Scream audio event detection systems commonly utilize supervised learning, including techniques like K-nearest neighbor (k-NN), linear discriminant analysis, Bayesian networks, support vector machines, and rule-based machine algorithms. Typically, these algorithms require training behavioral models using labeled data, leading to substantial resource consumption.

a) Instance-Based or K-Nearest-Neighbors (KNN): The K-nearest neighbor algorithm (KNN) is a non-parametric algorithm that is both simple and efficient, offering robust capabilities for organizing and segmenting audio streams. In scream classification, KNN, as employed by the author in [11], classifies based on the majority of neighbors within the k-nearest range. While KNN is straightforward to implement, it faces challenges related to memory and computation complexities.

b) Neural Networks: Modeled after biological neural networks, Artificial Neural Networks (ANNs) mimic human brain information processing. In supervised audio classification, Radial Basis Function (RBF) and Multi-Layer Perceptron (MLP) are employed in ANNs to minimize misclassification errors. MLP maps input datasets to corresponding output sets, frequently utilized in automatic phoneme recognition tasks. Radial Basis Function (RBF), a particular case of a feed-forward network, establishes a linear map from hidden space to the output space.

c) Rule-Based Classifiers: Machine learning based on rules identifies and employs a collection of relational rules representing the system's captured knowledge. Fuzzy rule-based classifiers (FRBC), a variant of this method, are effectively applied in diverse classification tasks. Nevertheless, fuzzy operators lack precise definitions, particularly for symbolic variables.

d) Bayesian Networks: A graphical model that depicts variables and their interconnections, a Bayesian network utilizes a directed acyclic graph (DAG). Its proficiency lies in swift supervised classification, rendering it appropriate for forecasting and classification assignments with extensive datasets.

e) Linear Discriminants: Linear discriminant analysis (LDA) identifies a linear combination of features to categorize two or more classes, transforming raw data into a feature space to achieve robust classification.

f) Support Vector Machines (SVM): SVMs prove beneficial for intricate data classification issues by segregating classes, maximizing the margin between class boundaries and the closest instances.

2) Unsupervised Learning Algorithms: Involving unlabeled input data, unsupervised learning algorithms deduce functions or conclusions, seeking to find and correlate labels. The primary goal is to scrutinize information and unveil similarities between objects. Cluster analysis is a prevalent technique in unsupervised learning, using heuristic data to discern hidden classes and patterns in audio data.

a) Gaussian Mixture Models (GMM): GMMs represent widely employed unsupervised classification techniques, presupposing that data points originate from a mixture comprising Gaussian distributions with undetermined parameters.

b) Clustering: Hierarchical clustering, a cluster analysis technique, seeks to construct a cluster hierarchy through recursive merging or dividing of patterns. Well-recognized methods in this domain include K-means and its variations.

c) Hidden Markov Models (HMM): Utilizing statistical Markov chains based on unobserved variable stats, HMMs serve as the simplest dynamic Bayesian network in scream classification.

d) Neural Networks: ANNs, consisting of a large number of processors and interconnections, are applied for reliable and efficient classification in unsupervised learning. Self-organizing Map (SOM) and Adaptive Resonance Theory (ART) are commonly used neural network models.

IV. RESULT AND DISCUSSION

A comprehensive examination involving 25 research articles centered on the themes of scream classification and detection reveals insightful comparisons based on problem domains and classification techniques. A succinct breakdown of the review for each case is outlined below:

A. Examination of Problem Domains

Scream classification primarily centers on three main problem domains: Surveillance, Speaker Identification, and Feature Enhancement. The applicable research articles are systematically organized for each problem domain, as outlined in Table IV. Out of the 25 articles, 18 are related to individual or public surveillance, 2 concentrate on speaker identification, and 5 explore methods to enhance the experimental outcomes of scream sound vocals. Subsequently, overall percentages are computed for these problem domains to pinpoint areas that may necessitate additional investigation (refer to Fig. 1).

Given the rising rates of public crime in streets and transportation, surveillance systems utilizing audio analysis of screams are rapidly gaining popularity. This increased interest is attributed to screams being widely recognized as signals of survival in humans.

TABLE IV. MACHINE LEARNING TECHNIQUES FOR SCREAM CLASSIFICATION

Sr.	Problem Domain	References	Total
1.	Surveillance	L. Gerosa et al. [2], Huang et al. [3], S. Chung et al. [9], S. Mun et al. [10], L. Girin [5], Vedant Kalbag et al.[6], Y. Li et al. [11], [12], L. H. Arnal et al.[7], M. Z. Zaheer et al. [14], M. Vacher et al. [16], M. Vacher et al. [17], B. Uzkent et al. [20], F.S. Saeed et al.[21], Ashutosh Shankhdhar et al.[22], R. O'Donovan et al.[23], Xuewen Yao et al.[24], J. Vandermeulen et al.[25]	18
2.	Speaker Identification	J. H. L. Hansen et al.[1], M. K. Nandwana et al. [15]	2
3.	Feature Enhancement	A. Pillai et al.[8], N. Hayasaka et al. [4], J. H. L. Hansen et al. [13], B. Lei et al. [18], K. Kato [19]	5

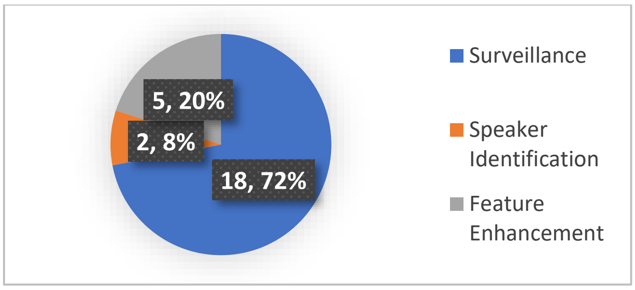

FIG. 1. PERCENTAGE USAGE OF SCREAM DETECTION IN VARIOUS PROBLEM DOMAINS.

systems possess considerable potential for applications in medical surveys, audio scene classification, embedded transport environments such as buses and trains, and continuous monitoring 24x7 for signs of distress in individuals' daily routines.

B. Examination of Classification Approaches

Sound event detection encompasses two distinct strategies: supervised, unsupervised, or a fusion of both. While these approaches have been perceptibly studied, a comprehensive analysis, particularly in the context of scream signal classification, is still lacking. This review categorizes scream classification into two subclasses, offering an in-depth examination of each class.

Table V presents this classification, displaying articles in distinct categories, meticulously observed and assigned to appropriate classes. Some methods utilize solely either supervised or unsupervised approaches, while others combine both. It's crucial to emphasize that direct comparison is hindered by dataset and sound feature variations; instead, the table serves as a comprehensive review of current illustrative approaches.

For a more nuanced view, Fig. 2 delineates that out of 25 research endeavors, 13 employed supervised approaches, 8 employed unsupervised methods, and 3 utilized a combined classification approach. Fig. 3 offers a generic analytical view, indicating that unsupervised approaches have been more successful in the past 15 years for scream detection and classification.

1. Supervised Learning Algorithms: This group comprises K-nearest-neighbors (KNN), neural networks, rule-based classifiers, linear discriminant, and support vector machines (SVM). The main aim is to present supervised learning methods for scream classification, serving as a reference for future researchers investigating automated acoustic environments. Recent experimental research works related to scream classification and detection based on supervised learning methods are summarized in Table VI.

Classifier accuracies are statistically compared by examining the total number of research studies and their classification outcomes. Figure 4 presents the average accuracies of supervised scream classification techniques, demonstrating that Linear discriminants achieve the highest accuracy rate at 96.1%, along with KNN.

2. Unsupervised Learning Algorithms: Unsupervised learning algorithms, a prominent learning approach, have received considerable focus in recent decades. Scream detection and classification using unsupervised methods encompass four categories: clustering, Gaussian Mixture Models (GMM), Hidden Markov Models (HMM), and neural networks (NN). Table VI presents notable research endeavors and their average accuracies, providing insights into addressing challenges in scream classification systems.

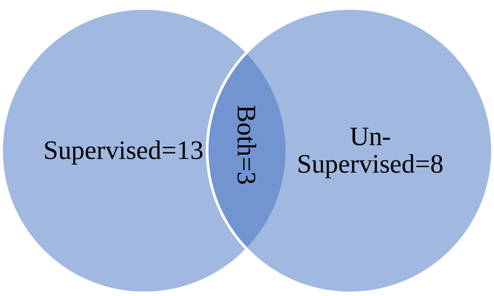

FIG. 2. REPRESENTATION OF SCREAM CLASSIFICATION TECHNIQUES

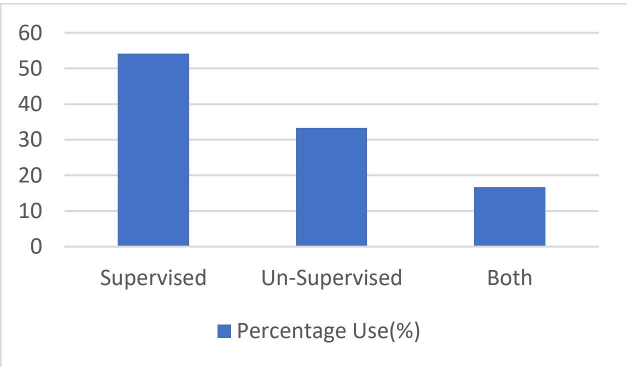

FIG. 3. PERCENTAGE USAGE OF MACHINE LEARNING TECHNIQUES.

Fig. 5 plots the overall average accuracies of the four unsupervised scream classifiers, revealing that GMMs lead with an average classification accuracy rate of 89.89%.

a) Combining Results: The synthesis of overall review results is presented in Table VII, encapsulating research articles spanning 15 years (2007-2022) based on specified sound parameters and classification techniques. Each article's accuracy percentage is mentioned.

Notably, Xuewen Yao et al.'s research in 2020 focused on classification but specifically dealt with infant crying. Another research by K. Kato [19] developed their own techniques for scream detection and classification, bypassing machine learning.

Hence, there is a distinct and broad opportunity to advance scream classification for improved situational comprehension, particularly in supporting sound-based systems, particularly surveillance, to ensure the safety of humans and animals from potential hazards.

TABLE V. MACHINE SCREAM CLASSIFICATION TECHNIQUES

Sr.	Feature Type	References
1.	Supervised	W. Huang et al. [3], L. Girin [5], Vedant Kalbag et al.[6], L. H. Arnal et al.[7], A. Pillai et al.[8], A. Sharma et al. [12], B. Lei et al. [18], B. Uzkent et al. [20], F.S. Saeed et al.[21], Ashutosh Shankhdhar et al.[22], R. O'Donovan et al.[23], Xuewen Yao et al.[24], J. Vandermeulen et al.[25]
2.	Un- Supervised	J. H. L. Hansen et al.[1], L. Gerosa et al. [2], S. Mun et al. [10], J. H. L. Hansen et al. [13], M. Z. Zaheer et al. [14], M. K. Nandwana et al. [15], M. Vacher et al. [16], M. Vacher et al. [17]
3.	Both	N. Hayasaka et al. [4], S. Chung et al. [9], Y. Li et al. [11]

TABLE VI. SCREAM CLASSIFICATION TECHNIQUES

Category	Classification Techniques	References	Total Articles	Average Accuracies
Supervised Learning	Instance-based or K-nearest-neighbors (KNN)	Y. Li et al. [11]	1	96.1%
	Neural Networks (RBF, MLP,CNN,PNN)	L. H. Arnal et al.[7], L. Girin [5], B. Uzkent et al. [20], R. O'Donovan et al.[23], J. Vandermeulen et al.[25]	5	73.1%
	Rule-based Classifiers	A. Pillai et al.[8]	1	84%
	Linear Discriminants	Y. Li et al. [11]	1	96.1%
	Support Vector Machines (SVM)	N. Hayasaka et al. [4], S. Chung et al. [9], A. Sharma et al. [12], B. Lei et al. [18], W. Huang et al. [3], B. Uzkent et al. [20], Vedant Kalbag et al.[6], F.S. Saeed et al.[21], Ashutosh Shankhdhar et al.[22], Xuewen Yao et al.[24]	10	88.47%
Unsupervised Learning Algorithms	Hierarchical and Partition Clustering	M. K. Nandwana et al. [15],	1	66.67%
	Gaussian Mixture Models (GMM)	J. H. L. Hansen et al.[1], N. Hayasaka et al. [4], M. Vacher et al. [17], S.Chung et al. [9], L. Gerosa et al. [2], S. Mun et al. [10], Y. Li et al. [11], M. Z. Zaheer et al. [14], M. Vacher et al. [16]	9	89.89%
	Hidden Markov Models (HMM)	M. Vacher et al. [16]	1	94%
	Neural networks (Self-organizing map)	S. Mun et al. [10], J. H. L. Hansen et al. [13]	2	73.15%

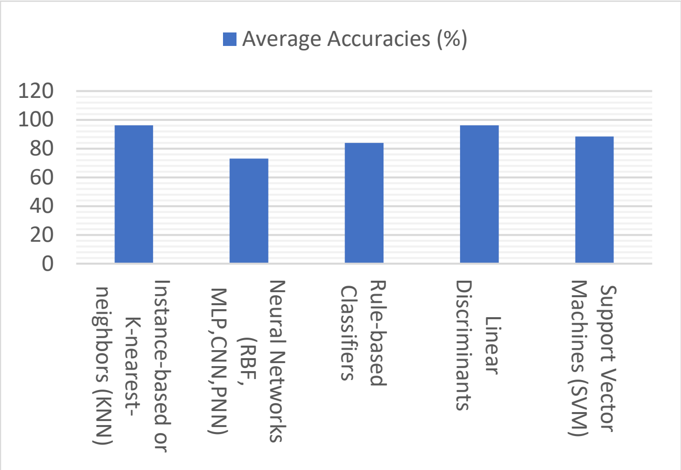

FIG. 4. AVERAGE ACCURACIES OF SUPERVISED SCREAM CLASSIFIERS.

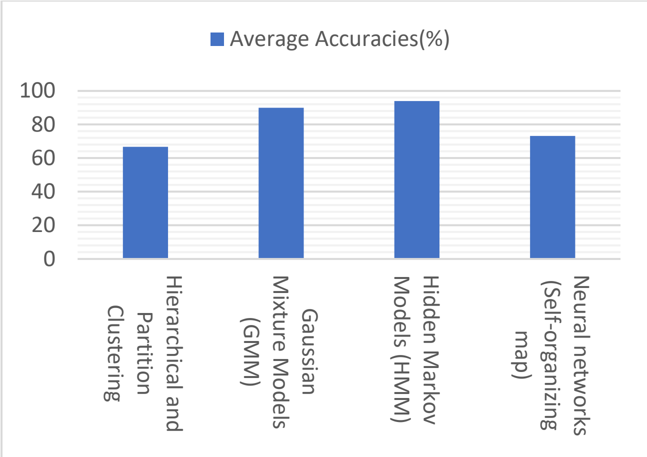

FIG. 5. AVERAGE ACCURACIES OF UN-SUPERVISED SCREAM CLASSIFIERS.

TABLE VII. BRIEF REVIEW OF SCREAM LITERATURE

#	Name/Ref	Year	Problem Domain			Classification Techniques		Detection/ Classification		Accuracy(%)
			Surveillance	**Speaker Identification**	**Feature Enhancement**	**Supervised**	**Un-Supervised**	**Detection**	**Classification**	
1.	Vedant Kalbag et al.[6]	2022	✓			✓		✓		87.6
2.	F.S. Saeed et al.[21]	2021	✓			✓		✓		95
3.	Ashutosh Shankhdhar et al.[22]	2021	✓			✓		✓		90
4.	R. O'Donovan et al.[23]	2020	✓			✓		✓		50
5.	Xuewen Yao et al.[24]	2020	✓			✓		✓	✓	61.2
6.	A. Pillai et al.[8]	2018			✓	✓		✓		84
7.	J. H. L. Hansen et al.[1]	2017		✓			✓	✓		66.67
8.	N. Hayasaka et al.[4]	2017			✓	✓	✓	✓		99.45
9.	S. Chung et al.[9]	2017	✓			✓	✓	✓		87.035
10.	S. Mun et al.[10]	2017	✓				✓	✓		86.3
11.	L. Girin et al.[5]	2016	✓			✓		✓		93.8
12.	Y. Li et al.[11]	2016	✓			✓	✓	✓		96.1
13.	A. Sharma et al.[12]	2015	✓			✓		✓		93.16
14.	L. H. Arnal et al.[7]	2015	✓			✓		✓		50

15.	M. Z. Zaheer et al.[14]	2015	✓				✓	✓		100
16.	J. H. L. Hansen et al.[13]	2015			✓	✓		✓		60
17.	J. Vandermeulen et al.[25]	2015	✓			✓		✓		83
18.	M. K. Nandwana et al.[15]	2014		✓			✓	✓		66.67
19.	M. Vacher et al.[16]	2014	✓				✓	✓		94
20.	M. Vacher et al.[17]	2014	✓				✓	✓		86.46
21.	B. Lei et al.[18]	2014			✓	✓		✓		92.76
22.	K. Kato[19]	2013			✓	N/A	N/A	✓		50
23.	B. Uzkent et al.[20]	2012	✓			✓		✓		88.7
24.	W. Huang et al.[3]	2010	✓			✓		✓		89.815
25.	L. Gerosa et al.[2]	2007	✓				✓	✓		93

V. CONCLUSION

This review thoroughly examines researchers' efforts in detecting and classifying screams, providing an in-depth taxonomy of systems dedicated to this task. The goal is to improve accuracy and understand the speaker's contextual situation. The primary focus is on machine learning, classification methods, and crucial sound parameters for scream-centric audio systems. While the recommended approach suggests using the unsupervised learning technique, particularly GMM, incorporating spectral sound features like MFCC in surveillance, it is crucial to recognize potential risks linked to high scream detection rates in surveillance systems [26]. These conclusions are derived from diverse datasets, using various combinations of sound parameters and classification techniques, and outcomes may vary based on dataset specifics and background noise levels.

As we progress in the domain of scream detection, numerous avenues emerge for further exploration and enhancement, such as Hybrid Models, Real-world Datasets, Explainable AI, and Edge Computing.

In summary, the evolution of scream detection research portrays a dynamic field marked by ongoing improvements and challenges. The future holds the promise of more precise, adaptable, and ethically sound scream detection systems, contributing to a range of applications from entertainment to public safety.

REFERENCES

[1] J. H. L. Hansen, M. K. Nandwana, and N. Shokouhi, "Analysis of human scream and its impact on text-independent," vol. 2957, 2017.

[2] L. Gerosa, M. Tagliasacchi, F. Antonacci, and I. Politecnico, "Scream and Gunshot Detection and Localization for Audio-Surveillance Systems," 2007.

[3] W. Huang, T. K. Chiew, H. Li, T. S. Kok, and J. Biswas, "Scream Detection for Home Applications," pp. 15–18, 2010.

[4] N. Hayasaka, A. Kawamura, and N. Sasaoka, "Noise-robust scream detection using band-limited spectral entropy," AEUE - Int. J. Electron. Commun., vol. 76, pp. 117-124, 2017.

[5] L. Girin, "Deep Neural Networks For Automatic Detection Of Screams And Shouted Speech In Subway Trains Ifsttar , COSYS , LEOST , Villeneuve d ' Ascq INRIA Grenoble Rhˆ," pp. 6460-6464, 2016.

[6] Vedant Kalbag, Alexander Lerch, "Scream Detection in Heavy Metal Music," arXiv:2205.05580v1 [cs.SD], May 11, 2022.

[7] L. H. Arnal et al., "Human Screams Occupy a Privileged Niche in the Communication Soundscape Report Human Screams Occupy a Privileged Niche in the Communication Soundscape," Curr. Biol., pp. 1– 6, 2015.

[8] A. Pillai and P. Kaushik, "AC : An Audio Classi fi er to Classify Violent Extensive Audios," 2018.

[9] S. Chung and Y. Chung, "Scream sound detection based on SVM and GMM," no. 1, pp. 1–4, 2017.

[10] S. Mun, S. Shon, W. Kim, D. K. Han, and H. Ko, "Deep Neural Network Based Learning and Transferring Mid-Level Audio Features for Acoustic Scene Classification," pp. 796–800, 2017.

[11] Y. Li and G. Liu, "Sound Classification Based On Spectrogram," no. 2, 2016.

[12] A. Sharma and S. Kaul, "Two-Stage supervised learning-based method to detect screams and cries in urban environments," IEEE/ACM Trans. Audio Speech Lang. Process., vol. 24, no. 2, pp. 290-299, 2016.

[13] J. H. L. H. Mahesh Kumar Nandwana, Ali Ziaei, "Robust Unsupervised Detection Of Human Screams In Noisy Acoustic Environments," pp. 161–165, 2015.

[14] M. Z. Zaheer, J. Y. Kim, H. G. Kim, and S. Y. Na, "A preliminary study on deep-learning based screaming sound detection," 2015 5th Int. Conf. IT Converg. Secur. ICITCS 2015 - Proc., 2015.

[15] M. K. Nandwana, J. H. L. Hansen, E. Jonsson, and C. Science, "Analysis and Identification of Human Scream : Implications for Speaker Recognition," no. September, pp. 2253–2257, 2014.

[16] M. Vacher and M. Vacher, "Sound Classification in a Smart Room Environment : an Approach using GMM and HMM Methods," 2014.

[17] M. Vacher et al., “Sound Detection and Classification for Medical Telesurvey,” 2014.

[18] B. Lei, “Sound-Event Partitioning and Feature Normalization for Robust Sound-Event Detection,” no. August, pp. 389-394, 2014.

[19] K. Kato and A. Ito, “Acoustic features and auditory impressions of death growl and screaming voice,” Proc. - 2013 9th Int. Conf. Intell. Inf. Hiding Multimed. Signal Process. IIH-MSP 2013, pp. 460–463, 2013.

[20] B. Uzkent, B. D. Barkana, and H. Cevikalp, “Non-speech environmental sound classification using SVMs with a new set of features,” no. January 2015, 2012.

[21] F.S. Saeed, A. Al Bashit, V. Viswanathan, D. Valles, "An Initial Machine Learning-Based Victim’s Scream Detection Analysis for Burning Sites," Appl. Sci. 2021, 11(18), 8425; https://doi.org/10.3390/app11188425, September 10, 2021.

[22] Ashutosh Shankhdhar, Rachit, Vinay Kumar, Yash Mathur, "Human Scream Detection Through Three-Stage Supervised Learning and Deep Learning," Springer Singapore, June 8, 2021.

[23] R. O'Donovan, E. Sezgin, S. Bambach, E. Butter, S. Lin, "Detecting Screams From Home Audio Recordings to Identify Tantrums: Exploratory Study Using Transfer Machine Learning," JMIR Publication - https://preprints.jmir.org/preprint/18279, February 17, 2020.

[24] Xuewen Yao, Megan Micheletti, Mckensey Johnson, Kaya de Barbaro, "Classification of Infant Crying in Real-World Home Environments Using Deep Learning," ResearchGate, May 13, 2020.

[25] J. Vandermeulen, C. Bahr, E. Tullo, I. Fontana, S. Ott, M. Kashiha, M. Guarino, C. P. H. Moons, F. A. M. Tuyttens, T. A. Niewold, D. Berckmans, "Discerning Pig Screams in Production Environments," April 29, 2015.

[26] S. Nazir, M. Awais, S. Malik, F. Nazir, "A Review on Scream Classification for Situation Understanding," IJACSA - International Journal of Advanced Computer Science and Applications, 2018.