

Article

A Review of Automated Bioacoustics and General Acoustics Classification Research

Leah Mutanu ¹, Jeet Gohil ¹, Khushi Gupta ², Perpetua Wagio ¹ and Gerald Kotonya ^{3,*}¹ Department of Computing, United States International University Africa, Nairobi P.O. Box 14634-0800, Kenya² Department of Computer Science, Sam Houston State University, Huntsville, TX 77341, USA³ School of Computing and Communications, Lancaster University, Lancaster LA1 4WA, UK

* Correspondence: g.kotonya@lancaster.ac.uk

Abstract: Automated bioacoustics classification has received increasing attention from the research community in recent years due its cross-disciplinary nature and its diverse application. Applications in bioacoustics classification range from smart acoustic sensor networks that investigate the effects of acoustic vocalizations on species to context-aware edge devices that anticipate changes in their environment adapt their sensing and processing accordingly. The research described here is an in-depth survey of the current state of bioacoustics classification and monitoring. The survey examines bioacoustics classification alongside general acoustics to provide a representative picture of the research landscape. The survey reviewed 124 studies spanning eight years of research. The survey identifies the key application areas in bioacoustics research and the techniques used in audio transformation and feature extraction. The survey also examines the classification algorithms used in bioacoustics systems. Lastly, the survey examines current challenges, possible opportunities, and future directions in bioacoustics.

Keywords: sound classification; bioacoustics; survey; review; acoustic detection; general acoustics



Citation: Mutanu, L.; Gohil, J.; Gupta, K.; Wagio, P.; Kotonya, G. A Review of Automated Bioacoustics and General Acoustics Classification Research. *Sensors* **2022**, *22*, 8361. <https://doi.org/10.3390/s22218361>

Academic Editor: Fabio Di Pietrantonio

Received: 1 October 2022

Accepted: 21 October 2022

Published: 31 October 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Automatic acoustic classification also referred to as audio or sound classification, involves the detection or recognition of sound using audio informatics for storage and retrieval, and machine learning techniques for autonomous classification [1–5]. Bioacoustics is the branch of acoustics that is concerned with sounds produced by or affecting living organisms. Bioacoustics is often used in acoustic sensing to monitor biodiversity, especially in visually inaccessible areas [6]. Animal acoustic emissions contain species-specific information that reflects the character and behavior of different living organisms [1]. There are three main application areas of bioacoustics [1]. The first focuses on the classification and analysis of sounds vocalized by different animal species. Its primary aim is to identify sounds that characterize species in different behavioral contexts. The second is concerned with integrating sound signals vocalized by animals with behavioral contexts to understand how the sounds affect the behavior and emotions of the receiver. The third explores the production mechanisms used in sound vocalization processes [1]. The survey presented in this paper explores how current research in automated bioacoustics classification differs from traditional acoustic classification with respect to the techniques used and application areas. We use the term “general acoustic studies” to refer to acoustic research whose primary focus is neither living or non-living organisms.

The scope of our survey is limited to studies that use machine learning as the primary tool for automating acoustic classification. The survey is intended to be a representative rather than an exhaustive review of the state of the research. The survey reviewed 124 publications, spanning 21 years, from 2000–2021. Only papers published in the English language were reviewed. To the best of our knowledge, no recent studies have been undertaken to examine the state of research in this important and fast-growing research area.

Our survey highlights the advances in automated bioacoustics classification, but also identifies the challenges and opportunities presented. For example, we note that the automated classification techniques used bioacoustics still lag behind those in general acoustics. A number of machine learning techniques that have been successfully used in general acoustics are yet to be tested in bioacoustics classification.

The survey sought to answer four questions relating to current bioacoustics research:

- RQ1: What are the main application areas?
- RQ2: What sound data processing and classification techniques are used?
- RQ3: How have the applications described in the studies been implemented?
- RQ4: To what extent have previously identified research problems been addressed by current studies?

Our findings show that current research in bioacoustics is mainly concerned with applications that involve species classification while general acoustic research is primarily concerned with identifying suitable machine-learning algorithms for classifying general sounds. The short-term Fourier transformation (STFT) technique was the most popular audio transformation technique for both bioacoustics and general acoustics studies. Although Mel-frequency cepstral coefficients (MFCCs) and feature extraction techniques were popular in both bioacoustics and general acoustics research, linear prediction cepstral coefficients (LPCCs) techniques were more popular in general acoustics. In bioacoustics research, ensemble classification algorithms were more popular while in general acoustic studies, convolutional neural networks (CNN) classifiers were more popular. Only half of the publications surveyed provided the implementation details of their systems (i.e., architectural design and theoretical background). Most general acoustic studies also described the system workflows, unlike bioacoustics studies. All the studies had a strong focus on results.

The rest of this paper is organized as follows; Section 2 provides a brief background on related work. Section 3 describes the methodology used in the review. Section 4 reports on the results of the review. Section 5 provides a summary of automated bioacoustics research and future trends. Section 6 provides some concluding thoughts.

2. Related Work

This section reviews existing surveys on acoustic classifications to provide the context and motivation for our work. The first survey on bioacoustics sound classification was published in 2010 [7], with the first general acoustics classification survey appearing four years later, in 2014 [8]. Since then, the number of surveys has steadily grown, as shown in Figure 1. The size of the circles indicate the number of surveys published in that year. However, while current surveys suggest significant growth in bioacoustics classification research, many research challenges remain. For example, most surveys focus on well-known taxonomic groups such as birds, and mammals [9] due to the lack of open-source datasets for other species [10]. Secondly, tropical regions are poorly represented in the surveys despite their rich diversity of flora and fauna [11]. Another challenge relates to the running costs of the IoT devices used in data collection. Most of the IoT devices are deployed in remote locations where they are intended to run autonomously for long periods of time, making their operational lifespan crucial in mitigating their running costs. As the devices are battery-operated it is important that effective ways are found improve their energy efficiency. An important aim of our survey was to establish the extent to which the research challenges identified in past surveys have been addressed by current work on acoustics classification.

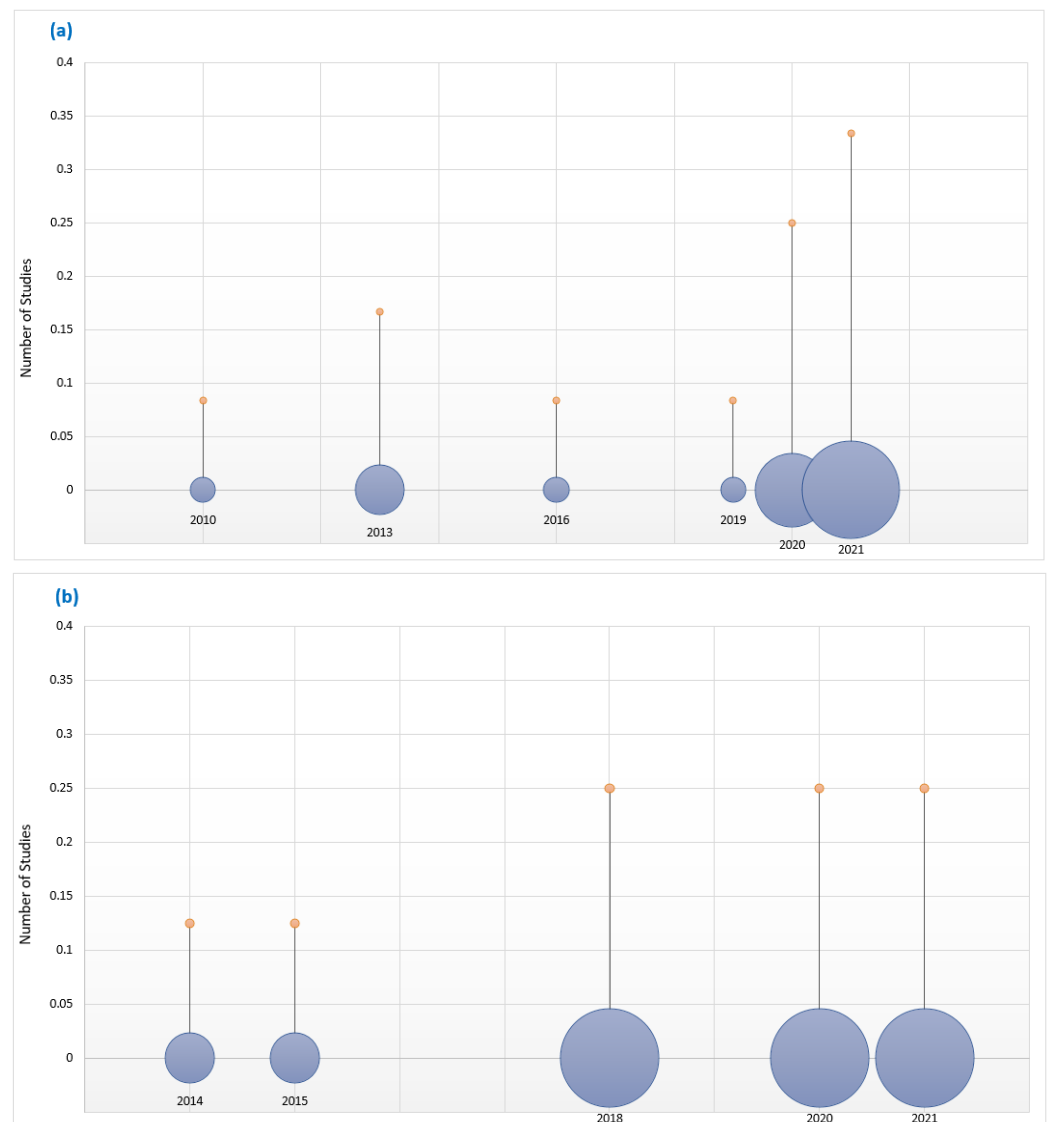


Figure 1. Growth of (a) bioacoustics and (b) general acoustics research over the years.

Current research in acoustics classification spans disciplines such as zoology, engineering, environmental sciences, physics, computer science, and medicine; thus, the range of datasets that we used to source the studies described here vary widely. Out of the 31 survey publications analyzed, twelve focused on bioacoustics sound and the rest on general acoustics. A significant number of bioacoustics survey publications (7) focused on the medical aspects, while general acoustic papers focused on the technology. However, there is growing interest in investigating the technical aspects of bioacoustics classification as highlighted in [9,10,12–15]. Early reviews [14,15] highlighted Mel-frequency cepstral coefficients (MFCCs) and hidden Markov model (HMM)-based classifiers as the popular acoustic preprocessing and classification techniques. However, recent surveys identify deep learning [13] and ensemble methods as better classification techniques. Other reviews note that widespread use of modern acoustic classification techniques is hindered by the lack of adequate datasets [10] and better de-noising techniques [9,12].

To establish the relevance of existing surveys to our own survey, we conducted a word cloud search to identify comparable surveys. The outcome indicates that the selected surveys used machine learning techniques to identify sounds made by animals. The word cloud search also shows that surveys on bioacoustics monitored biodiversity, characterized vocalizations, or investigated animal behavior. The search shows that the studies in general acoustics surveys focused largely on environmental awareness through sound recognition.

The results also show that the selected surveys are relevant and highlights the extent of surveys in acoustic sound classification.

Further analysis of the surveys revealed that most (55%) of the bioacoustics publications included study demographic information such as year, publisher, and implementation details, as shown in Figure 2. Additionally, most surveys focused on either bioacoustics [1–12,16–22] or general acoustics classification [13,23,24] without direct comparisons. This makes it difficult to share lessons and good practice between the two.

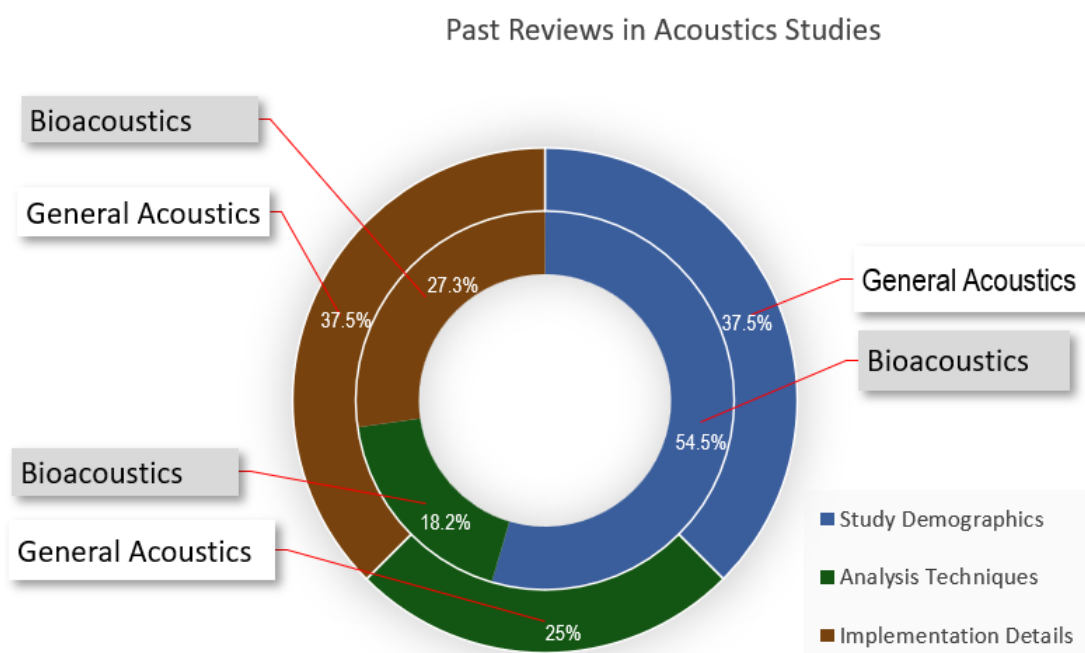


Figure 2. Analysis of previous reviews in acoustics classification.

3. Methodology

According to [25], reviews with an understanding goal focus more on interpretation than deductive logic. Understanding may be accomplished with the help of two types of reviews; scoping reviews and critical reviews [26]. This review uses a scoping approach where a broader perspective that strives to discern a subject's overall meaning and relationships is used. The analysis of survey papers consists of six key steps: problem formulation, literature research, screening for inclusion, quality assessment, data extraction, and data analysis and interpretation [26]. The scoping review methodology used in this study excludes quality assessment and therefore uses five of these steps as recommended by [26]. The process is described next.

3.1. Problem Formulation

The problem identification process was used to examine related work in past surveys. From this exercise, the research objectives identified were: (i) conducting a comparative analysis of acoustic classification techniques based on their application areas, (ii) highlighting the challenges (gaps) in current research on bioacoustics classification techniques, and (iii) making recommendations for a research agenda for bioacoustics classification techniques based on the application areas.

3.1.1. Literature Search

After examining past reviews, the study mined research papers that addressed the identified research objectives from publications in peer-reviewed research datasets. We

screened the relevant papers through an extensive review of literature on the design of bioacoustics and general acoustics classification techniques. This systematic review of the literature used various online databases that index computer science and technology research, namely: IEEE, Science Direct, PubMed, ACM Digital Library, Elsevier, MDPI, Nature, PLOS one, Taylor and Francis, and Springer. The search keywords used were: environmental sound classification, animal sound classification, bioacoustics sound classification, and general acoustics sound classification. To enhance the search process, synonyms complemented some of the keywords. For example, in place of bioacoustics, we also used terms such as animal or bird sounds. Table 1 summarizes the search terms used, the synonyms that complimented them, and the alternative terms used to refine the search.

Table 1. Literature search keywords.

Search Key	Acronyms	Search Refinement	Definition
Bioacoustics	Animals	birds, wildlife, pests	The branch of acoustics is concerned with sounds produced by or affecting living organisms, especially as relating to communication.
Non-Bioacoustics	Environment, artificial		Sounds are produced by artificial sources or both artificial and natural sources.
Sound	Noise		Vibrations that travel through the air or another medium, and can be heard when they reach a person's or animal's ear.
Classification	identification		The action or process of classifying something according to shared qualities or characteristics.
Technology	Sensors, Devices		Technology classification of sounds.
Machine Learning	Artificial Intelligence	CNN, SVM, Naïve Bayes	The use and development of computer systems that can learn and adapt without following explicit instructions, by using algorithms and statistical models to analyze and draw inferences from patterns in data.

We reviewed relevant articles published in the past 21 years (2000–2021). This time-frame was selected because practical machine learning techniques started gaining popularity during that time. Only papers written in the English language were included in the review process. The search criteria sought articles that involved sound classification and machine learning technology. Generic search terms (according to the thesaurus of each database) identified the relevant studies. The process of screening relevant studies used the inclusion and exclusion criteria tabulated in Table 2. The identification and elimination of duplicate studies followed. We categorized papers having the same titles or published by the same author on the same subject as duplicates. After the screening and duplicate elimination process, 124 (47 for environmental sound classification and 77 for bioacoustics sound classification) papers emerged as significant for the review.

Table 2. Literature exclusion and inclusion criteria.

Exclusion Criteria	Inclusion Criteria
Machine Learning Techniques Based on Images	bioacoustics classification.
Research not Published in English	General acoustic classification.
Research Published before 2000	Using machine learning technology
Sound classification in the medical sector that does not touch on technology	Peer-reviewed publications
Papers that were not considered original research, such as letters to the, editor comments, etc.	papers published in English

The initial search process yielded 166 articles (IEEE = 35, Elsevier = 23 Science Direct = 5, ACM Digital Library = 26, MDPI = 17 Springer = 20 and other = 40), with 101 articles for bioacoustics and 65 for general acoustics classification.

3.1.2. Screening for Inclusion

The Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) methodology [27] was used to screen relevant publications on acoustic classification for review. The PRISMA flow diagram in Figure 3 shows the number of papers identified, included, and excluded for the review and the databases used preferred Reporting Items for Systematic Reviews and Meta-Analyses. The identification and elimination of duplicate studies followed the search process. We categorized papers having duplicate titles or published by the same author on the same subject as duplicates. After excluding duplicated papers, 153 articles remained eligible for screening. The screening process resulted in the exclusion of 19 papers that were not in English and those published before 2000, when machine learning technology was still in its infancy. We further excluded ten papers that did not meet the inclusion criteria because they focused on the development of a dataset, monitoring sounds in the music industry or biologically, such as [28–30] through a full-text review of the articles.

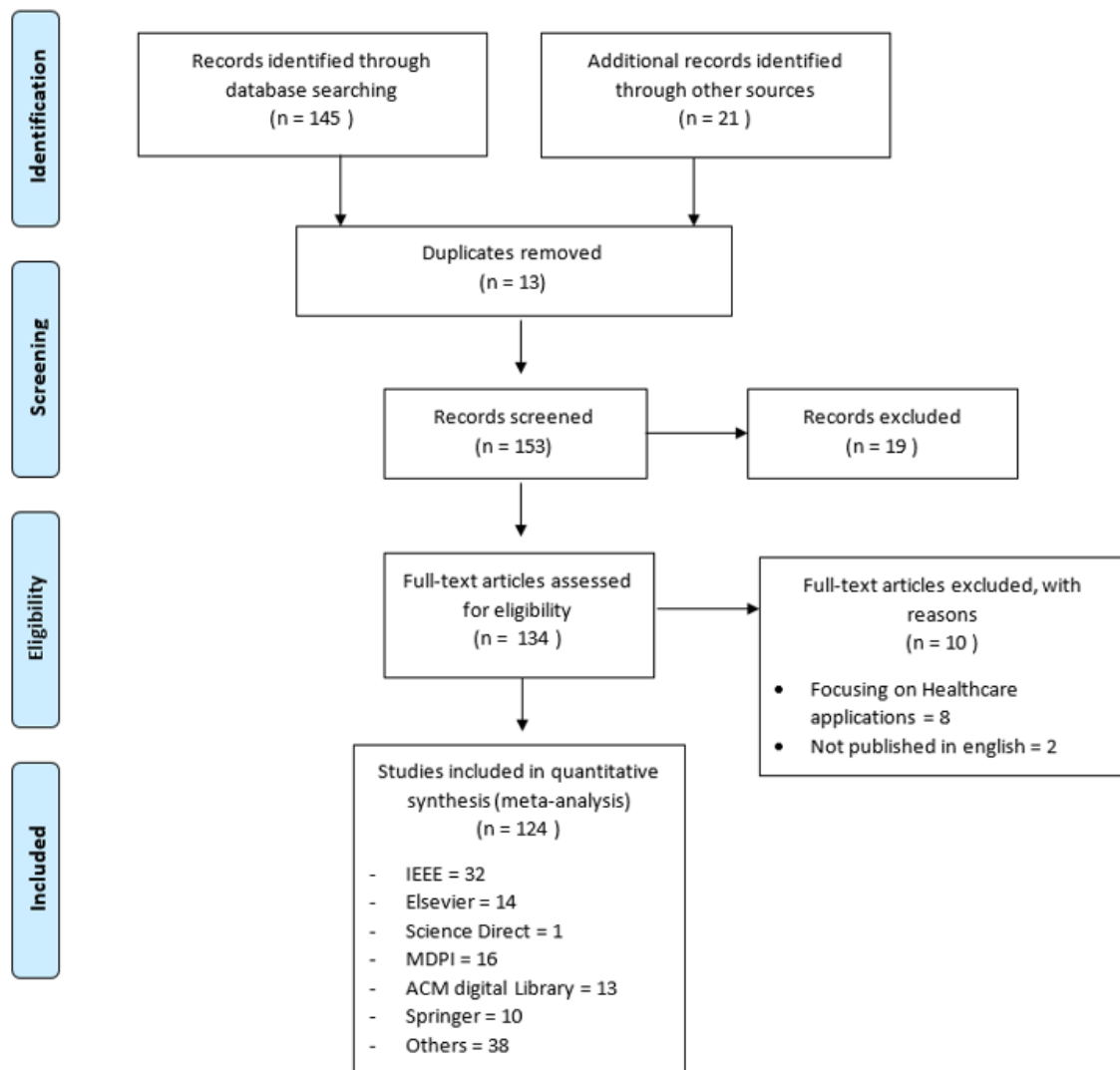


Figure 3. The study selection process.

From this process 124 papers (IEEE = 32, Elsevier = 14 Science Direct = 1, ACM Digital Library = 13, Springer = 10, MDPI = 16, and others = 38), emerged as significant for the final review. Most papers were retrieved from Computer Technology datasets such as Institute of Electrical and Electronics Engineers (IEEE) and the Association for Computing Machinery (ACM) for general acoustics papers. In contrast, bioacoustics papers were common in medical datasets PUBMED and multidisciplinary datasets such as MDPI, as shown in Figure 4. This unsurprising as bioacoustics classification integrates biology and technology disciplines while general acoustics classification focuses largely on technology-related disciplines. Our review used 77 papers representing bioacoustics classification and 47 papers representing general acoustic classification.

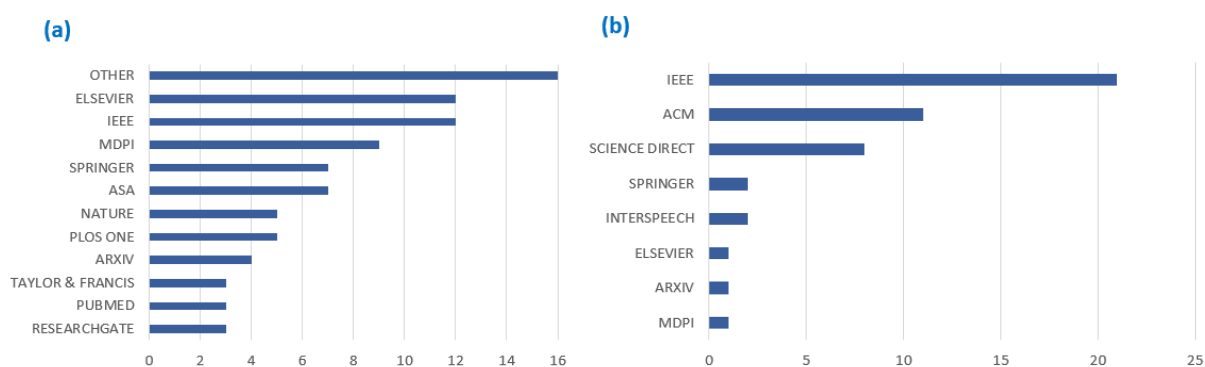


Figure 4. Databases used to retrieve (a) bioacoustics and (b) general acoustics classification papers.

3.1.3. Data Extraction

The screened articles were profiled next, in terms of keywords and year of publication to establish the nature and context of the research. The extracted data included: the year of publication, reference, publishers, algorithms used, datasets used, accuracy levels, application area, and the research contribution.

3.1.4. Data Analysis and Interpretation

Following the data extraction stage, research challenges (gaps) were identified using quantitative and qualitative descriptive techniques. Quantitative techniques involved numeric tabulation of observations from the reviews such as the number of datasets or machine learning techniques used in different studies. Qualitative techniques involved description of observations using words such as the limitations identified by previous studies. For example, some studies indicated that there was limited research in tropical geographic areas. These narrations were used to identify and describe the gaps. The results were collated and summarized. The analysis conducted on application areas of bioacoustics versus general acoustics studies provided insights on how research goals between the two areas differed. Additionally, a comparative analysis of acoustics technology revealed how these technologies differ across different application areas. The pre-processing techniques, datasets used, and machine learning algorithms adopted by different studies were tabulated for bioacoustics studies and compared to those used by general acoustic studies. The similarities and differences were documents and used to draw conclusions on preferences for different types of studies. The results of the analysis and interpretations are discussed in the next section.

3.1.5. Publication Demographics

For purposes of this survey, we classified acoustic sound classification publications into two broad categories; those that focused on bioacoustics (where the sound originated from living organisms in the animal kingdom) and general acoustics (where sounds originated from outside the animal kingdom). The word cloud generated from the publication keywords illustrates the relevance of the selected papers. Studies on bioacoustics focused

mainly on classifying animal sounds such as birds, insects, and whales, while those concerned with general acoustics were mostly environmental sound signals and not specific to particular species. We confined the scope of the survey to studies that used machine-learning algorithms for sound classification. It is worth noting that several studies also used image recognition techniques classify animals [31–38]. Those studies fell outside the scope of this review.

The survey revealed that both categories have received differing attention, with 62.0% of current acoustics classification research focused on bioacoustics and 38.0% on general acoustics, as illustrated in Table 3. This might be explained by the fact that research in bioacoustics classification started earlier than general acoustics research, with bioacoustics research picking up from 2009, as shown in Figure 5, while general acoustics research picked up from 2013. In both cases, the research output has steadily grown. However, the growth of acoustics classification in biology domains has been broader and faster than in technology domains.

Table 3. Literature exclusion and inclusion criteria.

Bioacoustics Research		General Acoustic Research	
Citations [14,15,39–107]	Number 77 (62.0%)	Citations [108–152]	Number 47 (38.0%)

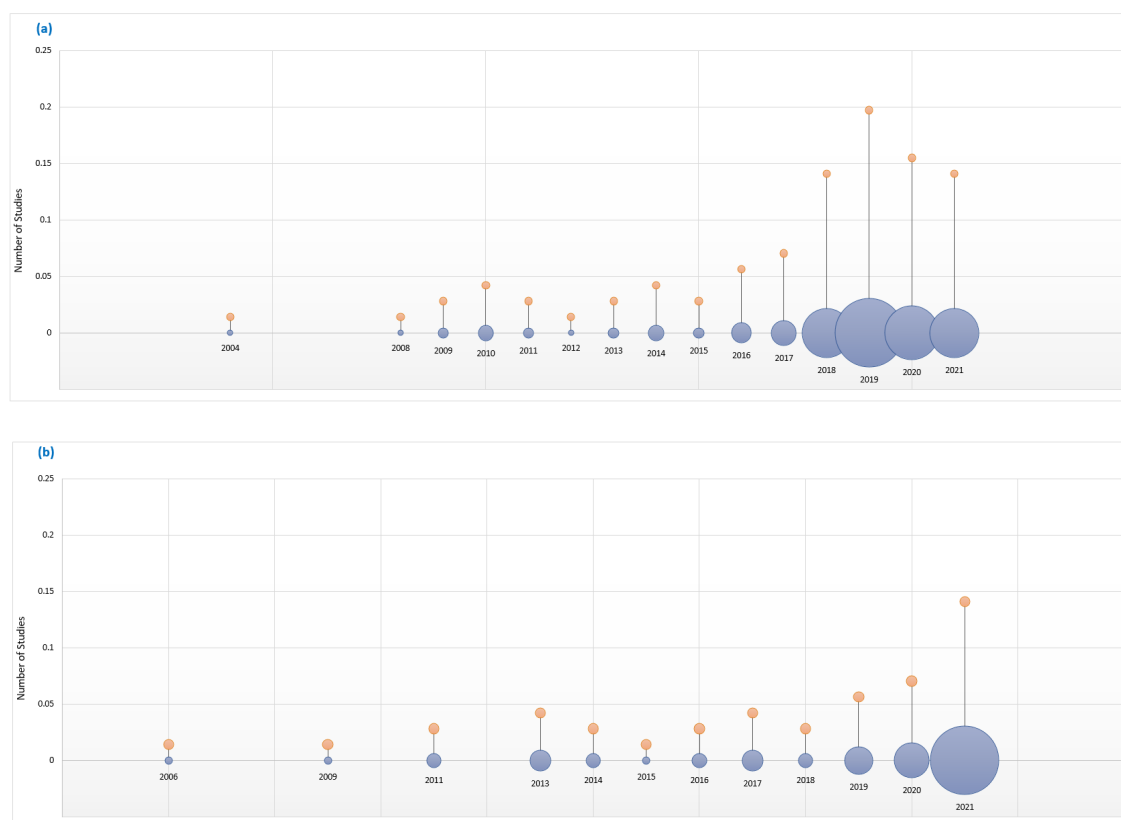


Figure 5. Progress of (a) bioacoustics and (b) general acoustics research output over the years.

4. Results

4.1. Application Areas

The survey shows that bioacoustics classification has found application in various botany and zoology fields such as: in conserving species [42–44,46]; monitoring of inter-species interaction [39,41,49,59,66,69]; understanding animal behavior [56,64,81]; agriculture in pest control [72,74]; and health in detecting sleep disorders [73]. General acoustics

classification has applications in: hearing aids [108,109,113,122,124,153]; analyzing machine-learning algorithms [110–112,114,154]; or for detecting the sources of sounds [155,156]. Monitoring of species formed the largest bioacoustics application area (84.2%), as shown in Figure 6. Most general acoustics research focused on technology improvement by evaluating machine learning algorithms (38.1%) and detecting the source of the sound through acoustic monitoring (33.3%), as illustrated in Figure 6. A few studies classified environmental sounds to support users with hearing impairments (19%).

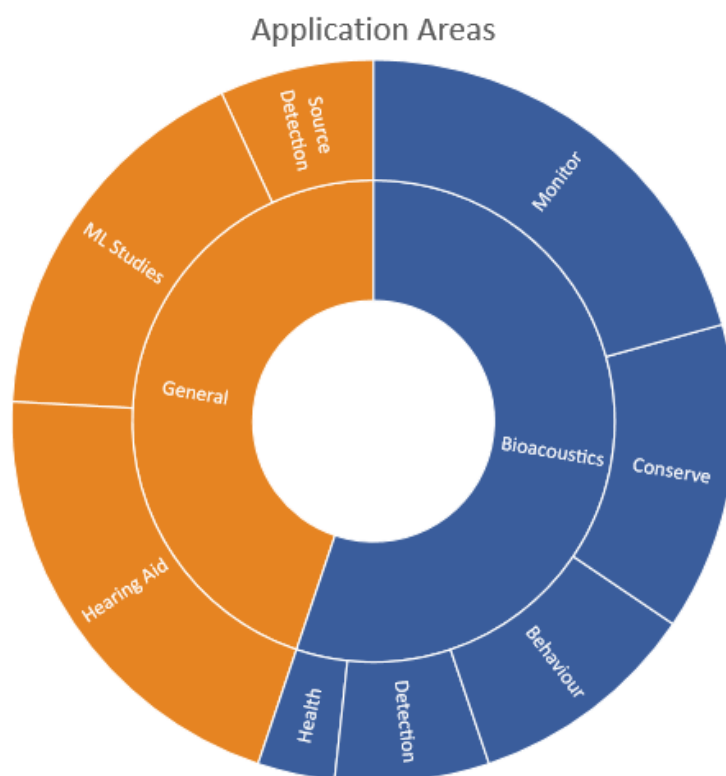


Figure 6. Application areas of research in bioacoustics and general acoustics classifications.

Most bioacoustics originated from animal vocals (74%) such as frogs croaking [40–42] or birds chirping [9,18,50,58,61] while a few originated from their locomotion (24%) such as bees [56,81,82] or mosquitoes [59] in flight as shown in Figure 7. Insects produce locomotion sounds in five different ways: stridulation, percussion, vibration, tymbal mechanism, or air expulsion [14]. Sounds originating from locomotion are low and sometimes not humanly audible thus, some studies have focused on image recognition to identify insects such as moths [28,33,34,36,38], which can be challenging if the insect is not within the field of vision. It is worth noting that some studies used both image and acoustic classification to classify bird sounds and observed that fusing these approaches achieved the better classification performance [79] compared to individual techniques. Some researchers have also noted that including features that provide visual-based discrimination, extending beyond the bio-acoustically, relevant parameters may offer improved performance [88].

Most publications surveyed (94%) dealt only with acoustic classification for humanly audible sounds. A similar observation was made for general acoustics sounds, where most studies focused on humanly audible sounds such as sounds made by a helicopter, chainsaw, or rain. Limited research existed for non-human inaudible sounds, as seen in Figure 7. This makes it difficult to assess the effectiveness of sound classification techniques for sounds that are not human audible from past studies. Thus, the acoustics research studies reviewed here are biased toward humanly audible sounds.

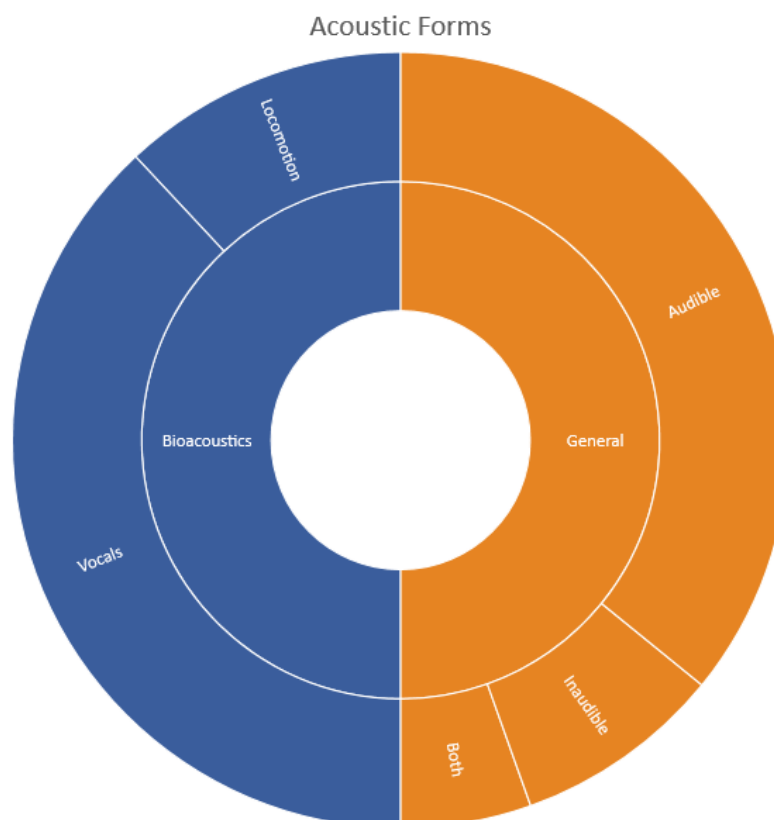


Figure 7. Forms of sound for bioacoustics and general acoustics studies.

It is worth noting that the general acoustics research reviewed here was concerned with sounds from both artificial sources, such as car alarms, gunshots, and construction equipment and natural sources, such as rain or animal sounds. Most of the studies examined [112,115,119,144,146] used the two types sound interchangeably, making it difficult to analyze general acoustic techniques exclusively on non-bioacoustics sounds. Establishing whether classification techniques differed for bioacoustics and non-bioacoustics techniques would provide better insight into the factors that influence the choice of classification techniques.

4.2. Techniques Used

Acoustic studies need datasets for training sound classification models. Most of the datasets used for bioacoustics classification are created by the researchers specifically for the study, as shown in Figure 8 [42,45,55,64,72,74]. This was common where publicly available datasets were unavailable. Datasets on insects were few, with the majority having sounds for birds, frogs, cats, whales, and dogs. For general acoustics, the most popular dataset was the US8K (Urban Sound 8K) which contains 8732 labeled sound excerpts of urban sounds [119,127,131,132,138,141,146,149] as shown in Figure 7. The ESC 50 and ESC 10 datasets were also among the popular datasets [119,130–132,141,143,144,146–149]. They contain a mixture of bioacoustics and general acoustics sounds. Most of the past general acoustics studies focused on a mixture of both bioacoustics and non-bioacoustics sounds. Therefore, targeted research is required to examine specific general acoustics based on their application areas.

Audio datasets present several challenges that influence the accuracy of the results obtained. For example, many real-world acoustic analysis problems are characterized by low signal-to-noise ratios and compounded by scarce data [59]. Another challenge is that most large-scale bioacoustics archives contain only a small percentage of animal vocalizations and a large amount of environmental noise, which makes it extremely difficult to retrieve

sufficient vocalizations for extended analysis [47]. The majority of the bioacoustics datasets examined had sounds exclusive to certain animal species, rendering them inappropriate for categorizing other different animal species [46]. Several studies also noted that [18] the species belong to specific geographic locations restricting the applications of the datasets. Beehives, for example, are found in various geographic locations with different acoustic backgrounds, and tests should represent each type of background [56]. Typically, locomotion sound falls under two behavioral contexts: (i) sonication (e.g., bees vibrating tomato flowers); and (ii) flight (e.g., bees between tomato flowers). The flight and sonication sound present pronounced differences in acoustic characteristics [82], which should be factored in during classification. A deeper experimental evaluation across multiple datasets is also required to improve the classification performance [107]. These datasets also do not factor in the animal age. Hence another challenge for the classifiers is to discriminate between species regardless of the age or stance [53].

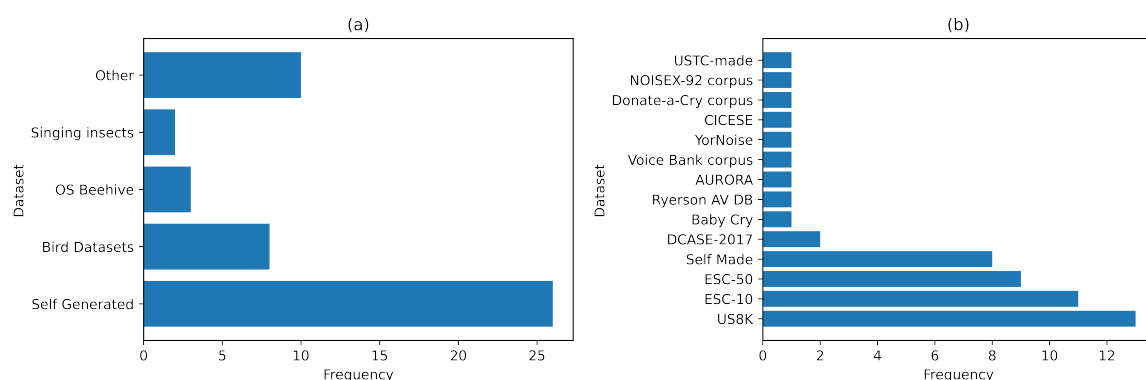


Figure 8. Data sets used for (a) bioacoustics and (b) general acoustics studies.

Our survey examined the impact of dataset size and classes on the accuracy obtained from acoustic classification. To achieve this, we assumed that all classes have the same number of instances; hence, we obtained an average of the instances per class. For bioacoustics, the results showed that higher accuracy levels occurred where fewer data (instances) existed, such as using the Cat Sound and Open-Source Beehive project datasets, as shown in Table 4. The number of classes also appeared to impact the accuracy, given that higher accuracy levels occurred where higher instance class ratios existed, as illustrated in Figure 8a.

Table 4. Bioacoustics dataset size and classification accuracy.

Dataset	Classes	Instances	Ratio	Average Accuracy
Cat Sound	2	440	220.00	91.13
Birdvox70k—CLO43SD	43	5428	126.20	90.00
Open Source Beehive Project	2	78	39.00	89.33
BIRDZ	50	602,512	12,050.20	89.04
Humboldt-University Animal Sound Archive	2530	120,000	47.40	81.30
MFCC dataset	10	7195	719.50	78.40
Zoological Sound Library	10,000	240,000	24.00	73.04
NIPS4Bplus	87	687	7.90	65.00

For general acoustics, the results showed that higher accuracy levels were obtained where fewer data (instances) existed, such as using the ESC-10 and DCASE datasets, as shown in Table 5. This is similar to the observations made for bioacoustics. However, the higher the number of classes, the higher the accuracy levels obtained, given that higher accuracy levels occurred where lower instance class ratios existed, as illustrated in Figure 9b. While these results point towards the number of classes having opposite impacts on the results' accuracy, it is difficult to verify them conclusively because existing

studies used only a single dataset. Most studies investigated how the type of algorithm influences the accuracy of the classification process. More research is required to investigate how other factors, such as the size or type of dataset, influence the accuracy of the classification process.

Table 5. General acoustics dataset size and classification accuracy.

Dataset	Classes	Instances	Ratio	Average Accuracy
ESC-10	10	400	40.00	90.66
DCASE	16	320	20.00	86.70
US8K	10	8732	873.20	83.67
ESC-50	50	2000	40.00	81.54
Ryerson AV DB	8	7356	919.50	71.30
CICESE	20	1367	68.35	68.10

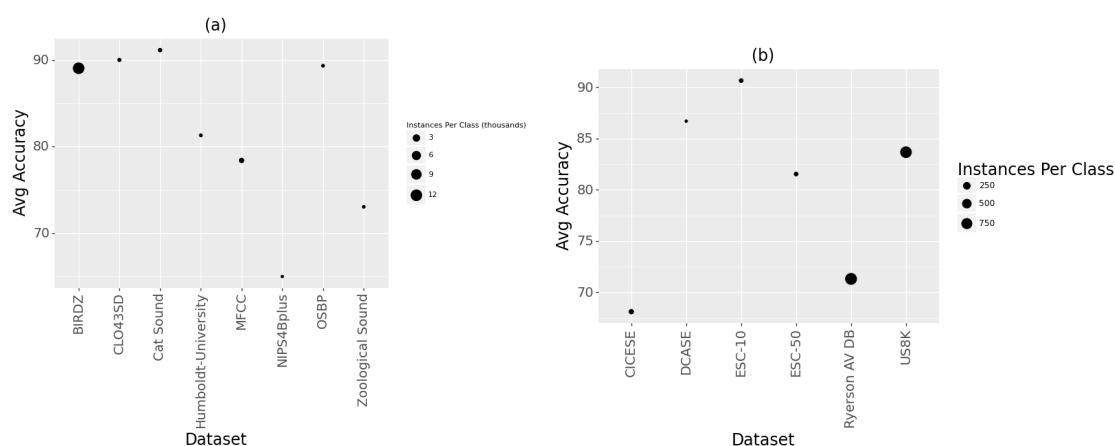


Figure 9. Impact of dataset size on classification accuracy for (a) bioacoustics and (b) general acoustics studies.

4.2.1. Data Preprocessing

After collecting audio data, they need to undergo preprocessing techniques that clean and transform them for classification. Most bioacoustics [48–51,53,54,61–64,68–75,81,82,86], and general acoustic [113–115,117,118,120,121,123,124,126,132,135,137–147,149,153,154] studies did not describe the preprocessing techniques that they used. An analysis of the studies that mentioned preprocessing revealed the most popular audio transformation technique as STFT (short-time Fourier transform) among both the bioacoustics [52,60,65,83] and general acoustic [111,119,130] studies (Figure 10). STFT is a powerful general-purpose tool for audio signal preprocessing [157–159] where a signal is broken into several signals of shorter duration and then transformed into frequency domains. The other popular technique mentioned was constant-Q transform (CQT) which was used in both bioacoustics [79] and general acoustic studies [148]. It transforms a data series into a frequency domain. The fast Fourier transform (FTT) was popular in bioacoustics studies [47,67]. It expands signals in terms of sinusoids. Both bioacoustics and general acoustic studies employed segmentation [14,39–43,46,83,84,110] to distinguish the sound in question from other sounds such as speech, music, environmental sounds, silence, and combinations of these sounds by automatically revealing semantically meaningful temporal segments in an audio signal [160].

Feature extraction helps derive the audio's short-time energy, zero-crossing rate, and bandwidth, among other useful features when classifying sound. It reduces the dimension of an audio input vector while retaining the important discriminating feature of the audio. This study revealed that the most popular feature extraction technique uses the cepstral coefficient, as illustrated in Figure 11. Mel frequency cepstral coefficients (MFCCs) use the MEL scale to divide the frequency band into sub-bands and then extract the Cepstral

Coefficients using a discrete cosine transform (DCT). The MEL scale is based on how humans distinguish between frequencies, making it a very effective approach for processing sounds. Before the introduction of MFCCs, linear prediction coefficients (LPCs) and linear prediction cepstral coefficients (LPCCs) were the primary feature type for automatic speech recognition, especially with hidden Markov model (HMM) classifiers.

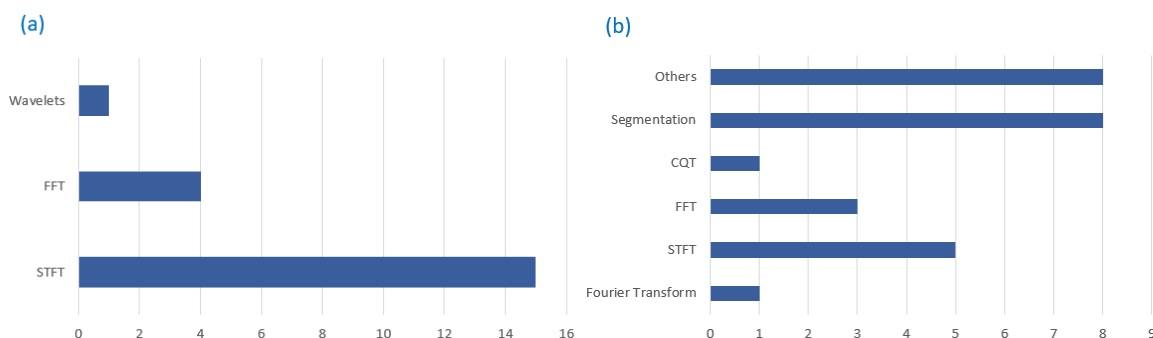


Figure 10. Audio pre-processing techniques used in (a) bioacoustics and (b) general acoustics studies.

The review observed that MFCCs was popular among bioacoustics studies [39,40,43,44,49,53,61,73,81–84,86] and general acoustic studies [112,116,125,133,141,148]. Linear frequency cepstral coefficients (LFCC) were popular among general acoustics studies [83,109,110,112] but found fewer applications in bioacoustics studies [14]. Few studies used LPCC [74,127] although it was used in both bioacoustics and general acoustics studies.

Feature Extraction techniques used in Bioacoustics and general acoustics studies

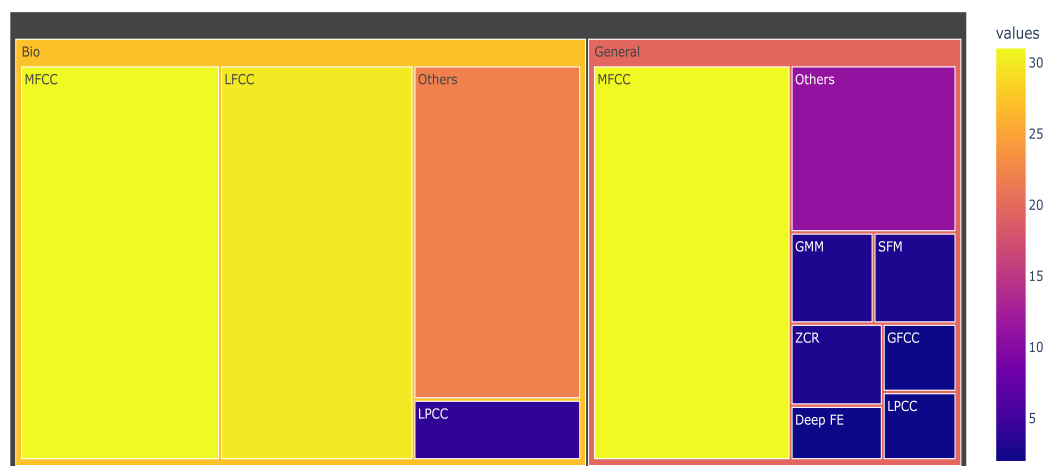


Figure 11. Feature extraction techniques used in bioacoustics and general acoustics studies.

4.2.2. Machine Learning Algorithms

Audio, sound, or acoustics classification is the process of analyzing audio recordings to identify their origin, type, or environment. The process is often automated using machine learning classification algorithms. Our survey showed that ensemble approaches are the most popular machine learning algorithms used in bioacoustics classification [39,40,43–45,48,50,51,53,56,76,77,79,81–84,86]. Convolutional neural networks (CNN) were the most popular algorithms for general acoustic classifications [113,114,119,121,125,133,136,137,139,141,144,146,148] as seen in Figure 12. The choice of particular classifiers was motivated by the performance of similar classification tasks from previous studies [110,111] or from experiments conducted to identify the most accurate algorithm [113,114]. Some studies did not specify the type of neural network they used; hence we classified them as DNN (Deep Neural Networks) [81,115,124,131,138,154].

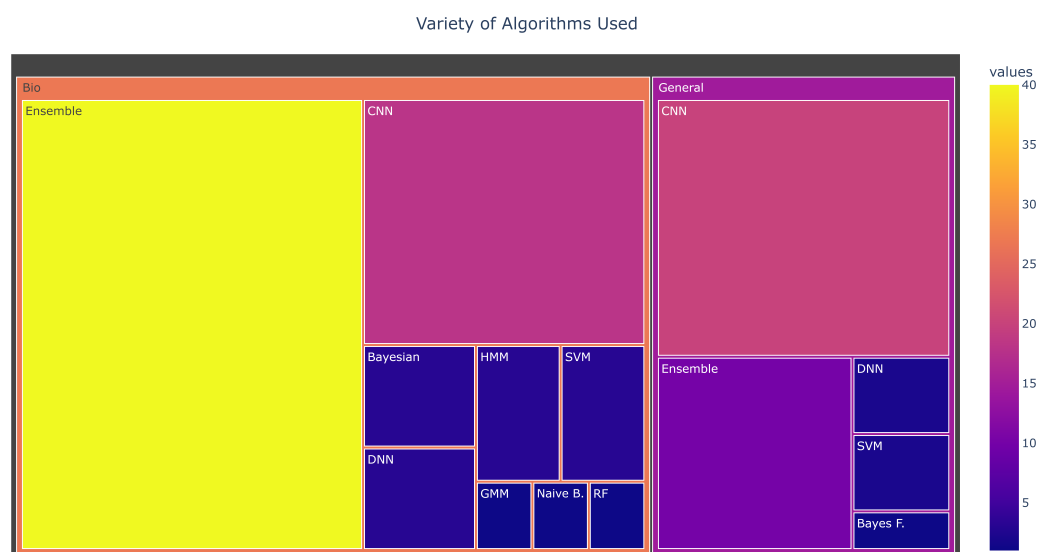


Figure 12. Classification algorithms used for bioacoustics and general acoustics studies.

Bayesian [58] and hidden Markov models [47] showed the best accuracy levels (based on the figures provided by the authors of these studies) for bioacoustics sounds. However, only a few studies used them, as seen in Figure 13a, due to (1) their high computational cost and (2) greater statistical expertise required than some other methods. This makes it difficult to generalize their efficacy. CNN algorithms and ensemble approaches were more popular; however, they had slightly lower accuracy (87–88%). Ensemble approaches showed better accuracy for classifying general acoustics than approaches based on CNN. However, only a few studies used them (Figure 13a). The SVM algorithm gave very high accuracy levels (84.5%), but was used only in a few studies [106,107], which makes it difficult to generalize. These results also show that CNN (at 88%) algorithms perform marginally better than ensemble (at 87%) approaches in bioacoustics studies. However, despite their popularity, they perform poorly (at 82%) in general acoustics studies compared to ensemble approaches (at 83.6%). Therefore, in general, in acoustic studies, ensemble approaches work better. Ensemble approaches also seem to be better at detecting some animal vocalizations, which might explain their accuracy [72]. For example, it has been shown that certain frog species are easily recognized by specific algorithms [71].

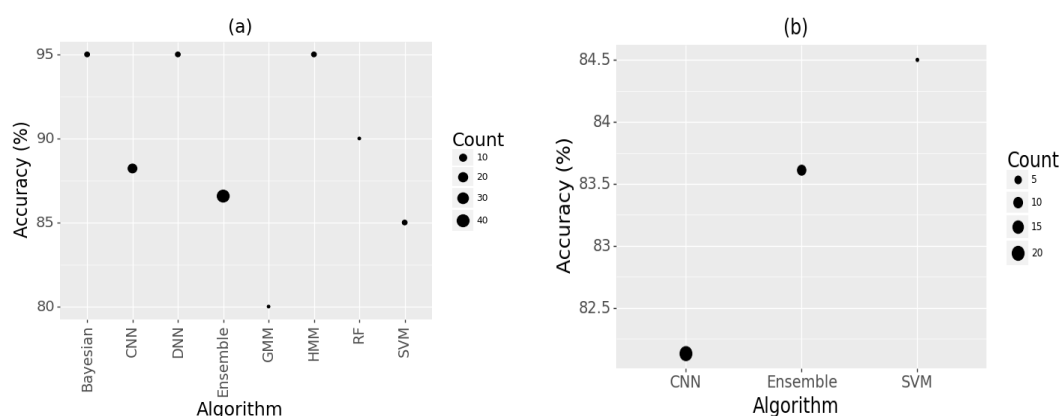


Figure 13. Classification algorithms used for (a) bioacoustics and (b) general acoustics studies.

Although more accurate, CNN demands large amounts of labeled raw acoustic data [68]. Learning directly from the raw waveform allows the algorithm to automatically select those elements of the sound that are best suited for the task, bypassing the onerous task of selecting feature extraction techniques and reducing possible biases [58].

However, due to the limited datasets available, solutions that yield effective classification results, even when only a small number of per-class training examples are available, should be explored [63]. For example, Ref. [64] proposes a deep learning approach that computes a perceptual embedding of animal vocalizations based on similarity judgments instead of class-specific labels. Similarly, a different study [80] combined transfer learning of a pre-trained deep convolutional neural network (CNN) model and a semi-supervised pseudo-labeling method with a custom loss function to address this challenge. They employ techniques to deal with the lack of class-labeled data, such as transfer learning from a (Multi-Dimensional Scaling) MDS space, attention pooling, and dynamic triplet loss. Combined with the ensemble approach, such techniques have produced better accuracy results [75].

Most acoustic studies did not address resource utilization as part of the algorithm's efficiency in terms of power and space. Hence, these approaches are unsuitable for real-time resource-constrained applications [76]. Most acoustic presentation approaches require extracting a large set of features, which consumes additional storage, processing, and communication resources.

The application areas and sources of sound can shed light on the preferred choice of classification techniques to establish the adequacy of an algorithm for a given role. The analysis results shown in Figure 14 reveal that CNN algorithms were predominantly used in general acoustics, where the research investigated ways of enhancing the classification algorithms or detecting the source of the sound. Support Vector Machine (SVM) approaches were also popular for detecting the source of sounds. Other roles, such as speech analysis and video captioning, preferred ensemble approaches. In bioacoustics studies, CNN and ensemble approaches were popular for all roles. However, some algorithms, such as Bayesian approaches were used in species detection.

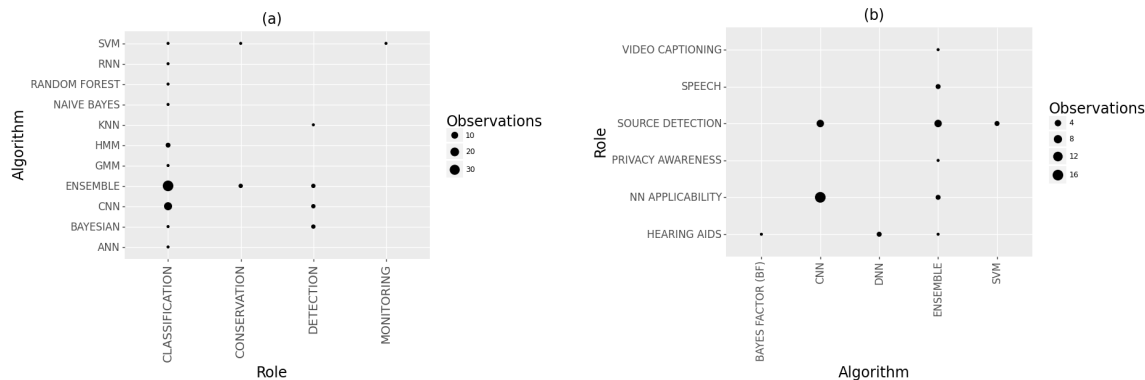


Figure 14. Algorithms used for different acoustic roles in (a) bioacoustics and (b) general acoustics studies.

Both CNN and Ensemble approaches were used to classify natural and artificial sound sources in general acoustic classifications, as shown in Figure 15b. No specific algorithm for natural sound classification was preferred, although such studies avoided CNN and SVM. Studies that investigated bioacoustics preferred CNN and ensemble approaches for analyzing locomotion. However, studies that analyzed vocals also used other algorithms, such as SVM and HMM (Figure 15a). Bayesian approaches were also preferred for analyzing locomotion.

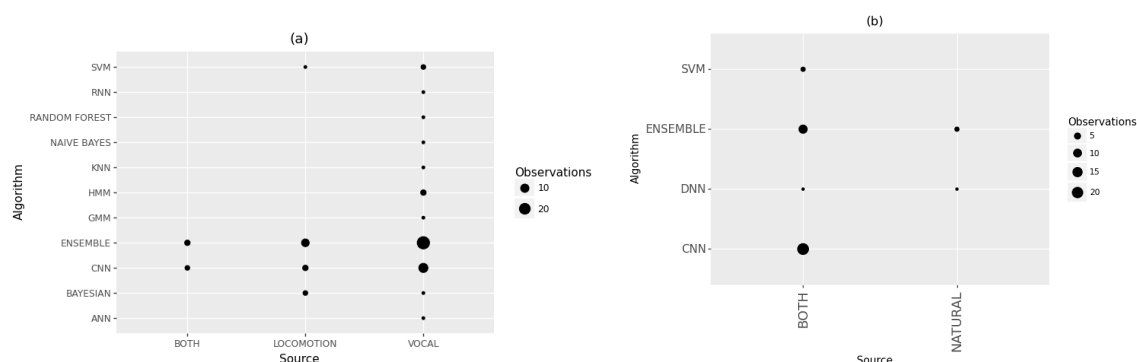


Figure 15. Algorithms used for different sources of sound in (a) bioacoustics and (b) general acoustics studies.

4.2.3. Overtones in Acoustic Techniques

Using the R Statistical analysis tool, we used the Cramer's V method to measure the strength of associations between preprocessing and classification algorithms. The Cramer's V values for the association between classification algorithms and the preprocessing techniques were obtained as 0.443 and 0.3274 for bioacoustics and general acoustics studies, respectively. While both values indicated a strong association, this value was only statistically significant for bioacoustics studies where the Pearson's correlation coefficient was 0.0414 ($p < 0.05$), as illustrated in Figure 16.

```
> assocstats(bio)
              X^2 df  P(> X^2)
Likelihood Ratio 27.652 20 0.117905
Pearson          32.178 20 0.041439

Phi-Coefficient   : NA
Contingency Coeff.: 0.663
Cramer's V        : 0.443
```

Figure 16. Cramer's V association test for bioacoustics and general acoustics studies.

Based on these findings, we identified the specific associations for bioacoustics studies using mosaic plots. The blue cells in Figure 17 contribute to the significance of the test of independence, therefore, demonstrating an association between artificial neural networks algorithms for classification and STFT techniques for audio transformation. Similarly, Gaussian mixture model (GMM) classification approaches were strongly associated with LFCC audio transformation techniques. Future studies should seek to understand these associations further through a comparative analysis of different classification algorithms and preprocessing techniques.

To understand how areas of focus varied among the bioacoustics studies that we reviewed, we conducted a cluster analysis of the studies. A cluster analysis groups the observations based on common characteristics to derive further insights from the observations. The results showed that most studies focused on one or two areas. For instance, most studies that examined neural network classifiers did not specify either the audio transformation techniques used (Cluster 5) or the feature extraction techniques (Cluster 3), as shown in Figure 18. Similarly, most of the studies that used ensemble classifiers did not specify the audio transformation techniques, instead they explored either ensemble feature extraction approaches (Cluster 1) or MFCC feature extraction approaches (Cluster 2). Only four studies explored all techniques (Cluster 4). In addition, most studies in Cluster 4 used MFCC and fast Fourier transform (FFT) preprocessing techniques for ensemble classification approaches. It is unclear from our findings how the choice of preprocessing techniques influenced the selection of classification techniques. However, this type of information

could benefit other researchers in the field. It would therefore be useful, if studies described the techniques used across all the phases of their bioacoustics classification.

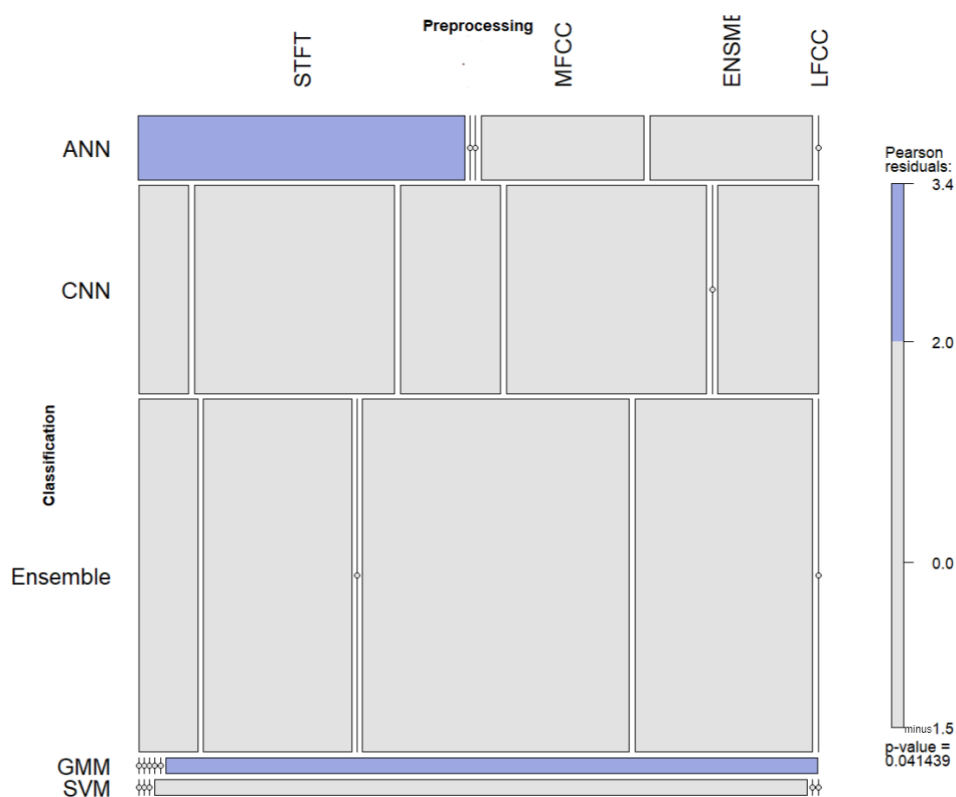


Figure 17. Associations between pre-processing and classification techniques for bioacoustics studies.

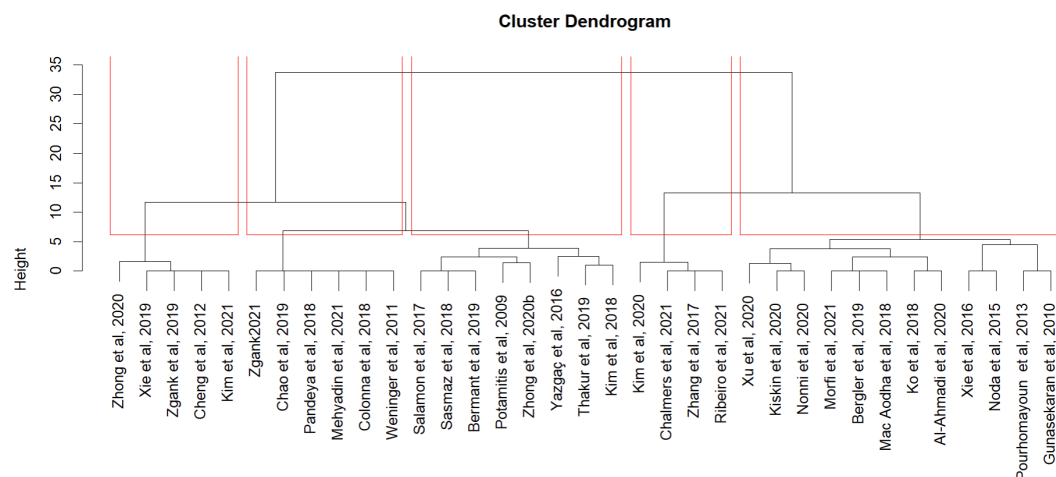


Figure 18. Focus Areas in bioacoustics studies.

4.3. Implementation and Evaluation

To understand how the applications identified in the survey were implemented, we examined the theoretical backgrounds, architectural designs, workflow descriptions, and the results presented. A comparison of bioacoustics and general acoustics studies revealed that in both cases, only half of the studies provided theoretical backgrounds or discussed architectural considerations. We attributed this to the fact that the studies prioritized the use of existing technology obtain results at the expense of other considerations. An interesting observation was the emphasis laid on the workflow description by general acoustic studies (85.1%). The results in Figure 19 show that most studies focused on presenting results

compared to providing implementation details. The ability to recreate results is a crucial aspect of evaluating the efficacy of any proposed solution, and future studies need to describe implementation as part of the research.

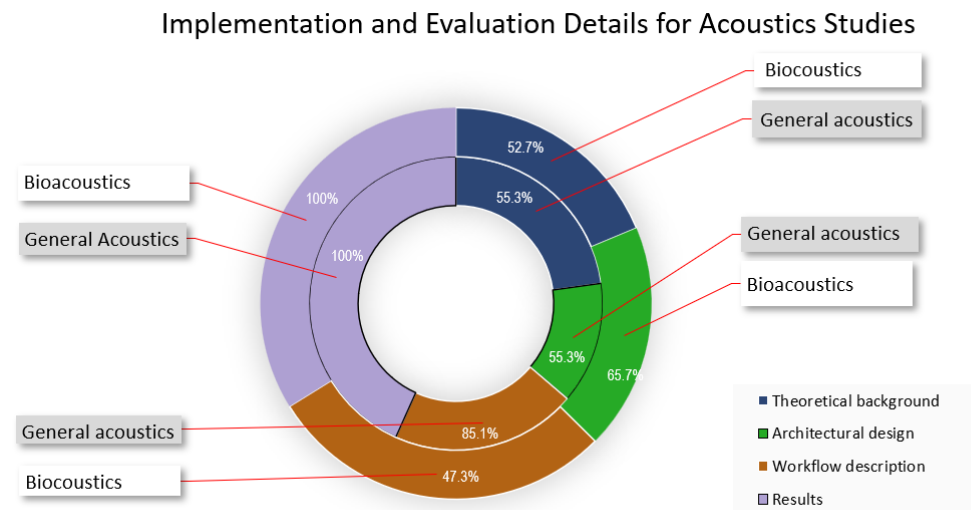


Figure 19. Implementation and evaluation details for acoustic studies.

5. Discussion and Open Questions

Our survey identified several open questions that might inform future research in bioacoustics. These research gaps are discussed next, and the emerging challenges and opportunities are summarized in Figure 20.

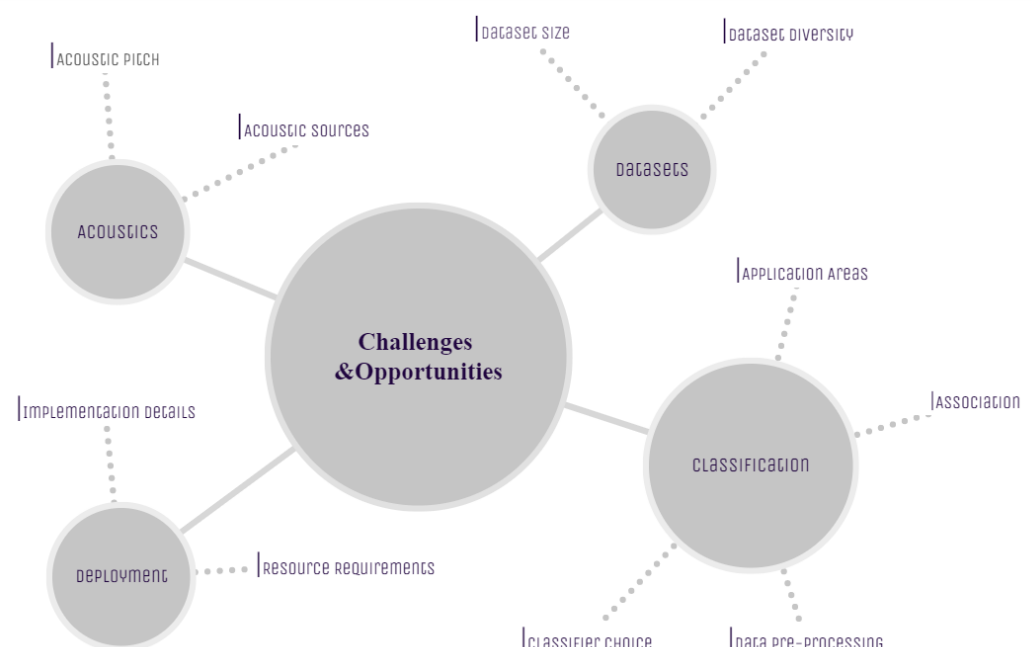


Figure 20. Summary of the challenges and opportunities.

5.1. Acoustics

The bioacoustics studies surveyed focused on sounds made vocally rather than through locomotion or other bodily movements. There is need for more research on classifying sounds generated through locomotion and bodily movements. Sonication and isolated motion present pronounced differences in acoustic characteristics, which should be

factored in during the classification. Both bioacoustics and general acoustic studies focused on humanly audible sounds, such as those made by frogs or birds, with limited research on less audible sounds made by insects such as moths.

5.2. Dataset

Most bioacoustics studies used datasets explicitly generated for the study. Publicly available datasets on insects, arachnids and arthropods were few. The majority of the datasets had sounds for birds, frogs, cats, whales, and dogs. More diverse datasets are needed to enhance research in this area. It is also useful that datasets include not just information on the species, but also geographic locations. Our survey examined the impact of dataset size and classes on the accuracy obtained from acoustic classification. However, it was difficult to verify the findings conclusively as existing studies used only a single dataset. Most studies investigated how the type of algorithm influences the accuracy of the classification process. More research is required to investigate how other factors, such as the size or type of dataset, influence the accuracy of the classification process. A deeper experimental evaluation across multiple datasets is required to enhance the classification performance. Existing datasets also do not factor in the age of the animal, gender, or season.

5.3. Classification

While bioacoustics applications in sound detection, species monitoring, and conservation are growing, the volume is still small. The current focus is mainly on classification. The most popular audio transformation and feature extraction techniques among bioacoustics studies were STFT (short-time Fourier transform) and MFCCs. However, few studies have investigated how these techniques' choices influenced the results' accuracy. Our survey observed that ensemble approaches were the most popular machine learning algorithms in bioacoustics classification; however, Bayesian and hidden Markov models presented higher accuracy levels. More research is needed on these techniques to generalize their efficacy. There is limited research on how the role or source of sound influence the effectiveness of selected algorithms. Additionally, there is limited understanding of the association between preprocessing techniques and the choice of classification algorithms.

5.4. Deployment

Most acoustic studies surveyed did not address resource utilization as part of the algorithm's efficiency in terms of processing power and memory space requirements. This makes it difficult to gauge their effectiveness for real-time resource-constrained applications. Most studies focused on presenting results compared to providing implementation details such as the theoretical background, architectural and workflow considerations. Further, most of the studies provided more information on feature extraction theoretical backgrounds compared to machine learning. The workflows presented focused more on machine learning compared to feature extraction phases. The ability to recreate results is a crucial aspect of evaluating the efficacy of any proposed solution, and future studies need to adequately describe feature extraction and machine learning implementation aspects as part of the research description.

Classification algorithms present challenges and opportunities for research in new application areas, preprocessing and selection. However, there is also need to investigate and create diverse bioacoustics sources and datasets.

6. Conclusions

This survey was a review of acoustic classification techniques based on their application areas to highlight the gaps in existing research on acoustic classification techniques. The results revealed the critical application areas as species classification, done using animal vocals. The popular audio transformation techniques are STFT, while the popular feature extraction techniques are MFCC. The most popular classification approaches are Ensemble and CNN machine learning algorithms. Studies that used ensemble approaches showed a

preference for MFCC feature extraction techniques and no specific audio transformation techniques. However, studies that used neural networks showed a preference for LFCC feature extraction techniques and STFT audio transformation techniques. The findings from the survey revealed that most studies focused on disseminating the results rather than implementation considerations. Finally, the study recommended a research agenda for bioacoustics classification techniques.

Author Contributions: Conceptualization, L.M. and G.K.; methodology, L.M. and G.K.; validation, L.M. and G.K.; formal analysis, L.M., G.K., J.G., K.G. and P.W.; investigation, L.M., G.K., J.G., K.G. and P.W.; resources, L.M., G.K. and J.G.; data, J.G., K.G. and P.W.; writing—original draft preparation, L.M., G.K. and J.G.; validation, L.M. and G.K.; writing—review and editing, L.M., G.K. and J.G.; supervision, L.M. and G.K.; project administration, L.M. All authors have read and agreed to the published version of the manuscript.

Funding: This research received on external funding.

Institutional Review Board Statement: This study is a review of published work on automated bioacoustics sensing, which does not require ethical review or approval.

Informed Consent Statement: Not applicable.

Data Availability Statement: <https://docs.google.com/spreadsheets/d/1Ca-okxXDCsxNt4blr8g1BDXqlxjV7Cv/edit#gid=1545484044>.

Conflicts of Interest: The authors declare no conflict of interest. There was external funding for this work.

References

1. Penar, W.; Magiera, A.; Kloczek, C. Applications of bioacoustics in animal ecology. *Ecol. Complex.* **2020**, *43*, 100847. [CrossRef]
2. Choi, Y.K.; Kim, K.M.; Jung, J.W.; Chun, S.Y.; Park, K.S. Acoustic intruder detection system for home security. *IEEE Trans. Consum. Electron.* **2005**, *51*, 130–138. [CrossRef]
3. Shah, S.K.; Tariq, Z.; Lee, Y. Iot based urban noise monitoring in deep learning using historical reports. In Proceedings of the 2019 IEEE International Conference on Big Data (Big Data), Los Angeles, CA, USA, 9–12 December 2019; pp. 4179–4184.
4. Vacher, M.; Serignat, J.F.; Chaillol, S.; Istrate, D.; Popescu, V. Speech and sound use in a remote monitoring system for health care. In *International Conference on Text, Speech and Dialogue*; Springer: Berlin/Heidelberg, Germany, 2006; pp. 711–718.
5. Olivieri, M.; Malvermi, R.; Pezzoli, M.; Zanoni, M.; Gonzalez, S.; Antonacci, F.; Sarti, A. Audio information retrieval and musical acoustics. *IEEE Instrum. Meas. Mag.* **2021**, *24*, 10–20. [CrossRef]
6. Schöner, M.G.; Simon, R.; Schöner, C.R. Acoustic communication in plant–animal interactions. *Curr. Opin. Plant Biol.* **2016**, *32*, 88–95. [CrossRef]
7. Obrist, M.K.; Pavan, G.; Sueur, J.; Riede, K.; Llusia, D.; Márquez, R. Bioacoustics approaches in biodiversity inventories. *Abc Taxa* **2010**, *8*, 68–99.
8. Chachada, S.; Kuo, C.C.J. Environmental sound recognition: A survey. *APSIPA Trans. Signal Inf. Process.* **2014**, *3*, 14015991. [CrossRef]
9. Kvsn, R.R.; Montgomery, J.; Garg, S.; Charleston, M. Bioacoustics data analysis—A taxonomy, survey and open challenges. *IEEE Access* **2020**, *8*, 57684–57708.
10. Mcloughlin, M.P.; Stewart, R.; McElligott, A.G. Automated bioacoustics: Methods in ecology and conservation and their potential for animal welfare monitoring. *J. R. Soc. Interface* **2019**, *16*, 20190225. [CrossRef]
11. Walters, C.L.; Collen, A.; Lucas, T.; Mroz, K.; Sayer, C.A.; Jones, K.E. Challenges of using bioacoustics to globally monitor bats. In *Bat Evolution, Ecology, and Conservation*; Springer: New York, NY, USA, 2013; pp. 479–499.
12. Xie, J.; Colonna, J.G.; Zhang, J. Bioacoustic signal denoising: A review. *Artif. Intell. Rev.* **2021**, *54*, 3575–3597. [CrossRef]
13. Chen, W.; Sun, Q.; Chen, X.; Xie, G.; Wu, H.; Xu, C. Deep learning methods for heart sounds classification: A systematic review. *Entropy* **2021**, *23*, 667. [CrossRef]
14. Potamitis, I.; Ganchev, T.; Kontodimas, D. On automatic bioacoustic detection of pests: The cases of *Rhynchophorus ferrugineus* and *Sitophilus oryzae*. *J. Econ. Entomol.* **2009**, *102*, 1681–1690. [CrossRef] [PubMed]
15. Stowell, D.; Plumbly, M.D. Automatic large-scale classification of bird sounds is strongly improved by unsupervised feature learning. *PeerJ* **2014**, *2*, e488. [CrossRef] [PubMed]
16. Bonet-Solà, D.; Alsina-Pagès, R.M. A comparative survey of feature extraction and machine learning methods in diverse acoustic environments. *Sensors* **2021**, *21*, 1274. [CrossRef]
17. Lima, M.C.F.; de Almeida Leandro, M.E.D.; Valero, C.; Coronel, L.C.P.; Bazzo, C.O.G. Automatic detection and monitoring of insect pests—A review. *Agriculture* **2020**, *10*, 161. [CrossRef]

18. Stowell, D.; Wood, M.; Stylianou, Y.; Glotin, H. Bird detection in audio: A survey and a challenge. In Proceedings of the 2016 IEEE 26th International Workshop on Machine Learning for Signal Processing (MLSP), Vietri sul Mare, Italy, 13–16 September 2016; pp. 1–6.
19. Qian, K.; Janott, C.; Schmitt, M.; Zhang, Z.; Heiser, C.; Hemmert, W.; Yamamoto, Y.; Schuller, B.W. Can machine learning assist locating the excitation of snore sound? A review. *IEEE J. Biomed. Health Inform.* **2020**, *25*, 1233–1246. [\[CrossRef\]](#)
20. Bhattacharya, S.; Das, N.; Sahu, S.; Mondal, A.; Borah, S. Deep classification of sound: A concise review. In *Proceeding of First Doctoral Symposium on Natural Computing Research*; Springer: Singapore, 2021; pp. 33–43.
21. Bencharif, B.A.E.; Ölçer, I.; Özkan, E.; Cesur, B. Detection of acoustic signals from Distributed Acoustic Sensor data with Random Matrix Theory and their classification using Machine Learning. In *SPIE Future Sensing Technologies*; SPIE: Anaheim, CA, USA, 2020; Volume 11525, pp. 389–395.
22. Sharma, G.; Umapathy, K.; Krishnan, S. Trends in audio signal feature extraction methods. *Appl. Acoust.* **2020**, *158*, 107020. [\[CrossRef\]](#)
23. Piczak, K.J. ESC: Dataset for environmental sound classification. In Proceedings of the 23rd ACM International Conference on Multimedia, Brisbane, Australia, 13 October 2015; pp. 1015–1018.
24. Gharib, S.; Derrar, H.; Niizumi, D.; Senttula, T.; Tommola, J.; Heittola, T.; Virtanen, T.; Huttunen, H. Acoustic scene classification: A competition review. In Proceedings of the 2018 IEEE 28th International Workshop on Machine Learning for Signal Processing (MLSP), Aalborg, Denmark, 17–20 September 2018; pp. 1–6.
25. Schryen, G.; Wagner, G.; Benlian, A.; Paré, G. A knowledge development perspective on literature reviews: Validation of a new typology in the IS field. *Commun. AIS* **2020**, *46*, 134–186. [\[CrossRef\]](#)
26. Templier, M.; Pare, G. Transparency in literature reviews: An assessment of reporting practices across review types and genres in top IS journals. *Eur. J. Inf. Syst.* **2018**, *27*, 503–550. [\[CrossRef\]](#)
27. Moher, D.; Shamseer, L.; Clarke, M.; Ghersi, D.; Liberati, A.; Petticrew, M.; Shekelle, P.; Stewart, L.A. Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015 statement. *Syst. Rev.* **2015**, *4*, 1. [\[CrossRef\]](#)
28. Barber, J.R.; Plotkin, D.; Rubin, J.J.; Homziak, N.T.; Leavell, B.C.; Houlihan, P.R.; Miner, K.A.; Breinholt, J.W.; Quirk-Royal, B.; Padrón, P.S.; et al. Anti-bat ultrasound production in moths is globally and phylogenetically widespread. *Proc. Natl. Acad. Sci. USA* **2022**, *119*, e2117485119. [\[CrossRef\]](#)
29. Bahuleyan, H. Music genre classification using machine learning techniques. *arXiv* **2018**, arXiv:1804.01149.
30. Sim, J.Y.; Noh, H.W.; Goo, W.; Kim, N.; Chae, S.H.; Ahn, C.G. Identity recognition based on bioacoustics of human body. *IEEE Trans. Cybern.* **2019**, *51*, 2761–2772. [\[CrossRef\]](#) [\[PubMed\]](#)
31. Bisgin, H.; Bera, T.; Ding, H.; Semey, H.G.; Wu, L.; Liu, Z.; Barnes, A.E.; Langley, D.A.; Pava-Ripoll, M.; Vyas, H.J.; et al. Comparing SVM and ANN based machine learning methods for species identification of food contaminating beetles. *Sci. Rep.* **2018**, *8*, 6532. [\[CrossRef\]](#) [\[PubMed\]](#)
32. Høye, T.T.; Årje, J.; Bjerger, K.; Hansen, O.L.; Iosifidis, A.; Leese, F.; Mann, H.M.; Meissner, K.; Melvad, C.; Raitoharju, J. Deep learning and computer vision will transform entomology. *Proc. Natl. Acad. Sci. USA* **2021**, *118*, e2002545117. [\[CrossRef\]](#) [\[PubMed\]](#)
33. Shankar, K.; Perumal, E.; Vidhyavathi, R. Deep neural network with moth search optimization algorithm based detection and classification of diabetic retinopathy images. *SN Appl. Sci.* **2020**, *2*, 748. [\[CrossRef\]](#)
34. Bjerger, K.; Nielsen, J.B.; Sepstrup, M.V.; Helsing-Nielsen, F.; Høye, T.T. An automated light trap to monitor moths (Lepidoptera) using computer vision-based tracking and deep learning. *Sensors* **2021**, *21*, 343. [\[CrossRef\]](#)
35. Valletta, J.J.; Torney, C.; Kings, M.; Thornton, A.; Madden, J. Applications of machine learning in animal behaviour studies. *Anim. Behav.* **2017**, *124*, 203–220. [\[CrossRef\]](#)
36. Feng, L.; Bhanu, B.; Heraty, J. A software system for automated identification and retrieval of moth images based on wing attributes. *Pattern Recognit.* **2016**, *51*, 225–241. [\[CrossRef\]](#)
37. Vasconcelos, D.; Nunes, N.J.; Gomes, J. An annotated dataset of bioacoustic sensing and features of mosquitoes. *Sci. Data* **2020**, *7*, 382. [\[CrossRef\]](#)
38. Mayo, M.; Watson, A.T. Automatic species identification of live moths. *Knowl.-Based Syst.* **2007**, *20*, 195–202. [\[CrossRef\]](#)
39. Cheng, J.; Xie, B.; Lin, C.; Ji, L. A comparative study in birds: Call-type-independent species and individual recognition using four machine-learning methods and two acoustic features. *Bioacoustics* **2012**, *21*, 157–171. [\[CrossRef\]](#)
40. Colonna, J.G.; Gama, J.; Nakamura, E.F. A comparison of hierarchical multi-output recognition approaches for anuran classification. *Mach. Learn.* **2018**, *107*, 1651–1671. [\[CrossRef\]](#)
41. Xu, W.; Zhang, X.; Yao, L.; Xue, W.; Wei, B. A multi-view CNN-based acoustic classification system for automatic animal species identification. *Ad. Hoc Netw.* **2020**, *102*, 102115. [\[CrossRef\]](#)
42. Gan, H.; Zhang, J.; Towsey, M.; Truskinger, A.; Stark, D.; Van Rensburg, B.J.; Li, Y.; Roe, P. A novel frog chorusing recognition method with acoustic indices and machine learning. *Future Gener. Comput. Syst.* **2021**, *125*, 485–495. [\[CrossRef\]](#)
43. Xie, J.; Towsey, M.; Zhang, J.; Roe, P. Acoustic classification of Australian frogs based on enhanced features and machine learning algorithms. *Appl. Acoust.* **2016**, *113*, 193–201. [\[CrossRef\]](#)
44. Kim, J.; Oh, J.; Heo, T.Y. Acoustic scene classification and visualization of beehive sounds using machine learning algorithms and grad-CAM. *Math. Probl. Eng.* **2021**, *2021*, 5594498. [\[CrossRef\]](#)

45. Kirkeby, C.; Rydhmer, K.; Cook, S.M.; Strand, A.; Torrance, M.T.; Swain, J.L.; Prangma, J.; Johnen, A.; Jensen, M.; Brydegaard, M.; et al. Advances in automatic identification of flying insects using optical sensors and machine learning. *Sci. Rep.* **2021**, *11*, 1555. [\[CrossRef\]](#)
46. Tacioli, L.; Toledo, L.; Medeiros, C. An architecture for animal sound identification based on multiple feature extraction and classification algorithms. In *Anais do XI Brazilian e-Science Workshop*; Sociedade Brasileira de Computação: Porto Alegre, Brazil, 2017; pp. 29–36.
47. Bergler, C.; Schröter, H.; Cheng, R.X.; Barth, V.; Weber, M.; Nöth, E.; Hofer, H.; Maier, A. ORCA-SPOT: An automatic killer whale sound detection toolkit using deep learning. *Sci. Rep.* **2019**, *9*, 10997. [\[CrossRef\]](#)
48. Zhang, L.; Saleh, I.; Thapaliya, S.; Louie, J.; Figueroa-Hernandez, J.; Ji, H. An empirical evaluation of machine learning approaches for species identification through bioacoustics. In Proceedings of the 2017 International Conference on Computational Science and Computational Intelligence (CSCI), Las Vegas, NV, USA, 14–16 December 2017; pp. 489–494.
49. Şaşmaz, E.; Tek, F.B. Animal sound classification using a convolutional neural network. In Proceedings of the 2018 3rd International Conference on Computer Science and Engineering (UBMK), Sarajevo, Bosnia and Herzegovina, 20–23 September 2018; pp. 625–629.
50. Nanni, L.; Brahnam, S.; Lumini, A.; Maguolo, G. Animal sound classification using dissimilarity spaces. *Appl. Sci.* **2020**, *10*, 8578. [\[CrossRef\]](#)
51. Romero, J.; Luque, A.; Carrasco, A. Animal Sound Classification using Sequential Classifiers. In Proceedings of the BIOSIGNALS, 21–23 February 2011; ScitePress Digital Library, 2017; pp. 242–247.
52. Kim, C.I.; Cho, Y.; Jung, S.; Rew, J.; Hwang, E. Animal sounds classification scheme based on multi-feature network with mixed datasets. *KSII Trans. Internet Inf. Syst. (TIIS)* **2020**, *14*, 3384–3398.
53. Weninger, F.; Schuller, B. Audio recognition in the wild: Static and dynamic classification on a real-world database of animal vocalizations. In Proceedings of the 2011 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), Prague, Czech Republic, 22–27 May 2011; pp. 337–340.
54. Chesmore, E.; Ohya, E. Automated identification of field-recorded songs of four British grasshoppers using bioacoustic signal recognition. *Bull. Entomol. Res.* **2004**, *94*, 319–330. [\[CrossRef\]](#)
55. Mac Aodha, O.; Gibb, R.; Barlow, K.E.; Browning, E.; Firman, M.; Freeman, R.; Harder, B.; Kinsey, L.; Mead, G.R.; Newson, S.E.; et al. Bat detective—Deep learning tools for bat acoustic signal detection. *PLoS Comput. Biol.* **2018**, *14*, e1005995. [\[CrossRef\]](#) [\[PubMed\]](#)
56. Zgank, A. Bee swarm activity acoustic classification for an IoT-based farm service. *Sensors* **2019**, *20*, 21. [\[CrossRef\]](#) [\[PubMed\]](#)
57. Zhong, M.; Castellote, M.; Dodhia, R.; Lavista Ferres, J.; Keogh, M.; Brewer, A. Beluga whale acoustic signal classification using deep learning neural network models. *J. Acoust. Soc. Am.* **2020**, *147*, 1834–1841. [\[CrossRef\]](#) [\[PubMed\]](#)
58. Bravo Sanchez, F.J.; Hossain, M.R.; English, N.B.; Moore, S.T. Bioacoustic classification of avian calls from raw sound waveforms with an open-source deep learning architecture. *Sci. Rep.* **2021**, *11*, 15733. [\[CrossRef\]](#) [\[PubMed\]](#)
59. Kiskin, I.; Zilli, D.; Li, Y.; Sinka, M.; Willis, K.; Roberts, S. Bioacoustic detection with wavelet-conditioned convolutional neural networks. *Neural Comput. Appl.* **2020**, *32*, 915–927. [\[CrossRef\]](#)
60. Pourhomayoun, M.; Dugan, P.; Popescu, M.; Clark, C. Bioacoustic signal classification based on continuous region processing, grid masking and artificial neural network. *arXiv* **2013**, arXiv:1305.3635.
61. Mehyaadin, A.E.; Abdulazeez, A.M.; Hasan, D.A.; Saeed, J.N. Birds sound classification based on machine learning algorithms. *Asian J. Res. Comput. Sci.* **2021**, *9*, 68530. [\[CrossRef\]](#)
62. Arzar, N.N.K.; Sabri, N.; Johari, N.F.M.; Shari, A.A.; Noordin, M.R.M.; Ibrahim, S. Butterfly species identification using convolutional neural network (CNN). In Proceedings of the 2019 IEEE International Conference on Automatic Control and Intelligent Systems (I2CACIS), Selangor, Malaysia, 29 June 2019; pp. 221–224.
63. Shamir, L.; Yerby, C.; Simpson, R.; von Benda-Beckmann, A.M.; Tyack, P.; Samarra, F.; Miller, P.; Wallin, J. Classification of large acoustic datasets using machine learning and crowdsourcing: Application to whale calls. *J. Acoust. Soc. Am.* **2014**, *135*, 953–962. [\[CrossRef\]](#)
64. Molnár, C.; Kaplan, F.; Roy, P.; Pachet, F.; Pongrácz, P.; Dóka, A.; Miklósi, Á. Classification of dog barks: A machine learning approach. *Anim. Cogn.* **2008**, *11*, 389–400. [\[CrossRef\]](#)
65. Gunasekaran, S.; Revathy, K. Content-based classification and retrieval of wild animal sounds using feature selection algorithm. In Proceedings of the 2010 Second International Conference on Machine Learning and Computing, Bangalore, India, 9–11 February 2010; pp. 272–275.
66. Nanni, L.; Maguolo, G.; Paci, M. Data augmentation approaches for improving animal audio classification. *Ecol. Inform.* **2020**, *57*, 101084. [\[CrossRef\]](#)
67. Ko, K.; Park, S.; Ko, H. Convolutional feature vectors and support vector machine for animal sound classification. In Proceedings of the 2018 40th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), Honolulu, HI, USA, 18–21 July 2018; pp. 376–379.
68. Bermant, P.C.; Bronstein, M.M.; Wood, R.J.; Gero, S.; Gruber, D.F. Deep machine learning techniques for the detection and classification of sperm whale bioacoustics. *Sci. Rep.* **2019**, *9*, 12588. [\[CrossRef\]](#) [\[PubMed\]](#)
69. Thakur, A.; Thapar, D.; Rajan, P.; Nigam, A. Deep metric learning for bioacoustic classification: Overcoming training data scarcity using dynamic triplet loss. *J. Acoust. Soc. Am.* **2019**, *146*, 534–547. [\[CrossRef\]](#) [\[PubMed\]](#)

70. Morfi, V.; Lachlan, R.F.; Stowell, D. Deep perceptual embeddings for unlabelled animal sound events. *J. Acoust. Soc. Am.* **2021**, *150*, 2–11. [\[CrossRef\]](#)
71. Bardeli, R.; Wolff, D.; Kurth, F.; Koch, M.; Tauchert, K.H.; Frommolt, K.H. Detecting bird sounds in a complex acoustic environment and application to bioacoustic monitoring. *Pattern Recognit. Lett.* **2010**, *31*, 1524–1534. [\[CrossRef\]](#)
72. Eliopoulos, P.; Potamitis, I.; Kontodimas, D.C.; Givropoulou, E. Detection of adult beetles inside the stored wheat mass based on their acoustic emissions. *J. Econ. Entomol.* **2015**, *108*, 2808–2814. [\[CrossRef\]](#)
73. Kim, T.; Kim, J.W.; Lee, K. Detection of sleep disordered breathing severity using acoustic biomarker and machine learning techniques. *Biomed. Eng. Online* **2018**, *17*, 16. [\[CrossRef\]](#)
74. Yazgaç, B.G.; Kırıcı, M.; Kivan, M. Detection of sunn pests using sound signal processing methods. In Proceedings of the 2016 Fifth International Conference on Agro-Geoinformatics (Agro-Geoinformatics), Tianjin, China, 18–20 July 2016; pp. 1–6.
75. Pandeya, Y.R.; Kim, D.; Lee, J. Domestic cat sound classification using learned features from deep neural nets. *Appl. Sci.* **2018**, *8*, 1949. [\[CrossRef\]](#)
76. Al-Ahmadi, S. Energy efficient animal sound recognition scheme in wireless acoustic sensors networks. *Int. J. Wirel. Mob. Netw. (IJWMN)* **2020**, *12*, 31–38.
77. Huang, C.J.; Yang, Y.J.; Yang, D.X.; Chen, Y.J. Frog classification using machine learning techniques. *Expert Syst. Appl.* **2009**, *36*, 3737–3743. [\[CrossRef\]](#)
78. Salamon, J.; Bello, J.P.; Farnsworth, A.; Kelling, S. Fusing shallow and deep learning for bioacoustic bird species classification. In Proceedings of the 2017 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), New Orleans, LA, USA, 5–9 March 2017; pp. 141–145.
79. Xie, J.; Zhu, M. Handcrafted features and late fusion with deep learning for bird sound classification. *Ecol. Inform.* **2019**, *52*, 74–81. [\[CrossRef\]](#)
80. Chao, K.W.; Hu, N.Z.; Chao, Y.C.; Su, C.K.; Chiu, W.H. Implementation of artificial intelligence for classification of frogs in bioacoustics. *Symmetry* **2019**, *11*, 1454. [\[CrossRef\]](#)
81. Zgank, A. IoT-based bee swarm activity acoustic classification using deep neural networks. *Sensors* **2021**, *21*, 676. [\[CrossRef\]](#) [\[PubMed\]](#)
82. Ribeiro, A.P.; da Silva, N.F.F.; Mesquita, F.N.; Araújo, P.d.C.S.; Rosa, T.C.; Mesquita-Neto, J.N. Machine learning approach for automatic recognition of tomato-pollinating bees based on their buzzing-sounds. *PLoS Comput. Biol.* **2021**, *17*, e1009426. [\[CrossRef\]](#)
83. Noda, J.J.; Travieso, C.M.; Sanchez-Rodriguez, D. Methodology for automatic bioacoustic classification of anurans based on feature fusion. *Expert Syst. Appl.* **2016**, *50*, 100–106. [\[CrossRef\]](#)
84. Chalmers, C.; Fergus, P.; Wich, S.; Longmore, S. Modelling Animal Biodiversity Using Acoustic Monitoring and Deep Learning. In Proceedings of the 2021 International Joint Conference on Neural Networks (IJCNN), Shenzhen, China, 18–22 July 2021; pp. 1–7.
85. Caruso, F.; Dong, L.; Lin, M.; Liu, M.; Gong, Z.; Xu, W.; Alonge, G.; Li, S. Monitoring of a nearshore small dolphin species using passive acoustic platforms and supervised machine learning techniques. *Front. Mar. Sci.* **2020**, *7*, 267. [\[CrossRef\]](#)
86. Zhong, M.; LeBien, J.; Campos-Cerqueira, M.; Dodhia, R.; Ferres, J.L.; Velev, J.P.; Aide, T.M. Multispecies bioacoustic classification using transfer learning of deep convolutional neural networks with pseudo-labeling. *Appl. Acoust.* **2020**, *166*, 107375. [\[CrossRef\]](#)
87. Kim, D.; Lee, Y.; Ko, H. Multi-Task Learning for Animal Species and Group Category Classification. In Proceedings of the 2019 7th International Conference on Information Technology: IoT and Smart City, Guangzhou, China, 2019; pp. 435–438.
88. Dugan, P.J.; Rice, A.N.; Urazghildiev, I.R.; Clark, C.W. North Atlantic right whale acoustic signal processing: Part I. Comparison of machine learning recognition algorithms. In Proceedings of the 2010 IEEE Long Island Systems, Applications and Technology Conference, Farmingdale, NY, USA, 7 May 2010; pp. 1–6.
89. Balemarthy, S.; Sajjanhar, A.; Zheng, J.X. Our practice of using machine learning to recognize species by voice. *arXiv* **2018**, arXiv:1810.09078.
90. Gradišek, A.; Slapničar, G.; Šorn, J.; Luštrek, M.; Gams, M.; Grad, J. Predicting species identity of bumblebees through analysis of flight buzzing sounds. *Bioacoustics* **2017**, *26*, 63–76. [\[CrossRef\]](#)
91. LOSTANLEN, V.; SALAMON, J.; FARNSWORTH, A.; KELLING, S.; BELLO, J.P. Robust sound event detection in bioacoustic sensor networks. *PLoS ONE* **2019**, *14*, e0214168. [\[CrossRef\]](#)
92. Xie, J.; Bertram, S.M. Using machine learning techniques to classify cricket sound. In *Eleventh International Conference on Signal Processing Systems*; SPIE: Chengdu, China, 2019; Volume 11384, pp. 141–148.
93. Nanni, L.; Rigo, A.; Lumini, A.; Brahnam, S. Spectrogram classification using dissimilarity space. *Appl. Sci.* **2020**, *10*, 4176. [\[CrossRef\]](#)
94. Salamon, J.; Bello, J.P.; Farnsworth, A.; Robbins, M.; Keen, S.; Klinck, H.; Kelling, S. Towards the automatic classification of avian flight calls for bioacoustic monitoring. *PLoS ONE* **2016**, *11*, e0166866. [\[CrossRef\]](#) [\[PubMed\]](#)
95. Kawakita, S.; Ichikawa, K. Automated classification of bees and hornet using acoustic analysis of their flight sounds. *Apidologie* **2019**, *50*, 71–79. [\[CrossRef\]](#)
96. Li, L.; Qiao, G.; Liu, S.; Qing, X.; Zhang, H.; Mazhar, S.; Niu, F. Automated classification of Tursiops aduncus whistles based on a depth-wise separable convolutional neural network and data augmentation. *J. Acoust. Soc. Am.* **2021**, *150*, 3861–3873. [\[CrossRef\]](#) [\[PubMed\]](#)

97. Ntalampiras, S. Automatic acoustic classification of insect species based on directed acyclic graphs. *J. Acoust. Soc. Am.* **2019**, *145*, EL541–EL546. [\[CrossRef\]](#)
98. Chen, Y.; Why, A.; Batista, G.; Mafra-Neto, A.; Keogh, E. Flying insect classification with inexpensive sensors. *J. Insect Behav.* **2014**, *27*, 657–677. [\[CrossRef\]](#)
99. Zhu, L.-Q. Insect sound recognition based on mfcc and pnn. In Proceedings of the 2011 International Conference on Multimedia and Signal Processing, Guilin, China, 14–15 May 2011; Volume 2, pp. 42–46.
100. D'mello, G.C.; Hussain, R. Insect Inspection on the basis of their Flight Sound. *Int. J. Sci. Eng. Res.* **2015**, *6*, 49–54.
101. Aide, T.M.; Corrada-Bravo, C.; Campos-Cerqueira, M.; Milan, C.; Vega, G.; Alvarez, R. Real-time bioacoustics monitoring and automated species identification. *PeerJ* **2013**, *1*, e103. [\[CrossRef\]](#)
102. Rathore, D.S.; Ram, B.; Pal, B.; Malviya, S. Analysis of classification algorithms for insect detection using MATLAB. In Proceedings of the 2nd International Conference on Advanced Computing and Software Engineering (ICACSE), Sultanpur, India, 8–9 February 2019.
103. Ovaskainen, O.; Moliterno de Camargo, U.; Somervuo, P. Animal Sound Identifier (ASI): Software for automated identification of vocal animals. *Ecol. Lett.* **2018**, *21*, 1244–1254. [\[CrossRef\]](#)
104. Müller, L.; Marti, M. Bird Sound Classification using a Bidirectional LSTM. In Proceedings of the CLEF (Working Notes), Avignon, France, 10–14 September 2018.
105. Supriya, P.R.; Bhat, S.; Shivani, S.S. Classification of birds based on their sound patterns using GMM and SVM classifiers. *Int. Res. J. Eng. Technol.* **2018**, *5*, 4708–4711.
106. Oikarinen, T.; Srinivasan, K.; Meisner, O.; Hyman, J.B.; Parmar, S.; Fanucci-Kiss, A.; Desimone, R.; Landman, R.; Feng, G. Deep convolutional network for animal sound classification and source attribution using dual audio recordings. *J. Acoust. Soc. Am.* **2019**, *145*, 654–662. [\[CrossRef\]](#) [\[PubMed\]](#)
107. Nanni, L.; Ghidoni, S.; Brahnam, S. Ensemble of convolutional neural networks for bioimage classification. *Appl. Comput. Inform.* **2020**, *17*, 19–35. [\[CrossRef\]](#)
108. Hong, R.; Wang, M.; Yuan, X.T.; Xu, M.; Jiang, J.; Yan, S.; Chua, T.S. Video accessibility enhancement for hearing-impaired users. *ACM Trans. Multimed. Comput. Commun. Appl. (TOMM)* **2011**, *7*, 1–19. [\[CrossRef\]](#)
109. Wang, W.; Chen, Z.; Xing, B.; Huang, X.; Han, S.; Agu, E. A smartphone-based digital hearing aid to mitigate hearing loss at specific frequencies. In Proceedings of the 1st Workshop on Mobile Medical Applications, Seattle, WA, USA, 2014; pp. 1–5.
110. Bountourakis, V.; Vrysis, L.; Papanikolaou, G. Machine learning algorithms for environmental sound recognition: Towards soundscape semantics. In Proceedings of the Audio Mostly 2015 on Interaction with Sound, Thessaloniki Greece, 2015; pp. 1–7.
111. Li, M.; Gao, Z.; Zang, X.; Wang, X. Environmental noise classification using convolution neural networks. In Proceedings of the 2018 International Conference on Electronics and Electrical Engineering Technology, Tianjin, China, 19–21 September 2018; pp. 182–185.
112. Alsouda, Y.; Pillana, S.; Kurti, A. Iot-based urban noise identification using machine learning: Performance of SVM, KNN, bagging, and random forest. In Proceedings of the International Conference on Omni-Layer Intelligent Systems, Crete, Greece, 5–7 May 2019; pp. 62–67.
113. Kurnaz, S.; Aljabery, M.A. Predict the type of hearing aid of audiology patients using data mining techniques. In Proceedings of the Fourth International Conference on Engineering & MIS 2018, Istanbul, Turkey, 19–20 June 2018; pp. 1–6.
114. Wang, W.; Seraj, F.; Meratnia, N.; Havinga, P.J. Privacy-aware environmental sound classification for indoor human activity recognition. In Proceedings of the 12th ACM International Conference on Pervasive Technologies Related to Assistive Environments, Island of Rhodes, Greece, 5–7 June 2019; pp. 36–44.
115. Seker, H.; Inik, O. CnnSound: Convolutional Neural Networks for the Classification of Environmental Sounds. In Proceedings of the 2020 4th International Conference on Advances in Artificial Intelligence, London, UK, 9–11 October 2020; pp. 79–84.
116. Sigitia, S.; Stark, A.M.; Krstulović, S.; Plumbly, M.D. Automatic environmental sound recognition: Performance versus computational cost. *IEEE/ACM Trans. Audio Speech Lang. Process.* **2016**, *24*, 2096–2107. [\[CrossRef\]](#)
117. Van De Laar, T.; de Vries, B. A probabilistic modeling approach to hearing loss compensation. *IEEE/ACM Trans. Audio Speech Lang. Process.* **2016**, *24*, 2200–2213. [\[CrossRef\]](#)
118. Salehi, H.; Suelzle, D.; Folkeard, P.; Parsa, V. Learning-based reference-free speech quality measures for hearing aid applications. *IEEE/ACM Trans. Audio Speech Lang. Process.* **2018**, *26*, 2277–2288. [\[CrossRef\]](#)
119. Demir, F.; Abdullah, D.A.; Sengur, A. A new deep CNN model for environmental sound classification. *IEEE Access* **2020**, *8*, 66529–66537. [\[CrossRef\]](#)
120. Ridha, A.M.; Shehieb, W. Assistive Technology for Hearing-Impaired and Deaf Students Utilizing Augmented Reality. In Proceedings of the 2021 IEEE Canadian Conference on Electrical and Computer Engineering (CCECE), Virtual Conference, 12–17 September 2021; pp. 1–5.
121. Ayu, A.I.S.M.; Karyono, K.K. Audio detection (Audition): Android based sound detection application for hearing-impaired using AdaBoostM1 classifier with REPTree weaklearner. In Proceedings of the 2014 Asia-Pacific Conference on Computer Aided System Engineering (APCASE), South Kuta, Indonesia, 10–12 February 2014; pp. 136–140.
122. Chen, C.Y.; Kuo, P.Y.; Chiang, Y.H.; Liang, J.Y.; Liang, K.W.; Chang, P.C. Audio-Based Early Warning System of Sound Events on the Road for Improving the Safety of Hearing-Impaired People. In Proceedings of the 2019 IEEE 8th Global Conference on Consumer Electronics (GCCE), OSAKA, Japan, 15–18 October 2019; pp. 933–936.

123. Bhat, G.S.; Shankar, N.; Panahi, I.M. Automated machine learning based speech classification for hearing aid applications and its real-time implementation on smartphone. In Proceedings of the 2020 42nd Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC), Montreal, QC, Canada, 20–24 July 2020; pp. 956–959.
124. Healy, E.W.; Yoho, S.E. Difficulty understanding speech in noise by the hearing impaired: Underlying causes and technological solutions. In Proceedings of the 2016 38th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), Orlando, FL, USA, 16–20 August 2016; pp. 89–92.
125. Jatturas, C.; Chokkoedsakul, S.; Ayudhya, P.D.N.; Pankaew, S.; Sopavanit, C.; Asdornwiset, W. Recurrent Neural Networks for Environmental Sound Recognition using Scikit-learn and Tensorflow. In Proceedings of the 2019 16th International Conference on Electrical Engineering/Electronics, Computer, Telecommunications and Information Technology (ECTI-CON), Pattaya, Thailand, 10–13 July 2019; pp. 806–809.
126. Saleem, N.; Khattak, M.I.; Ahmad, S.; Ali, M.Y.; Mohmand, M.I. Machine Learning Approach for Improving the Intelligibility of Noisy Speech. In Proceedings of the 2020 17th International Bhurban Conference on Applied Sciences and Technology (IBCAST), Islamabad, Pakistan, 14–18 January 2020; pp. 303–308.
127. Davis, N.; Suresh, K. Environmental sound classification using deep convolutional neural networks and data augmentation. In Proceedings of the 2018 IEEE Recent Advances in Intelligent Computational Systems (RAICS), Thiruvananthapuram, India, 6–8 December 2018; pp. 41–45.
128. Chu, S.; Narayanan, S.; Kuo, C.C.J. Environmental sound recognition with time–frequency audio features. *IEEE Trans. Audio Speech Lang. Process.* **2009**, *17*, 1142–1158. [\[CrossRef\]](#)
129. Chu, S.; Narayanan, S.; Kuo, C.C.J.; Mataric, M.J. Where Am I? Scene recognition for mobile robots using audio features. In Proceedings of the 2006 IEEE International Conference on Multimedia and Expo, Toronto, ON, Canada, 9–12 July 2006; pp. 885–888.
130. Ullo, S.L.; Khare, S.K.; Bajaj, V.; Sinha, G.R. Hybrid computerized method for environmental sound classification. *IEEE Access* **2020**, *8*, 124055–124065. [\[CrossRef\]](#)
131. Zhang, X.; Zou, Y.; Shi, W. Dilated convolution neural network with LeakyReLU for environmental sound classification. In Proceedings of the 2017 22nd International Conference on Digital Signal Processing (DSP), London, UK, 23–25 August 2017; pp. 1–5.
132. Piczak, K.J. Environmental sound classification with convolutional neural networks. In Proceedings of the 2015 IEEE 25th International Workshop on Machine Learning for Signal Processing (MLSP), Boston, MA, USA, 17–20 September 2015; pp. 1–6.
133. Han, B.J.; Hwang, E. Environmental sound classification based on feature collaboration. In Proceedings of the 2009 IEEE International Conference on Multimedia and Expo, New York, NY, USA, 28 June–3 July 2009; pp. 542–545.
134. Wang, J.C.; Lin, C.H.; Chen, B.W.; Tsai, M.K. Gabor-based nonuniform scale-frequency map for environmental sound classification in home automation. *IEEE Trans. Autom. Sci. Eng.* **2013**, *11*, 607–613. [\[CrossRef\]](#)
135. Salamon, J.; Bello, J.P. Deep convolutional neural networks and data augmentation for environmental sound classification. *IEEE Signal Process. Lett.* **2017**, *24*, 279–283. [\[CrossRef\]](#)
136. Wang, J.C.; Wang, J.F.; He, K.W.; Hsu, C.S. Environmental sound classification using hybrid SVM/KNN classifier and MPEG-7 audio low-level descriptor. In Proceedings of the 2006 IEEE International Joint Conference on Neural Network Proceedings, Vancouver, BC, Canada, 16–21 July 2006; pp. 1731–1735.
137. Zhao, Y.; Li, J.; Zhang, M.; Lu, Y.; Xie, H.; Tian, Y.; Qiu, W. Machine learning models for the hearing impairment prediction in workers exposed to complex industrial noise: A pilot study. *Ear Hear.* **2019**, *40*, 690. [\[CrossRef\]](#) [\[PubMed\]](#)
138. Tokozume, Y.; Harada, T. Learning environmental sounds with end-to-end convolutional neural network. In Proceedings of the 2017 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), New Orleans, LA, USA, 5–9 March 2017; pp. 2721–2725.
139. Nossier, S.A.; Rizk, M.; Moussa, N.D.; el Shehaby, S. Enhanced smart hearing aid using deep neural networks. *Alex. Eng. J.* **2019**, *58*, 539–550. [\[CrossRef\]](#)
140. Abdoli, S.; Cardinal, P.; Koerich, A.L. End-to-end environmental sound classification using a 1D convolutional neural network. *Expert Syst. Appl.* **2019**, *136*, 252–263. [\[CrossRef\]](#)
141. Mushtaq, Z.; Su, S.F. Environmental sound classification using a regularized deep convolutional neural network with data augmentation. *Appl. Acoust.* **2020**, *167*, 107389. [\[CrossRef\]](#)
142. Chen, Y.; Guo, Q.; Liang, X.; Wang, J.; Qian, Y. Environmental sound classification with dilated convolutions. *Appl. Acoust.* **2019**, *148*, 123–132. [\[CrossRef\]](#)
143. Mushtaq, Z.; Su, S.F.; Tran, Q.V. Spectral images based environmental sound classification using CNN with meaningful data augmentation. *Appl. Acoust.* **2021**, *172*, 107581. [\[CrossRef\]](#)
144. Ahmad, S.; Agrawal, S.; Joshi, S.; Taran, S.; Bajaj, V.; Demir, F.; Sengur, A. Environmental sound classification using optimum allocation sampling based empirical mode decomposition. *Phys. A Stat. Mech. Appl.* **2020**, *537*, 122613. [\[CrossRef\]](#)
145. Medhat, F.; Chesmore, D.; Robinson, J. Masked conditional neural networks for environmental sound classification. In *International Conference on Innovative Techniques and Applications of Artificial Intelligence*; Springer: Cham, Switzerland, 2017; pp. 21–33.
146. Zhang, Z.; Xu, S.; Cao, S.; Zhang, S. Deep convolutional neural network with mixup for environmental sound classification. In *Chinese Conference on Pattern Recognition and Computer Vision (PRCV)*; Springer: Cham, Switzerland, 2018; pp. 356–367.

147. Sailor, H.B.; Agrawal, D.M.; Patil, H.A. Unsupervised Filterbank Learning Using Convolutional Restricted Boltzmann Machine for Environmental Sound Classification. In Proceedings of the INTERSPEECH 2017, Stockholm, Sweden, 20–24 August 2017; Volume 8, p. 9.
148. Sharma, J.; Granmo, O.C.; Goodwin, M. Environment Sound Classification Using Multiple Feature Channels and Attention Based Deep Convolutional Neural Network. In Proceedings of the INTERSPEECH 2020, Shanghai, China, 25–29 October 2020; pp. 1186–1190.
149. Mohaimenuzzaman, M.; Bergmeir, C.; West, I.T.; Meyer, B. Environmental Sound Classification on the Edge: A Pipeline for Deep Acoustic Networks on Extremely Resource-Constrained Devices. *arXiv* **2021**, arXiv:2103.03483.
150. Toffa, O.K.; Mignotte, M. Environmental sound classification using local binary pattern and audio features collaboration. *IEEE Trans. Multimed.* **2020**, *23*, 3978–3985. [[CrossRef](#)]
151. Khamparia, A.; Gupta, D.; Nguyen, N.G.; Khanna, A.; Pandey, B.; Tiwari, P. Sound classification using convolutional neural network and tensor deep stacking network. *IEEE Access* **2019**, *7*, 7717–7727. [[CrossRef](#)]
152. Su, Y.; Zhang, K.; Wang, J.; Madani, K. Environment sound classification using a two-stream CNN based on decision-level fusion. *Sensors* **2019**, *19*, 1733. [[CrossRef](#)] [[PubMed](#)]
153. Bragg, D.; Huynh, N.; Ladner, R.E. A personalizable mobile sound detector app design for deaf and hard-of-hearing users. In Proceedings of the 18th International ACM SIGACCESS Conference on Computers and Accessibility, Reno, NV, USA, 23–26 October 2016; pp. 3–13.
154. Jatturas, C.; Chokkoedsakul, S.; Avudhva, P.D.N.; Pankaew, S.; Sopavanit, C.; Asdornwised, W. Feature-based and Deep Learning-based Classification of Environmental Sound. In Proceedings of the 2019 IEEE International Conference on Consumer Electronics-Asia (ICCE-Asia), Bangkok, Thailand, 12–14 June 2019; pp. 126–130.
155. Smith, D.; Ma, L.; Ryan, N. Acoustic environment as an indicator of social and physical context. *Pers. Ubiquitous Comput.* **2006**, *10*, 241–254. [[CrossRef](#)]
156. Ma, L.; Smith, D.J.; Milner, B.P. Context awareness using environmental noise classification. In Proceedings of the INTERSPEECH, Geneva, Switzerland, 1–4 September 2003.
157. Allen, J. Short term spectral analysis, synthesis, and modification by discrete Fourier transform. *IEEE Trans. Acoust. Speech Signal Process.* **1977**, *25*, 235–238. [[CrossRef](#)]
158. Allen, J.B.; Rabiner, L.R. A unified approach to short-time Fourier analysis and synthesis. *Proc. IEEE* **1977**, *65*, 1558–1564. [[CrossRef](#)]
159. Allen, J. Applications of the short time Fourier transform to speech processing and spectral analysis. In Proceedings of the ICASSP’82. IEEE International Conference on Acoustics, Speech, and Signal Processing, Paris, France, 3–5 May 1982; Volume 7, pp. 1012–1015.
160. Lu, L.; Hanjalic, A. Text-like segmentation of general audio for content-based retrieval. *IEEE Trans. Multimed.* **2009**, *11*, 658–669. [[CrossRef](#)]