



Automatic synthesis of insects bioacoustics using machine learning: a systematic review

Henry Kyalo^{1,3} · Henri Tonnang^{1,4} · James Egonyu² · John Olukuru³ · Chrysantus Tanga¹ · Kennedy Senagi¹

Received: 28 February 2024 / Accepted: 29 November 2024 / Published online: 8 January 2025
© The Author(s) 2025

Abstract

The application of machine learning has received increasing attention in the synthesis of insect sounds to preserve biodiversity. This study reviewed current literature on the application of these techniques in the automatic synthesis of insect bioacoustic and their applications in insects as food and feed, improving pest management, and as well as managing pollinators. To achieve this, the study used Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) methodology to identify, screen, and include the final articles used in this review, based on criteria such as papers addressing machine learning in insect acoustics, biodiversity, ecology conservation, etc. This study revealed that most of the researchers used secondary data and the microphone was the common tool used to record sound signals. Sound signals were mainly pre-processed using techniques such as denoising, segmentation, and windowing. Sound signal classification algorithms were categorized mainly as shallow and deep machine learning algorithms. In the shallow machine learning algorithms, the most common method of feature extraction was the Mel-Frequency Cepstral Coefficient (MFCC) and the Support Vector Machine (SVM) was the most commonly used algorithm. In deep learning, spectrogram image features were widely extracted and the Convolutional Neural Network (CNN) was mostly used to synthesize the spectral features. This paper also reviewed recent developments in insect bioacoustics signals processing, applications, and future directions. Generally, machine learning algorithms can be applied and deployed successfully to different insects' automatic synthesis problems to improve the production of insects (as food and/or feed), and improve/preserve diversity and life on Earth.

Keywords Insects · Bioacoustics · Machine learning · Insects as food · Insects as feed · Ecological conservation

Introduction

Background

Insects belong to the class Insecta within the phylum Arthropoda (Auerswald and Lopata 2005). They are the largest and most diverse group of organisms on earth, with estimates of millions of species yet to be discovered and

described. Insects are incredibly varied in terms of their morphology, behavior, and ecological roles. They encompass a wide range of forms, from tiny parasitic wasps to large beetles and butterflies. Depending on the species, insects play an important role in ecology and conservation of biodiversity (Riyaz et al. 2022). Their impact includes pollination where insects facilitate the transfer of pollen between flowers and enable the production of fruits and seeds (Kevan et al. 1990; Ghisbain et al. 2021). Further, insects play a key role in decomposition and nutrient cycling by breaking down dead organic matter such as leaves, wood, and carcasses (Crespo-Pérez et al. 2020). This decomposition process is essential for recycling nutrients back into the ecosystem, allowing plants to access essential elements for growth (Prasad et al. 2021; Astrid 2018). Moreover, insects are a source of food for many animals, including humans (Lange and Nakamura 2021). Insects are known to cause crop damage (Manosathiyadevan et al. 2017), and vector diseases such as malaria (Karunamoorthi et al. 2010), dengue

✉ Kennedy Senagi
ksenagi@icipc.org

¹ Data Management, Modelling and Geo-Information Unit, International Centre of Insect Physiology and Ecology, Nairobi, Kenya

² Department of Agriculture, Busitema University, Tororo, Uganda

³ iLab, Strathmore University, Nairobi, Kenya

⁴ School of Agricultural, Earth, and Environmental Sciences, University of KwaZulu-Natal, Durban, South Africa

fever (Shrestha and Lee 2020), etc. Insects also cause structural damage to wooden structures (Horwood et al. 2022). Insects such as grain weevils and flour beetles infest stored food products leading to waste and contamination (Srivastava and Subramanian 2016). Some insect species are known predators that feed on other insects or smaller organisms. These predators help regulate populations of other species and can contribute to biological pest control by keeping pest populations in check (Kremen and Chaplin-Kramer 2007).

Despite the important roles played by insects in ecology and biodiversity conservation management, the population of insects has been on the decline (Hallmann et al. 2017; Eggleton 2020). Some of the factors attributed to the decline of insects include pesticides (Riyaz et al. 2022), urbanization (Riyaz et al. 2022), light pollution (Owens et al. 2020), deforestation (Riyaz et al. 2022), habitat destruction (Sánchez-Bayo and Wyckhuys 2019), global warming (Harris et al. 2019; Soroye et al. 2020), agriculture extension (Riyaz et al. 2022) and fertilizers (Riyaz et al. 2022). Therefore, ecology and conservation management of insects are imperative to preserve and manage biodiversity and natural resources.

Ecology and conservation management can further be understood from insect sounds i.e., bioacoustics. Insect bioacoustics involves studying sound production, dispersion, and reception in insects. Insect bioacoustics plays a crucial role in understanding insects' survival, i.e., attracting mates, territory establishment, predator avoidance, and other behaviors (Ponce 2014). Among others, these sounds aid in species identification and biodiversity monitoring (Aide et al. 2013; Ponce 2014). Insect acoustics can also be applied in ecology, pest management, and conservation management, whereby researchers studying these signals, have developed intelligent acoustic traps to attract and capture selective pests (Hermosilla et al. 2021; Mankin 2012); this offers a more environmentally friendly approach to pest control. Sounds produced during insect pollination activities can be assessed to better understand and preserve forest ecosystems within the context of climate change. Folliot et al. (2022) analyzed sounds produced by pollination by flying insects and drumming woodpecker birds during different seasons. They found that insects were flying in bright, warm, and dry conditions, after dawn and before dusk during spring and summer while the woodpeckers were mainly drumming around March at the time of pair formation in cool and wet conditions. By studying weather information, Folliot et al. (2022) found that an increase in temperature favored pollination by flying insects and not woodpecker drumming while an increase in relative humidity favored woodpecker drumming but not pollinators. Therefore, these bioacoustics analytics results provided vital information about

biodiversity and environment description in the context of weather changes. Insects have robust sensory systems that enable them to adapt to simple and complex behaviors (Heller et al. 2017; Wang et al. 2016). For instance, insects like bees have olfactory receptors to detect pheromones and floral scents, which are crucial for foraging and mating (Uthoff et al. 2023). Moreover, the olfactory receptors allow them to perform complex behaviors such as navigating to specific flowers and communicating with other bees through the "waggle dance". Further, insects such as dragonflies, have compound eyes that provide a wide field of view and detect fast movements. This is essential for hunting prey and avoiding predators. Nonetheless, insects, like butterflies, can see ultraviolet light, which helps them locate nectar-rich flowers (Lunau et al. 2020).

Insects produce sounds, that can be perceived by humans (Potamitis et al. 2007). Manual identification of insect bioacoustics is quite tedious, time-consuming, and subject to human error (Alonso et al. 2017; Zhang et al. 2021). However, with advancements in technology, insect recognition and classification using bioacoustics can be done more accurately and precisely by the use of machine learning techniques. Machine learning has revolutionized the field of insect bioacoustic identification by automating and enhancing the process, and has achieved considerable success.

Following the advancement of automatic synthesis of insect bioacoustics using machine learning, it was necessary to do a state-of-the-art literature review of the compilation of the various machine-learning techniques for processing insect bioacoustics and their application to ecological and biodiversity conservation, pest management, and insects as food and feed.

The rest of this paper is organized as follows: Section "[Literature and research questions](#)" gives a brief overview of related work. Section "[Materials and methods](#)" discusses the methodology, Section "[Auto-processing of bioacoustics](#)" outlines the results and discussions about methodological procedures of automatic synthesis of bioacoustic, challenges, and future direction, Section "[Automatic synthesis of insects bioacoustic: current and future prospects](#)" discusses applications of automatic insects bioacoustic synthesis; current and future prospects. Section "[Conclusion](#)" concludes this paper.

Literature and research questions

Previous works

The use of machine learning and artificial intelligence techniques to analyze and classify sounds produced by living organisms has gained attention from the scientific

community due to the diverse applications and the cross-disciplinary nature of the field. For instance, some of the applications include conservation and ecology (Kampichler et al. 2010), animal behavior (Valletta et al. 2017), and communications such as specific vocalizations associated with mating, aggression, and other behaviors (Potamitis et al. 2007). Further, automated classification of acoustics can be applied in pest management, by identifying and tracking pests which can help in the development of targeted and environmentally friendly pest management strategies (Yazgaç et al. 2016). Uthoff et al. (2023) conducted a review of bioacoustics and vibration methodologies to automatically monitor bee colony behavior and health. They aimed to address how acoustics and vibration recordings have been used to determine queen presence and indicators of swarming, by presenting common methods and analyses. The review was guided by questions such as the application areas of the acoustics, methods used to record acoustics, and data analysis methods employed by different researchers ranging from machine learning and statistical algorithms, denoising methods, etc.

Stowell (2022) conducted a review focused on the current state of deep learning in computational bioacoustics, with the aim of clarifying key concepts. They highlighted knowledge gaps and under-explored areas that present opportunities for further research in the coming years. The author unearthed the deep learning techniques applied by different researchers, the evaluation metrics of the models, and how different studies have solved the problem of insufficient data since deep learning requires voluminous data, and signal preprocessing using deep learning. All these studies enhanced the knowledge of general acoustics.

Research questions

This research conducted a meta-analysis of existing literature on the use and application of machine learning tools and algorithms in insect bioacoustics with a focus on answering the questions outlined in Table 1.

Materials and methods

This section describes the materials and methods applied in this systematic literature review. To achieve this we applied the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) methodology to identify and screen scientific papers based on specific criteria. Figure 1 summarises the PRISMA methodology used in this paper. From the diagram, the methodology has 3 key steps which are: search and identification of articles, screening of identified articles for eligibility, and inclusion of final articles for review in this study.

Google Scholar, Scopus, and Web of Science were used as the main web search engines since they index articles published in most journals. The search engines allowed us

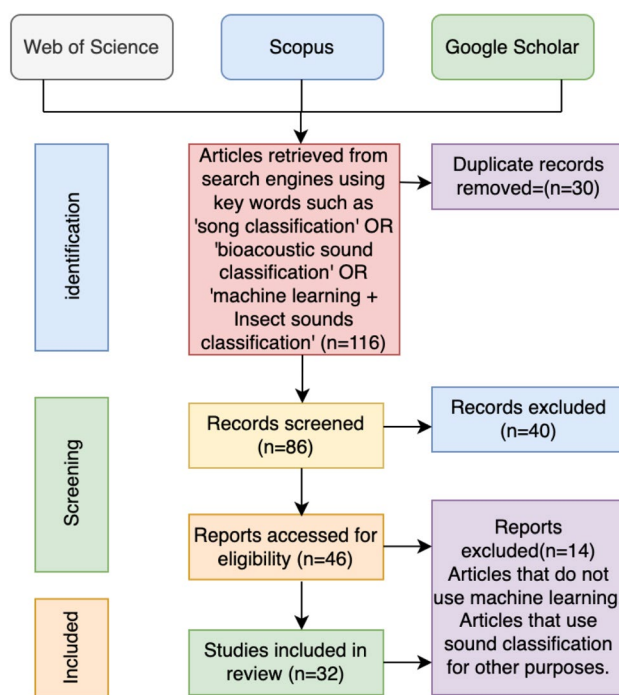


Fig. 1 Article selection criteria

Table 1 Outline of questions addressed in this study

Tag	Research Question
RQ1	Which are the data sources and data recording tools of insect bioacoustics?
RQ2	What data preprocessing techniques have been used in insect bioacoustics signals?
RQ3	Which feature extraction techniques have been used to extract features from insect bioacoustics signals?
RQ4	Which machine learning techniques have been used to process extracted insect bioacoustics signal features?
RQ5	How are the machine learning methods evaluated?
RQ6	What are the challenges and future directions of applications of machine learning in insect bioacoustics signals?
RQ7	How has automatic synthesis of insect bioacoustics been applied to ecological and biodiversity conservation?

to filter the sources based on the range of years indicated and the topics of interest as well. The search strategy used was based on keywords provided in Table 2. We note that keywords were derived from existing and related articles that were at the intersection of insect bioacoustics and machine learning (Wei et al. 2022; Fernandes et al. 2021; Paim et al. 2023; Trivedi and Shroff 2021; Robles-Guerrero et al. 2023; Várkonyi et al. 2023; Karar et al. 2021; Amlathe 2018; Hodzic and Hoang 2021; Potamitis et al. 2007; Kim et al. 2021b; Phung et al. 2017). This provided an exhaustive search of other related articles within a specified search period. Ideally, the search keywords were combined into search queries using the conjunction AND and OR operators. This study reviewed articles published in the last 7 years, i.e., from 2016 to 2023. The inclusion and exclusion criteria for the relevant materials are tabulated in Table 3.

Auto-processing of bioacoustics

A total of 32 articles met the selection criteria for inclusion in the review of insects' bioacoustics. Figure 2 summarises the count of articles published in each year between 2016 and 2023. Figure 3 shows the publishers where the majority of the articles were published.

This paper systematically reviews the selected papers and builds Sections “Data sources, data collection, sound signals pre-processing, and feature extraction” and “Machine learning in sound signal processing”. Section “Data sources, data collection, sound signals pre-processing, and feature extraction” highlights data collection and sources, sound signal processing, and feature extraction while Section “Machine learning in sound signal processing” introduces machine learning algorithms frequently used in insect acoustics and the evaluation metrics used to gauge the performance of these models. Figure 4 gives a summary of the acoustics data collection and automatics processing steps which are discussed as follows.

Data sources, data collection, sound signals pre-processing, and feature extraction

Data sources and data collection tools

The success of insect classification based on sound signals processing mainly relies on the quality of data recorded and how it is preprocessed to make the data machine-learning-ready. Generally, there are two sources of data: primary (collect first-hand records, e.g., in the laboratory or the field) and secondary (reuse available data, e.g., from

Table 2 Search queries applied on various search engines

Database	Search query
Scopus	TITLE (“acoustic classification” OR “bioacoustic signals” OR “sound signal processing” OR “MFCC” OR “PNN” OR “SVM” OR “LFCC” OR “bioacoustics” OR “acoustics”) OR ABS (“sound recognition” OR “insect sounds” OR “acoustic features” OR “CNN” OR “MFCC” OR “machine learning”) OR KEY (“Automatic acoustic identifications” OR “Deep convolutional neural networks” OR “Acoustic insect classification” OR “Sound analysis”)
Web of Science	TI=(“acoustic classification” OR “bioacoustic signals” OR “sound signal processing” OR “MFCC” OR “PNN” OR “SVM” OR “LFCC” OR “bio-acoustics” OR “acoustics”) OR AB=(“sound recognition” OR “insect sounds” OR “acoustic features” OR “CNN” OR “MFCC” OR “machine learning”) OR ALL=(“Automatic acoustic identifications” OR “Deep convolutional neural networks” OR “Acoustic insect classification” OR “Sound analysis”)
Google Scholar	“With all of the words” field: acoustic signal insect Sound * AND machine acoustic classification OR bioacoustic OR signals OR sound OR processing OR bioacoustics OR acoustics OR sound OR recognition OR insect OR sounds OR acoustic OR features OR Acoustic OR classification OR Sound OR analysis OR Insect OR learning OR Automatic OR identifications OR Deep OR learning OR Artificial OR intelligence OR AI)

Relevant articles from 2016 to 2023 were filtered

Table 3 Inclusion and exclusion criteria

Inclusion Criteria	Exclusion Criteria
Insects bioacoustics classification	Research not published in English
Publications that are peer-reviewed	Papers that were classified as not original research
Papers published in English	Sound classification not related to insects
Papers on insect acoustics using machine learning	Papers on bioacoustics and acoustics but not on insects
Papers which are not review articles	Papers which are review articles
Papers that do not have editorial material	Papers that are editorial materials
Papers that are not proceeding papers or correction papers	Papers that are proceeding papers or correction papers.

Fig. 2 A count of bioacoustic insect synthesis using machine learning papers published in each year between 2016 and 2023

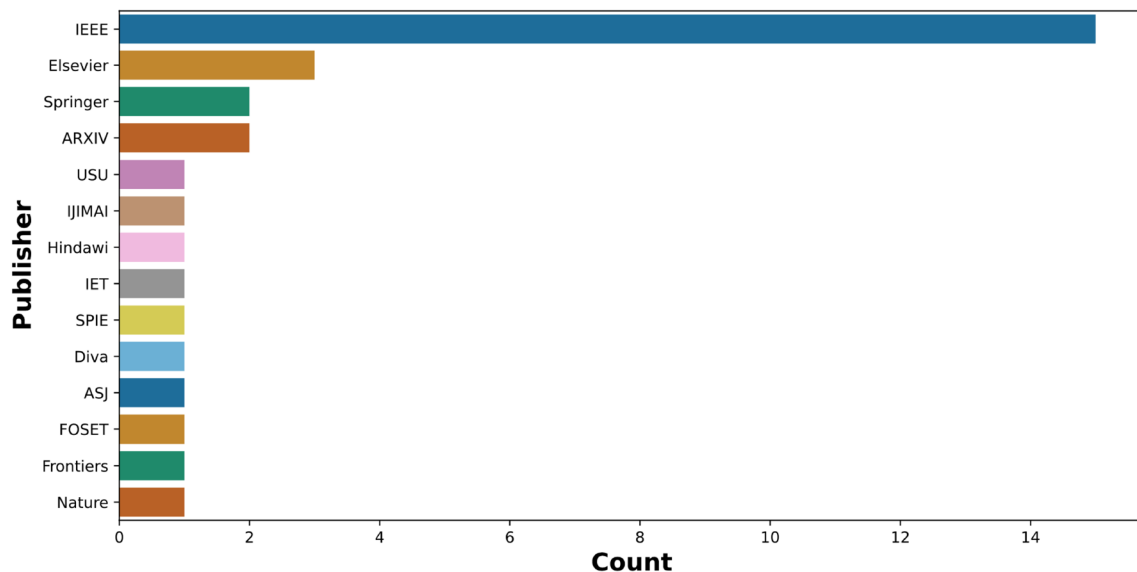
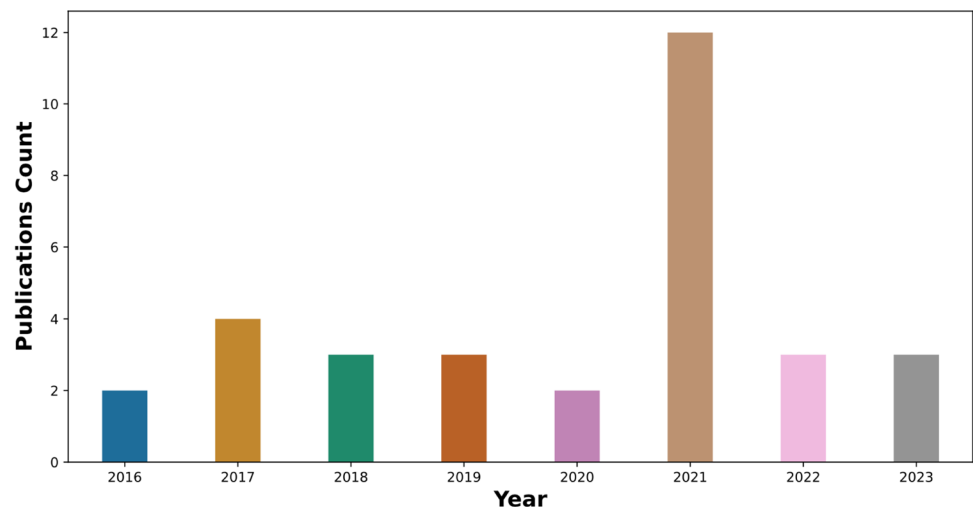


Fig. 3 A bar graph of the count of articles selected from different scholarly publishers

online repositories). This review notes that the majority of authors (Potamitis et al. 2007; Dong et al. 2018; Arpitha et al. 2021; Kim et al. 2021a; Zgank 2022; Terenzi et al. 2021; Basak et al. 2022; Nolasco et al. 2019; Phung et al. 2017; Noda et al. 2016; Bhattacharya et al. 2016; Zhang and Guo 2010) used secondary data to analyze and build their models. As shown in Table 5, the secondary data was sourced from various databases such as NU-HIVE (Cecchi et al. 2018), SINA (Thomas 2022), USDA (Mankin 2023), Kaggle repository (Potamitis 2023), OSBH (OSBeehives 2020), etc. Authors such as (Yazgaç et al. 2016; Silva et al. 2013; Fanioudakis et al. 2018; Ferreira et al. 2023; Hibino et al. 2021; Cruz et al. 2020) set up experiments in a laboratory or in outdoor settings to collect primary data.

Audio recording tools and audio editing software

To capture insect sounds accurately, appropriate audio recording tools are needed. We note that the widely used audio recording tool was the microphone (Amlathe 2018; Yazgaç et al. 2016; Arpitha et al. 2021; Terenzi et al. 2021; Nolasco et al. 2019; Kiskin et al. 2020; Hodzic and Hoang 2021; Várkonyi et al. 2023; Kawakita and Ichikawa 2019). After recording, the audio files can be cleaned (e.g., denoised, split, etc.) using software like Audacity (Audacity 2023), Adobe Audition (Kawakita and Ichikawa 2019) or Matlab (Kiskin et al. 2020) to clean up, edit and analyze the recorded sounds. Table 4 gives an overview of the

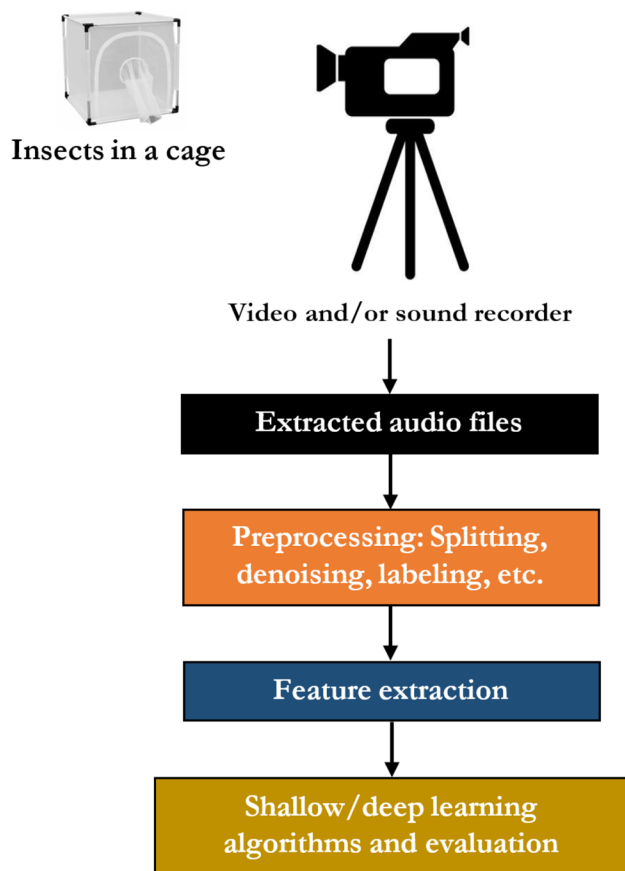


Fig. 4 A workflow of automatic synthesis of bioacoustics

audio recording tools and audio editing software used by various authors.

Pre-processing of insect songs

In this context, pre-processing refers to the cleaning of bio-acoustic recordings before they are subjected to analysis. Clean data contains quality and valuable information about the insects' vocalizations and improves data reliability making it suitable for accurate analysis and interpretation. From the literature, we discovered that pre-processing majorly consisted of 4 major steps: denoising, normalization, data labeling, and data splitting. These steps are covered in detail below.

a) Denoising Insects' songs are often recorded in natural environments, where background noise from wind, water, other animals, or human activity can be present. Reducing/removing the noise is critical to isolate and focus on the insect sounds of interest. By amplifying the insect sounds and attenuating background noise, researchers can improve the signal-to-noise ratio, making the calls more audible and easier to analyze (Amlathe 2018; Kim et al. 2021a; Silva et al. 2013). Various techniques can be used for noise removal such as wavelet transform technique (Amlathe 2018), audio editing software (Cruz et al. 2020; Kawakita and Ichikawa 2019; Varma et al. 2021).

Table 4 Overview of audio recording and audio editing software used by authors under review

Author	Audio recording tool	Audio editing software
Amlathe (2018)	Microphone embedded in electronic Beehive Monitoring System	Audacity software
Yazgaç et al. (2016)	Microphone with acoustics and ultrasonic sound levels	Not provided
Arpitha et al. (2021)	Microphone	Not provided
Kim et al. (2021a)	Audio sensors	Not provided
Terenzi et al. (2021)	Behringer UCA 22 sound card and ADMP 401 MEMS microphones	Not provided
Nolasco et al. (2019)	Microphone	Not provided
Silva et al. (2013)	Sensor	Not provided
Fanioudakis et al. (2018)	Optical sensors	Not provided
Ferreira et al. (2023)	Digital acoustic recorder	Not provided
Cruz et al. (2020)	Piezoelectric disc sensor	Audacity software
Hibino et al. (2021)	Recorder	Not provided
Kiskin et al. (2020)	Microphone embedded on low cost smartphones	MATLAB software
Hodzic and Hoang (2021)	Microphone embedded on Raspberry Pi	Not provided
Várkonyi et al. (2023)	Microphone	Not provided
Kawakita and Ichikawa (2019)	Microphone	Adobe Audition CC
Fernandes et al. (2021)	Microphone embedded in a smartphone	Audacity software
Parvathy et al. (2021)	Data Acquisition System (DAS)	Not provided
Varma et al. (2021)	Not provided	Audacity Software
Zamanian and Pourghassem (2017)	ZOOM-H4 audio portable recorder	Not provided

Table 5 Primary and secondary data source, data size, pre-processing steps, and splitting length

Study	Data source	Data size	Pre-processing	Splitting length
a) Primary data				
Amlathe (2018)	Recorded data	12600 audio samples, 9110 time-of-day samples of bee sounds	Denoising, data labeling	2 secs
Yazgaç et al. (2016)	Laboratory, outdoor data	118 Individual per each Sunn Pests class	Not provided	1 sec
Cruz et al. (2020)	Recorded data	5990 data samples of mango pulp weevil pupa sounds	Segmentation, data labeling	Not provided
Paim et al. (2023)	Recorded data	537 recordings of <i>A.e aegypti</i> mosquito species	Standardization, segmentation	Not provided
Ferreira et al. (2023)	Recorded data	518 audio samples of pollinating bee species	Data labeling, data splitting, data augmentation	Not provided
Fanioudakis et al. (2018)	Recorded data	279566 wing beat mosquito sounds	Sound splitting	Not provided
Hibino et al. (2021)	Recorded data	100 pieces for each type of the 7 cicada and cricket species	Segmentation, data augmentation	5 secs
Várkonyi et al. (2023)	Recorded data	10000 recordings of beehive sounds	Not provided	Not provided
Kawakita and Ichikawa (2019)	Recorded Data	200 samples of <i>A.mellifera</i> and <i>B.arden</i> sounds, 160 samples of <i>T.nipponensis</i> sounds, 120 samples of <i>V.s.xanthoptera</i> sounds, 200 recordings of background sounds and unspecified bird song	Not provided	Not provided
Kiskin et al. (2017)	Recorded data	6000-60000 training samples of mosquito sounds	Data annotation	2 secs
b) Secondary data				
Potamitis et al. (2007)	SINA	220 species of crickets and cicadas	Segmentation	Not provided
Zhang et al. (2021)	ARS	Not mentioned	Normalization, pre-emphasis	Not provided
Dong et al. (2018)	USDA	47 types of insect sounds, each including 1-10 records	Normalization, pre-accentuation, framing, windowing	Not provided
Arpitha et al. (2021)	Kaggle repository	1200 samples of mosquito wingbeats	Framing, windowing	50 milliseconds
Zgank (2022)	Open source beehives project	2443 recordings of bee swarm activities	Data labeling	Not provided
Terenzi et al. (2021)	NU-HIVE	345000 files of honey bees sounds	Segmentation, normalization	3 secs
Basak et al. (2022)	Sound library on the mobcup, quick sounds, pixabay websites	9 insect species of cicada, bees, crickets, mosquito, grasshopper, bugs, termites, beetles, ant	Segmentation, data labeling, framing, pre-emphasis	Not provided
Nolasco et al. (2019)	NU-Hive	576 files of beehive sounds	Not provided	
Phung et al. (2017)	USDA	66000 samples of insect sound activities	Pre-emphasis, normalization, framing	Not provided
Noda et al. (2016)	InsectSingers	88 cicada species sounds	Segmentation	Not provided
Kim et al. (2021b)	Open source beehives project	95 beehive sound files	Noise filtering, centering and zero-padding	Not provided
Wei et al. (2022)	Wingbeats, Abuzz	wingbeats dataset had 6 mosquito species of 3 different genera recordings, Abuzz had 20 mosquito species of 4 different genera recordings	Noise removal, data padding, segmentation, data augmentation	Not provided

Table 5 (continued)

Study	Data source	Data size	Pre-processing	Splitting length
Fernandes et al. (2021)	Abuzz	1285 audio files of 20 mosquito species	Denoising, normalization, segmentation, data augmentation	Not provided

c) Data labeling This refers to the process of annotating/tagging the insect song recordings with relevant information that machine-learning algorithms can learn from. Labeling provides the necessary context and ground truth for the machine learning algorithms to understand the relationship between the input (acoustic features) and the desired output (insect song synthesis). Some of the key information associated with data labeling in insect song synthesis are species (Zamanian and Pourghassem 2017), type of song according to wingbeats (Kiskin et al. 2020), background noise (Amlathe 2018), etc.

d) Sound signal splitting Due to the sound signal length, a long signal cannot be processed in one instant. For effective sound signal analysis in machine learning, the clean sound signals must be split into equal lengths. Different split lengths result in different sizes of features extracted and affect machine learning modeling as explained in Sections “Feature extraction” and “Machine learning in sound signal processing” respectively. In the literature reviewed, researchers have used different split lengths for their audio data. For instance, authors such as Kiskin et al. (2020) split mosquito-recorded sound signals to 2 seconds with an overlap of 1 second. Amlathe (2018) extracted a 2-second sample clip from the 30-second length of each bee sound record. Potamitis et al. (2007) split sound signals into 15 seconds each. Arpitha et al. (2021) split mosquito wingbeats into equal lengths of 50 milliseconds for each category. While most of the studies reviewed did not mention the sound-splitting ratio, other researchers emphasized the importance of sound splitting into equal lengths, e.g., Towsey (2017) and Weiping et al. (2017) split sound records into 4 seconds and 60 seconds clips respectively, Hibino et al. (2021) and Hodzic and Hoang (2021) split the audio files into 5-second log files. Other authors such as Yazgaç et al. (2016); Terenzi et al. (2021); Varma et al. (2021) segmented all the longer files into recordings of 1, 3, and 20 seconds respectively. Further, Hershey et al. (2017) recommends a splitting rate of at least 0.96s clips for a variety of audio detection tasks.

Generally, different split lengths give different feature matrices, and huge feature matrices make machine learning take more time to train the models. Table 5 summarizes the pre-processing steps and splitting length carried out by various studies.

Feature extraction

Feature extraction in acoustic analysis is the process of extracting mainly numerical and/or image values from sound signals for machine learning tasks. These features serve as input to the synthesis model, allowing it to learn characteristics of the insect songs which can be used to classify/identify insects based on their different species, wing beats etc. Effective feature extraction is crucial in determining the quality and accuracy of the synthesized songs. The choice of features depends on the specific characteristics of the insect songs and the requirements of the synthesis model. There are several libraries that can extract sound signal features e.g. Librosa (McFee et al. 2015) and Spafe (Malek 2023) libraries in Python programming languages, Sound Object (Lazzarini 2001) Library, or Essentia (Dmitry et al. 2013) in C++ programming language. Some of the common features which are extracted from the sound signals are the Mel-Scale Frequency Cepstral Coefficients (MFCC), chroma features, spectrogram, spectral centroid, and frequency-band energy features (Barchiesi et al. 2015). A summary of the features extracted in various studies is outlined in Table 6. Some of the common features identified are discussed below:

a) Mel-scale frequency cepstral coefficients (MFCC) feature MFCCs are widely used in audio signal processing and speech recognition. They capture the spectral characteristics of the sound and represent the short-term power spectrum of the audio signal after being transformed into Mel frequency scale (Le-Qing 2011). MFCCs are based on the human auditory system’s perception of sound, which makes them well-suited for various audio-related applications. This approach has been used by several studies (Amlathe 2018; Yazgaç et al. 2016; Phung et al. 2017; Noda et al. 2016; Zhang and Guo 2010; Silva et al. 2013; Kawakita and Ichikawa 2019) to extract sound signal features.

b) Chroma feature A chroma feature is a type of audio feature whose tonal variation of an audio signal is in a condensed form (Kattel et al. 2019). Kattel et al. (2019) notes that the tonal variations are represented by ‘color’ which refers to the quality of a pitch class, namely: C, C#, D, D#, E, F, F#, G, G#, A, A# and B. A chroma vector is represented by 12 bins of pitches which indicate the amount of energy

in each pitch class present in the signal. Thus, the 12 bins represent the magnitude of the 12 features/tone in a standard chromatic scale. Figure 5 is a visualization of the chroma feature of *Cacama valvata* (Kathy and David 2023) insect bioacoustics extracted using the Librosa (McFee et al. 2015) library. Amlathe (2018) explored this approach in insect bioacoustics. This has been explored in other domains such as in construction (Shi et al. 2019) and music genre (Shi et al. 2019) classification respectively.

c) Spectrogram feature This is a representation of 2D images with time along one axis representing sequences of spectra, with color brightness on the other axis representing the frequency of the strength of each component at each time frame. Spectrograms show where there is more or less energy, and how energy levels vary over time. In automatic insect song synthesis, spectrograms are common as they

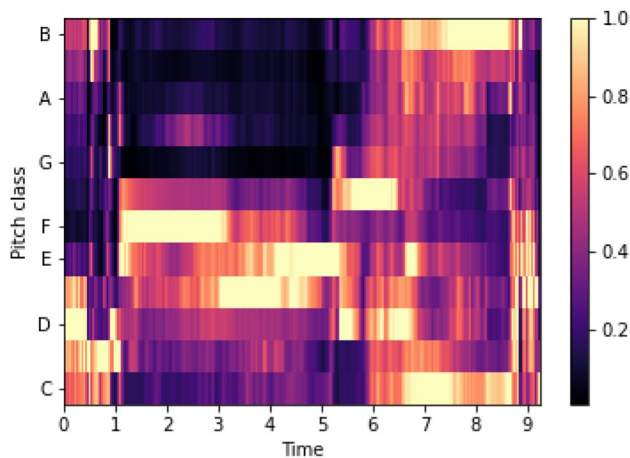
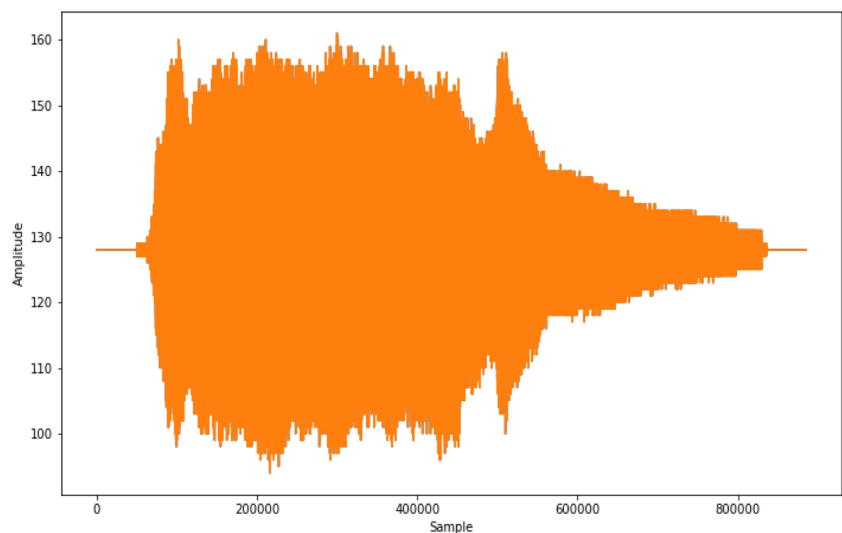


Fig. 5 Visualization of chroma features extracted from *Cacama valvata* (Kathy and David 2023) insect bioacoustics

Fig. 6 Spectrogram feature extracted from *Cacama valvata* (Kathy and David 2023) insect bioacoustics



capture the temporal and spectral characteristics of the insect sounds. Figure 6 is a visualization of a spectrogram image extracted using the Librosa Library. Several researchers (Dong et al. 2018; Arpitha et al. 2021; Terenzi et al. 2021; Nolasco et al. 2019; Kiskin et al. 2020; Karar et al. 2021) have used this approach to extract sound signal features.

d) Spectral centroid feature A spectral centroid is used to characterize the center mass of the spectral distribution of a sound. The spectral centroid is computed from the magnitude spectrum of an audio signal, typically obtained using the Short-time Fourier Transform (STFT). The *feature.spectral_centroid* function can be used to compute spectral centroid in Python programming language. The formula for computing the spectral centroid is as follows:

$$\text{SpectralCentroid} = \frac{\sum_k f(k) \cdot \text{Magnitude}(k)}{\sum_k \text{Magnitude}(k)} \quad (1)$$

Figure 7 is a visualization of spectral centroid features. Spectral centroid has not been widely used in insect bioacoustics analysis. However, this feature extraction method has been used by Várkonyi et al. (2023); Xie and Bertram (2019); Kavitha and Manikandan (2021) in their audio processing and classification tasks.

Machine learning in sound signal processing

Machine learning techniques are becoming increasingly popular in insects songs signal processing due to their ability to extract complex patterns, classify different sounds and insect species. Machine learning can be applied in species identification, bioacoustic monitoring, classification of call types, etc. Automatic sound/audio signal processing can be

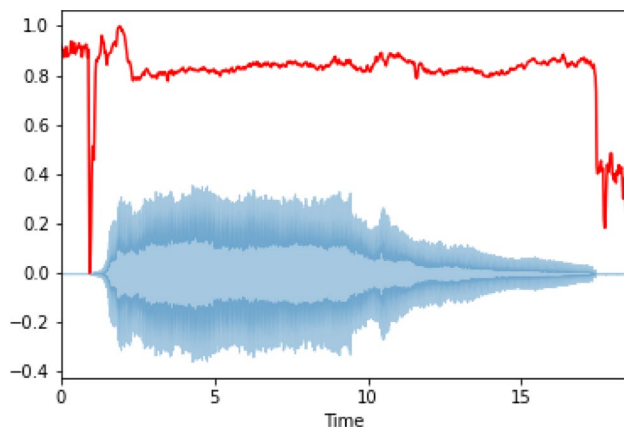


Fig. 7 Normalized spectral centroid features extracted from *Cacama valvata* (Kathy and David 2023) insect bioacoustics

improved using machine learning (Phung et al. 2017; Noda et al. 2016; Kawakita and Ichikawa 2019). The previous section discussed some approaches used in extracting features from sound signals. These features are then ingested in machine learning algorithms for processing. Section “[Applications of machine learning models on insects bioacoustic analysis](#)” describes various machine learning algorithms used by various authors in sound signals analysis.

Applications of machine learning models on insects bioacoustic analysis

Generally, machine learning algorithms fall into two broad categories, namely, shallow and deep machine learning methods. Shallow learning algorithms are algorithms such as random forest (RF), k-nearest neighbor (KNN), and Support Vector Machine (SVM), including neural networks that have only one hidden layer. While deep learning algorithms are those neural network models with multiple hidden layers. Examples of deep learning models are convolutional neural networks (CNN), Recurrent Neural Networks (RNNs), Long Short-Term Memory (LSTM), etc. Compared to shallow learning, deep learning algorithms require huge amounts of training data and subsequently give a better performance in various domains such as image and natural language processing (Yayin et al. 2021). Applications of these algorithms in insect bioacoustics are summarized in Table 6.

a) Shallow machine learning models Shallow machine learning falls under three major categories: supervised, unsupervised, and reinforcement learning. Supervised machine learning uses labeled data sets to train algorithms to classify categorical data or do regression on continuous data. Some of the algorithms used in supervised learning include logistic regression, perceptron, and KNN (Xu et al. 2021). In automatic insect bioacoustic synthesis, the widely

used supervised learning algorithms were SVM (Amlathe 2018; Kim et al. 2021a; Basak et al. 2022; Nolasco et al. 2019; Noda et al. 2016; Silva et al. 2013; Cruz et al. 2020; Robles-Guerrero et al. 2023), RF (Kim et al. 2021a; Ferreira et al. 2023; Robles-Guerrero et al. 2023) and KNN (Yazgaç et al. 2016; Basak et al. 2022; Phung et al. 2017; Xie and Bertram 2019; Robles-Guerrero et al. 2023). In all these cases, they have been used for call types or species classification. Unsupervised learning algorithms automatically decipher hidden patterns in unlabelled data and build clusters. Common unsupervised machine learning algorithms are K-means, principal component analysis (PCA), and Density-Based Spatial Clustering of Applications with Noise (DBSCAN); for instance, Hodzic and Hoang (2021) used DBSCAN to cluster bee buzz. Clustering is a means by which DBSCAN differentiates the bee buzz and groups the homogeneous instances into clusters on its own without supervision by a human. Reinforcement learning involves training computational algorithms based on rewards and punishments (Mitchell 1997). This review did not come across applications of reinforcement learning in automatic insect bioacoustic synthesis. The commonly used supervised algorithms are discussed below and Table 6 provides a summary of studies that applied them.

- i) **SVM:** SVM is effective in handling high dimensional data and can efficiently find a hyperplane that best separates different insect species or call types. This algorithm can be used for both binary and multi-class classification tasks. The algorithm can find a hyperplane that maximizes the margin between classes; this helps in achieving better generalization and robustness, especially when dealing with noisy or overlapping acoustic data. It handles both linear and non-linear boundaries by using different inbuilt kernel functions. This algorithm has been applied in various insect songs classification such as Amlathe (2018); Kim et al. (2021a); Basak et al. (2022); Nolasco et al. (2019); Phung et al. (2017); Ferreira et al. (2023); Cruz et al. (2020); Kawakita and Ichikawa (2019); Robles-Guerrero et al. (2023).
- ii) **RF:** The algorithm creates multiple decision trees from different random subsets of the data and features. To make a prediction, the algorithm averages the prediction of each decision tree, hence improving the predictive accuracy of the dataset. The algorithm has an internal mechanism for identifying and ranking the most important features that distinguish different insect species based on their bioacoustic signals. This algorithm has been used in various studies (Amlathe 2018; Kim et al. 2021a; Xie and Bertram 2019; Robles-Guerrero et al. 2023) for insect species classification and identification.

iii) KNN: This is a supervised algorithm used in classification tasks. The algorithm uses distance metrics to measure the similarity between feature vectors. KNN requires selecting a value for K, which is the number of nearest neighbors to consider when making predictions. This algorithm has been used in studies such as Amlathe (2018); Basak et al. (2022); Xie and Bertram (2019); Robles-Guerrero et al. (2023).

b) Deep learning models Insect sound detection and classification by use of deep learning is increasingly becoming popular. These models can learn intricate patterns and features from large and complex datasets, making them well-suited for tasks that involve analyzing audio signals. Furthermore, with advancements in technology, the models have been popular among researchers due to sufficient computing power for training these models. In insect bioacoustics, deep learning models are trained with spectrogram features which are extracted from insect sounds as images. Further, deep-learning models have been used in environmental sound classification to identify different sounds in the environment, thereby enhancing audio monitoring in smart devices. Deep learning models such as generative adversarial networks (GANs) and autoencoders have been used to clean up noisy audio, thereby enhancing the data quality for analysis. The majority of the insect audio data collected is not sufficient for analysis and model development due to the presence of noise which has to be removed. Deep learning models such as GANs can/have been used to generate synthetic data to complement the available data to ensure there is sufficient data for machine learning synthesis. Some of the common deep learning architectures used are multilayer perceptron (MLP), Convolutional Neural Networks (CNN), and Recurrent Neural Networks (RNNs) (Zhang et al. 2021; Dong et al. 2018; Arpitha et al. 2021; Zamanian and Pourghasem 2017; Várkonyi et al. 2023). These architectures are described below:

i) **MLP:** The MLP is a feed-forward neural network classification algorithm that is organized into 3 layers. The three layers are the input layer, the hidden layer, and the output layer. The input layer receives the input data, i.e., the extracted features from the bioacoustic signals of insects. Hidden layers contain multiple neurons, and the connections between them have associated weights. Hidden layers allow the network to capture complex patterns in the data. The output layer is the predicted class label, i.e. the label of the insect species or the insect song call type. Figure 8 illustrates the 3 layers of the MLP classifier and how they are interconnected (Heidari et al. 2020). This algorithm has been used by Kiskin et al. (2020).

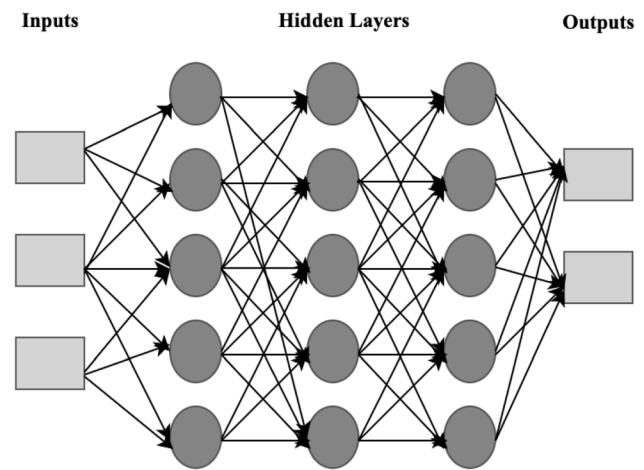
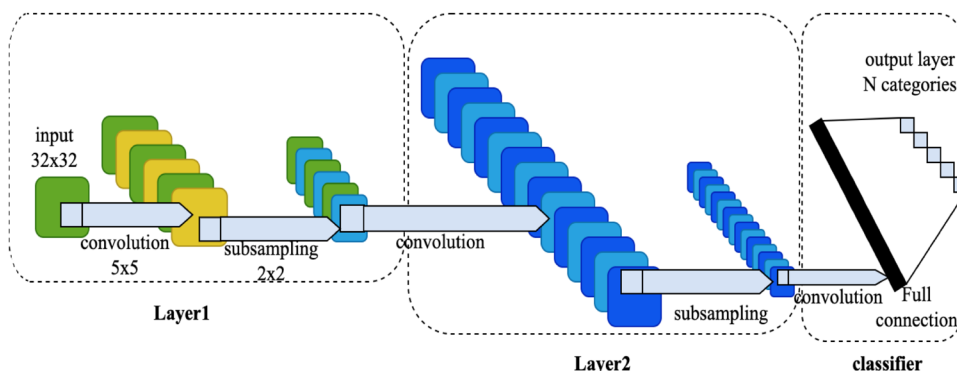


Fig. 8 Multilayer perceptron; as shown with 3 layers, input, hidden and output layer (Heidari et al. 2020)

ii) **CNN:** This is a specialized type of artificial neural network designed for processing grid-like data such as images, audio spectrograms, and time-series data. The algorithm works by assigning weights and biases to the input features based on the importance of various objects in the input image. CNN processes data in a 1D array and 2D array. The 1D array could include data such as speech and biomedical signals whereas the 2D array includes data such as images and audio spectrograms (Weiping et al. 2017). CNN has four important layers, (a) the input layer- this layer holds the input pixels of the image, (b) the convolutional layer - which captures different aspects of the input data such as edges, textures, and shapes, (c) the pooling layer- downsamples the spatial dimensions of the feature maps obtained from the convolutional layers, reducing the number of parameters within that activation, (d) fully connected layers - combine the learned features and make high-level predictions based on the extracted features (O'Shea and Nash 2015). Figure 9 is a diagrammatic flow of the working of CNN. This algorithm has been used successfully in studies such as Terenzi et al. (2021); Nolasco et al. (2019); Parvathy et al. (2021); Varma et al. (2021); Paim et al. (2023); Trivedi and Shroff (2021).

b) Transfer Learning Transfer learning entails the reuse of pre-trained machine learning models to solve new problems. Pre-trained machine learning models are models that have been trained on large datasets but can be fine-tuned to solve specific tasks. Transfer learning is becoming popular in deep learning tasks since it can be fine-tuned with relatively small amounts of data. These make them quite relevant in most real-world problems that typically do not have huge amounts of data points to train such complex models.

Fig. 9 Architecture of the CNN (Jiuxiang et al. 2018)



Pre-trained models are diverse depending on the kind of task you want to solve (Zhuang et al. 2021; Karl et al. 2016). For instance, in the domain of insect bioacoustics, as discussed in Section “Feature extraction”, sound signal features can be extracted as spectrogram or spectral centroid images. These images can be trained using various transfer-learning pre-trained models such as VGG16 (Qassim et al. 2018), VGG19 (Simonyan and Zisserman 2015), RESNet152 (Kora Venu 2021), Xception (Sharma and Kumar 2022), etc. Notably, Torky et al. (2023); Boulila et al. (2023) converted sound data of the Red Palm Weevil insect into images and used deep learning pre-trained models to classify the infestation (or no infestation) of the insects on the palm trees.

Evaluation metrics for insects songs synthesis

To evaluate the performance of different machine learning algorithms, various evaluation metrics have been used across different studies. The metrics are described below and Table 6 summarises the various evaluation metrics used by different studies in the automatic classification of insect acoustics using machine learning algorithms.

a) Accuracy It measures the proportion of correctly classified instances in the entire dataset. This evaluation metric is a straightforward and intuitive metric that provides an overall measure of how well the model performs on the classification task. Accuracy can be computed from the confusion matrix, which is a tabular representation of the performance of a classification model. In the confusion matrix, true positive (TP) represents the number of instances correctly classified as the positive class, false positive (FP) is the number of instances incorrectly classified as the positive class, false negative (FN) is the number of instances incorrectly classified as the negative class and true negative (TN) is the other class identified as the other class. Figure 10 shows the

	Predicted Label	
True Label	TP	FN
	FP	TN

Fig. 10 Species classification on 10-class insect recognition dataset

described variables in a confusion matrix table while (2) describes how accuracy is calculated from the confusion matrix.

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad (2)$$

b) Precision This is calculated as the ratio of the number of samples correctly predicted as the class to the total number of samples in the class as indicated in (3). Precision is used to evaluate a model’s ability to correctly identify the class.

$$Precision = \frac{TP}{TP + FP} \quad (3)$$

c) Recall This also known as sensitivity or true positive rate (TPR) measures the ability of a model to correctly identify all positive instances out of the total positive instances as indicated in (4).

$$Recall = \frac{TP}{TP + FN} \quad (4)$$

d) F1-score F1-score is an evaluation metric commonly used to evaluate imbalanced datasets, where some classes have a much larger instance than others. F1-score combines both precision and recall to achieve a balance of both. The F1-score is calculated using the formula in (5). When dealing with multi-class classification, you can compute the

Table 6 Summary of authors, features, machine learning classifier algorithm, and evaluation metrics

Author	Features	Classifier	Evaluation Metrics
a) Deep Learning			
Wei et al. (2022)	Spectrograms	ResNet	Accuracy, precision, recall, F1-score
Fernandes et al. (2021)	Spectrograms	CNN	Accuracy, precision, recall, F1-score
Paim et al. (2023)	Spectrograms	Residual CNN	Accuracy, precision, recall, F1-score
Trivedi and Shroff (2021)	Melfilter bank energy	Artificial neural networks	Accuracy
Robles-Guerrero et al. (2023)	MFCC	Neural Network	Confusion matrix, accuracy, precision, recall, F1-score, ROC curve, misclassification
Várkonyi et al. (2023)	STFT, chroma, MFCC, MFCC delta, spectral centroid, zero-crossing rate	Recurrent neural networks (long short-term memory networks)	Accuracy
Karar et al. (2021)	Spectrograms	InceptionResNetV2	Accuracy, precision, recall, F1-score
Nolasco et al. (2019)	Mel spectrograms	CNN	AUC
Parvathy et al. (2021)	Mel spectrogram	Convolutional autoencoder	Mean squared error, F1-score, ROC curve, AUC value, accuracy, precision, recall
Varma et al. (2021)	Mel-spectrogram, 1-D convolution	2D-CNN, 1D-CNN, SincNet	Accuracy, precision, recall, F1-score
Terenzi et al. (2021)	STFT spectrograms, Mel spectrograms, MFCC, Hilbert Huang transform (HHT), Continuous Wavelet Transform (CWT), Discrete Wavelet Transform (DWT)	CNN	AUC, precision, recall, F1 measure
Hibino et al. (2021)	Spectrogram (mel spectrogram, power spectrogram)	CNN	Accuracy
Cruz et al. (2020)	MFCC	Artificial neural network	Accuracy
Zhang et al. (2017)	Spectrum images	Deep CNN	Accuracy
Fanioudakis et al. (2018)	Spectrogram	CNN	Accuracy
Kim et al. (2021b)	MFCC, CQT, Mel spectrogram	CNN	Accuracy
Kiskin et al. (2017)	STFT, wavelet spectrogram	MLP, CNN	F1-score, ROC area, precision, recall
Ferreira et al. (2023)	Log mel-spectrogram	CNN	Accuracy, precision, recall, F1-score
Zamanian and Pourghassem (2017)	Spectral and temporal	MLP	Accuracy
Dong et al. (2018)	Enhanced spectrograms, chromatic spectrograms, MFCC	CNN	Accuracy
Arpitha et al. (2021)	Spectrogram images	CNN	Precision, recall, F1-score
Zhang et al. (2021)	MFCC	CNN	Accuracy
Kim et al. (2021a)	MFCC	CNN	Accuracy
Kiskin et al. (2020)	Spectrograms	CNN	Confusion matrix
b) Shallow Learning			
Amlathe (2018)	MFCC, chroma-stft, spectral contrast, tonnetz, Mel spectrogram, STFT	RF, SVM, KNN, Logistic Regression	Accuracy
Yazgaç et al. (2016)	LFCC, MFCC, Linear Spectral Frequencies (LSF)	KNN	confusion matrix, precision, recall, accuracy

Table 6 (continued)

Author	Features	Classifier	Evaluation Metrics
Xie and Bertram (2019)	Temporal features (zero-crossing rate, short-time energy, entropy), frequency features (MFCC, spectral centroid, spectral roll-off, spectral flux, spectral entropy)	KNN, Logistic Regression, Gaussian naive Bayes, random forest	F1-score
Nolasco et al. (2019)	MFCC, Hilbert Huang Transform (HHT)	SVM	AUC score
Cruz et al. (2020)	MFCC	SVM	Accuracy
Kawakita and Ichikawa (2019)	MFCC	SVM	Precision, recall
Hodzic and Hoang (2021)	MFCC	Density-based spatial clustering of applications with noise (DBSCAN)	Calculate the number of data points of the deviating bee buzz for each cluster observed during the analysis
Zgank (2022)	MFCC	Hidden Markov Models (HMM)	Accuracy
Basak et al. (2022)	MFCC	SVM, KNN, decision tree	Accuracy
Ferreira et al. (2023)	MFCC	SVM, RF, decision tree, ensemble	Accuracy, precision, recall, F1-score
Phung et al. (2017)	LL, MFCC	KNN, SVM, bagged tree	Accuracy
Noda et al. (2016)	LFCC, MFCC	SVM	Accuracy, F1-score
Potamitis et al. (2007)	LFCC	GMM, PNN	Accuracy
Zhang and Guo (2010)	MFCC	HMM combined GMM	Accuracy
Silva et al. (2013)	LPC, MFCC, LFCC, LSF, temporal, spectral features	SVM	Accuracy
Kim et al. (2021b)	MFCC, Mel spectrogram, Constant-Q-Transform (CQT),	SVM, RF, XGBoost	Accuracy

F1score separately for each class or take the macro-average or micro-average of F1scores across all the classes to measure the model's performance. The macro-average gives equal weight to each class, while the micro-average considers the overall performance across all instances in the dataset. The choice of averaging method depends on the classification task's specific requirements and the desired evaluation strategy.

$$F1 - score = 2 * \frac{Precision * Recall}{Precision + Recall} \quad (5)$$

e) Receiver operating characteristic (ROC) and area under the curve (AUC) The ROC and AUC are evaluation tools used to assess the performance of binary classification models. The AUC quantifies the overall performance of the model by measuring the model's ability to correctly classify positive instances of TPR while minimizing the FPR. The AUC value ranges from 0 to 1, where an AUC of 0.5 indicates random performance and an AUC of 1 indicates perfect classification. The ROC curve is created by plotting the TPR against the FPR and shows the trade-off between the TPR and FPR

as the classification threshold varies. Equations (6) and (7) show how TPR and FPR are calculated.

$$TPR = \frac{TruePositives}{TruePositives + FalseNegatives} \quad (6)$$

$$FPR = \frac{FalsePositives}{FalsePositives + TrueNegatives} \quad (7)$$

Summary: bioacoustics automatic synthesis

The summary can be outlined in point form as follows:

- i) Secondary data were used by most studies compared to primary data. The microphone sensor was the most preferred device for recording insects' bioacoustics; refer to Table 5.
- ii) Different pre-processing techniques were used to prepare bioacoustics data for machine learning applications. The common sound signal pre-processing steps were denois-

ing, segmentation, framing, and windowing; refer to Table 5.

- iii) A majority of authors used deep learning algorithms to process spectrogram images of insects' bioacoustics signals. It was seen that CNN was the widely used deep learning algorithm for image processing; refer to Table 6.
- iv) Some researchers used shallow machine-learning algorithms to process insect bioacoustic signals in the form of numerical data. The most common numerical feature was the MFCC since it partly mimics the human perception of sound color and thus makes it more relevant in sound and signal analysis (Montalvão and Araujo 2012). SVM was the most commonly used algorithm to process numerical features since it gave better performance metric results; refer to Table 6a.
- v) The widely used performance metric was accuracy for both shallow and deep learning algorithms; refer to Table 6.

Challenges and future research directions: bioacoustics in machine learning

While the automatic synthesis of bioacoustics has achieved considerable success, a number of challenges were identified in this review.

- i) As shown in Table 5, a majority of authors have used secondary data since some of the data are open access. However, secondary data often lacks originality and may not align with specific research needs. Additionally, it may be incomplete or inaccurate depending on the primary data collection protocol (Baldwin et al. 2022). The use of secondary data can be attributed to the cost implication of having an audio sensor, time factor, and human personnel required to collect primary data collection.
- ii) Pre-processing bioacoustics signals is labor-intensive and time-consuming. Bioacoustics pre-processing procedures such as denoising can be automated. However, Amlathe (2018) noted automating denoising using algorithms smoothens the sound signals to the extent that it distorts the original data, especially the low-frequency sounds. Therefore, more research can be invested to improve these approaches.
- iii) A majority of authors discussed in detail the machine learning algorithms used but scanty and at times no information is provided on the algorithms deployment strategies used. This, therefore, makes it hard to determine the resources required as part of the algorithm's processing power and memory space required to gauge the effectiveness of the algorithm's performance.

- iv) As shown in Table 5, a few authors mention the final size of the dataset that was used to build the intelligent models. As such, this review found it difficult to evaluate the fit of the respective models to the respective tasks. The study noted that the size of datasets is very key in determining whether the model is overfitting or underfitting.
- v) Authors do not provide the reasoning behind the choice of their different evaluation metrics. This is very important as it will explain the nature of the dataset.
- vi) Insect acoustics research programming outputs are often in an executable form which hinders experiments reproducibility. However, it could be more valuable if the scripts were made accessible to the open science community. By sharing these tools/scripts, researchers could improve/innovate the models and help address specific social challenges more effectively.

The following are some recommendations that can be pursued by researchers in future works:

- i) Research on the ecological significance of bioacoustics synthesis by studying how natural weather (e.g., temperature and humidity) variables affect sounds produced by insects. This can bring forth a better understanding of insects in natural ecological setups.
- ii) Insects produce songs in different tonal variations. This variation can also be captured in chroma features. Classification of sound signals using chroma features can be explored. This can be compared with the commonly used MFCC features. The results can also be compared to the spectral bioacoustics features synthesized using deep learning approaches.
- iii) Future studies can explore sound signal feature combinations (e.g., Chroma+MFCC, LFCC+LSF, etc.) to improve the performance of machine learning models.
- iv) Future studies to develop user interfaces to bridge the gap between the development of solutions and making them available in production to be used by the end user. This will make the solutions more meaningful and impactful.

Automatic synthesis of insects bioacoustic: current and future prospects

Automatic synthesis of insect bioacoustics has been applied considerably in detecting and monitoring insect behavior. In this section, we will discuss applications of automatic synthesis of bioacoustics in pollinators and pest management services for purposes of biodiversity preservation, and in optimizing insects as food and feed

Production of insects as food and/or feed

A variety of insects produce sounds to signify specific phenomena and/or behavior. These sounds automatically be learned using machine learning to improve the health of the edible insects to improve their production for food and feed. For instance, the edible cricket, *Gryllus bimaculatus*, produces sounds to signify specific behavior/activities (e.g., calling, aggression, courtship, etc.) (Lin and Hedwig 2021; Miyashita et al. 2016; Alexander 1961). These sounds can inform the behavior/health of the crickets and can be automatically synthesized to improve their production as food for humans and feed for animals (such as poultry, fish, etc.). The edible palm weevils, *Rhynchophorus* spp. (Coleoptera: Curculionidae) is not only a devastating invasive pest of palms but it is also a traditional and sustainable source of food in many communities (Egonyu et al. 2021). When the palm weevils are feeding on the palm trees, they produce sounds that inform their early detection after infesting palm trees (Torky et al. 2023; Boulila et al. 2023). These sounds can be classified to inform farmers about the collapse of the colony, diseases, and/or harvesting for purposes of optimizing production. The black soldier fly (BSFL), *Hermetia illucens*, is invaluable for organic waste recycling and as food for animals. Physiological changes in relation to the surrounding noise effects showed that high-frequency sound, at or above 85 dB, acts as a stress and negatively affects the development of the black soldier fly in terms of pupation and survival of the adults (Park et al. 2013). Classifying and optimizing these surrounding noises/sounds can optimize the production of BSFL for feed purposes.

Pollinator conservation

Pollination is one of the most important mechanisms in ecological conservation that promotes biodiversity and the preservation of life on Earth. The overall importance of pollinators' health and their crucial service in supporting diversity in many habitats, such as agroecosystems, cannot be overstated. In essence, pollinators are at the center of facilitating reproduction in more than 80% of the world's flowering plants in the fragile agroecosystems fabric that is compounded by many challenges including climate change and pollution. Without a doubt, insects are the largest and most efficient pollinators and their habitats have rich ecological, economic, cultural, and social benefits. Bees are the most important group of pollinating insects (Kaur and Kaleka 2022; Khalifa et al. 2021; Baldock 2020).

The health of pollinators such as bees is paramount in securing food supply and a stable economy. However, bee populations have dwindled due to colony collapse disorder and unsustainable farming practices (Amlathe 2018). Agricultural interventions like frequent hive inspections

negatively impact honey production (Várkonyi et al. 2023). Sensors and machine learning approaches have been employed to understand bee acoustics in relation to their health with minimal disturbance to their natural activities in the habitat (Hodzic and Hoang 2021; Kim et al. 2021b). This has automated a) bee hive management and health (Amlathe 2018; Nolasco et al. 2019; Kim et al. 2021b; Abdollahi et al. 2022; Terenzi et al. 2020), b) classification of different bee species (Ferreira et al. 2023), c) differentiation of bee predators (Hodzic and Hoang 2021), d) uncovering insights into bee colony health and queen activity (Uthoff et al. 2023), e) studies of bees to detect air pollutants (Zhao et al. 2021), f) predicting swarming in honeybee colonies (Zgank 2022, 2020; Michael-Thomas et al. 2020) studying bee pollination services for sustainable tomato cultivation (Ribeiro et al. 2021), g) distinguishing of bees from hornets (a predator) (Kawakita and Ichikawa 2019), etc. These have promoted the sustainability, diversity, and ecological conservation of life on Earth.

Pest management

Integrated pest management (IPM) is a comprehensive and sustainable approach to pest management that aims to effectively control pest populations while minimizing the negative impact on the environment, human health, and non-target organisms. IPM takes into account the complex interactions between pests, crops, natural predators of pests/crops, and the surrounding ecosystem, and seeks to strike a balance between pest control and its habitat to promote overall ecological health and diversity (Marinko et al. 2023; Muniz-Junior et al. 2023).

Accurate identification of insect species holds crucial importance. For example, appropriate pesticides can be applied accurately after knowing the corresponding insect it will neutralize. This aids in preventing environmental pollution and resource wastage due to improper pesticide use (Dong et al. 2018; Zamanian and Pourghassem 2017). Phung et al. (2017) investigated insect detection through activity-generated sounds for purposes of pest management. Noda et al. (2016) employed intelligent systems like SVM for insect classification to enhance pest control strategies. Zhang and Guo (2010) focused on discerning the identification of destructive fruit fly strains affecting fruits and vegetables through their vibrational sounds; outcomes of the experiments highlighted the practicality and efficiency of proactively detecting these fruit flies, offering a potential approach for averting damage to vegetables and fruits. In the context of red palm weevils, Karar et al. (2021) proposed an IoT-based sound detection method to aid in quarantining infested date palm trees and preserving them; date palm trees are a source of food and also one of the major sources of income in Arab countries, hence the pressing need for

their preservation. Yazgaç et al. (2016) addressed detecting sunn pests (notorious for cereal crop damage) using signal processing enabling effective population monitoring and reduced insecticide usage. Intelligent traps that automatically differentiate and trap specific insects to preserve biodiversity have been developed (Silva et al. 2013). Intelligent traps can selectively target harmful pests while sparing beneficial or neutral insects. This minimizes ecological disruption. The targeted approach reduces reliance on broad-spectrum pesticides, protecting biodiversity and supporting ecosystem health. Precise identification of insect species using machine learning can guide in controlling pests such as cicadas (Zamanian and Pourghassem 2017) and mango pulp weevil (Cruz et al. 2020).

Mosquito-borne diseases like malaria, zika, and dengue are significant health threats to humans (Kim et al. 2021a). By studying mosquito wingbeat sounds, Kim et al. (2021a); Kiskin et al. (2020, 2017); Trivedi and Shroff (2021) demonstrated machine learning's potential to identify mosquitoes hence a step into automated approaches to vector and disease management. Moreover, Wei et al. (2022) employed ResNet to classify different mosquito species by studying wing-beat sounds; different mosquito species transmit different diseases, therefore, this technique facilitates targeted disease prevention strategies (Trivedi and Shroff 2021). Different automatic bioacoustic synthesis approaches classify the *Aedes aegypti* mosquitoes species and created acoustic tools to monitor and manage vector-borne diseases (Fernandes et al. 2021; Paim et al. 2023).

Conclusion

This article provides a comprehensive review of methodological approaches in the automatic (using machine learning) synthesis of insect bioacoustics and their application in biodiversity and ecological conservation, pest management, and insects as food and/or feed. We found that the basic device for collecting bioacoustics signals was the microphone sensor. Most researchers used secondary data to build machine learning models. Different approaches (such as denoising, segmentation, windowing, framing, and splitting) were used to pre-process and transform the bioacoustics signals to be machine-learning-ready. MFCC was the most extracted numerical feature from bioacoustics signals while SVM was the most commonly used shallow machine learning algorithm. Spectral image features were also extracted from bioacoustics signals to build deep-learning models, and the CNN algorithm was the deep-learning model used by many researchers. Automatic synthesis of insect bioacoustics has been successful and has a promising future in the areas such as pollinator management, and integrated pest management/

preservation. These approaches can be applied to the conservation of biodiversity and ecological conservation.

Funding The authors gratefully acknowledge the financial support for this research by the following organizations and agencies: Kenya Education Network Trust (KENET), grant number AGMT1170; Australian Centre for International Agricultural Research (ACIAR) (ProteinAfrica: LS/2020/154); Novo Nordisk Foundation (RefIPro: NNF22SA0078466), IKEA Foundation (G-2204-02144); the Rockefeller Foundation (WAVE-IN: 2021 FOD 030); Bill and Melinda Gates Foundation (INV-032416); European Commission (NESTLER Project: 101060762 and INNOECOFOOD project: 101136739), the Curt Bergfors Foundation [Food Planet Prize Award]; Norwegian Agency for Development Cooperation, the Section for Research, Innovation, and Higher Education grant number RAF–3058 KEN–18/0005 (CAP–Africa); the Swedish International Development Cooperation Agency (Sida); the Swiss Agency for Development and Cooperation (SDC); the Australian Centre for International Agricultural Research (ACIAR); the Government of Norway; the German Federal Ministry for Economic Cooperation and Development (BMZ); and the Government of the Republic of Kenya. The views expressed herein do not necessarily reflect the official opinion of the donors. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript. The views expressed herein don't necessarily reflect the official opinion of the donors.

Declarations

Conflict of Interest The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Abdollahi M, Giovenazzo P, Falk TH (2022) Automated beehive acoustics monitoring: A comprehensive review of the literature and recommendations for future work. *Appl Sci* 12(8). <https://doi.org/10.3390/app12083920>
- Aide TM, Corrada-Bravo C, Campos-Cerqueira M, Milan C, Vega G, Alvarez R (2013) Real-time bioacoustics monitoring and automated species identification. *PeerJ* 1:e103. <https://doi.org/10.7717/peerj.103>
- Alexander RD (1961) Aggressiveness, territoriality, and sexual behavior in field crickets (orthoptera: Gryllidae). *Behaviour* 130–223. <https://doi.org/10.1163/156853961X00042>
- Alonso JB, Cabrera J, Shyamnani R, Travieso CM, Bolaños F, García A, Villegas A, Wainwright M (2017) Automatic anuran identification using noise removal and audio activity detection. *Expert Syst Appl* 72:83–92. <https://doi.org/10.1016/j.eswa.2016.12.019>

- Amlathe P (2018) Standard machine learning techniques in audio bee-hive monitoring: Classification of audio samples with logistic regression, K-nearest neighbor, random forest and support vector machine. Ph. D. thesis, Utah State University
- Arpitha M, Rani SK, Lavanya M (2021) CNN based framework for classification of mosquitoes based on its wingbeats. In: 2021 Third international conference on intelligent communication technologies and virtual mobile networks (ICICV), pp 1–5. IEEE
- Astrid J (2018) The importance of insects in agricultural ecosystems. *Adv Entomology* 6(2):1. <https://doi.org/10.4236/ae.2018.62006>
- Audacity (2023) Audio and video editing software
- Auerswald L, Lopata A (2005) Insects-diversity and allergy. *Curr Allergy Clin Immunol* 18(2):58–60
- Baldock KC (2020) Opportunities and threats for pollinator conservation in global towns and cities. *Curr Opin Insect Sci* 38:63–71. <https://doi.org/10.1016/j.cois.2020.01.006>
- Baldwin JR, Pingault JB, Schoeler T, Sallis HM, Munafò MR (2022) Protecting against researcher bias in secondary data analysis: challenges and potential solutions. *Eur J Epidemiol* 37(1):1–10. <https://doi.org/10.1007/s10654-021-00839-0>
- Barchiesi D, Giannoulis D, Stowell D, Plumbley MD (2015) Acoustic scene classification: Classifying environments from the sounds they produce. *IEEE Signal Process Mag* 32(3):16–34
- Basak S, Ghosh A, Saha D, Dutta C, Maiti A (2022) Insects sound classification with acoustic features and k-nearest algorithm. PREPARE@u@I FOSET Conferences. https://doi.org/10.36375/prepare_u.foset.a291
- Bhattacharya S, Basu P, Sajal Bhattacharya C (2016) The southern house mosquito, *Culex quinquefasciatus*: profile of a smart vector. *J Entomol Zool Stud* 4(2):73–81
- Boulila W, Alzahem A, Koubaa A, Benjdira B, Ammar A (2023) Early detection of red palm weevil infestations using deep learning classification of acoustic signals. *Comput Electron Agri* 212:10815. <https://doi.org/10.1016/j.compag.2023.108154>
- Cecchi S, Terenzi A, Orcioni S, Riolo P, Ruschioni S, Isidoro N (2018) A preliminary study of sounds emitted by honey bees in a bee-hive. In: Audio Engineering Society Convention 144. Audio Engineering Society
- Crespo-Pérez V, Kazakou E, Roubik DW, Cárdenas RE (2020) The importance of insects on land and in water: a tropical view. *Curr Opin Insect Sci* 40:31–38. <https://doi.org/10.1016/j.cois.2020.05.016>
- Cruz JCD, Calalay MAA, Gonzaga MRA, Medina PJG, Banlawe IAP (2020) Characterization of mango pulp weevil pupa using acoustics and machine learning. In: 2020 IEEE 12th International conference on humanoid, nanotechnology, information technology, communication and control, environment, and management (HNICEM), pp 1–6. IEEE
- Dmitry B, Nicolas W, Emilia G, Sankalp G, Perfecto H, Oscar M, Gerard R, Justin S, José ZXS (2013) Essentia: an open-source library for sound and music analysis. In: Proceedings of the 21st ACM international conference on multimedia
- Dong X, Yan N, Wei Y (2018) Insect sound recognition based on convolutional neural network. In: 2018 IEEE 3rd International conference on image, vision and computing (ICIVC), pp 855–859. IEEE
- Eggleton P (2020) The state of the world's insects. *Ann Rev Environ Resour* 45:61–82. <https://doi.org/10.1146/annurev-envir-012420-050035>
- Egonyu J, Gitonga K, Khamis F, Copeland R, Finyange P, Odhiambo R, Ddamulira G, Tanga C, Subramanian S (2021) Trapping, identification and rearing of edible palm weevils in kenya and uganda. *J Insects Food Feed* 7(8):1243–1253. <https://doi.org/10.3920/JIFF2021.0018>
- Fanioudakis E, Geismar M, Potamitis I (2018) Mosquito wingbeat analysis and classification using deep learning. pp 2410–2414
- Fernandes MS, Cordeiro W, Recamonde-Mendoza M (2021) Detecting aedes aegypti mosquitoes through audio classification with convolutional neural networks. *Comput Biol Med* 129:104152. <https://doi.org/10.1016/j.combiomed.2020.104152>
- Ferreira AIS, da Silva NFF, Mesquita FN, Rosa TC, Monzón VH, Mesquita-Neto JN (2023) Automatic acoustic recognition of pollinating bee species can be highly improved by deep learning models accompanied by pre-training and strong data augmentation. *Front Plant Sci* 14
- Folliot A, Hauptert S, Ducret M, Sèbe F, Sueur J (2022) Using acoustics and artificial intelligence to monitor pollination by insects and tree use by woodpeckers. *Sci Total Environ* 838:155883. <https://doi.org/10.1016/j.scitotenv.2022.155883>
- Ghisbain G, Gérard M, Wood TJ, Hines HM, Michez D (2021) Expanding insect pollinators in the anthropocene. *Biol Rev* 96(6):2755–2770. <https://doi.org/10.1111/brev.12777>
- Hallmann CA, Sorg M, Jongejans E, Siepel H, Hofland N, Schwan H, Stenmans W, Müller A, Sumser H, Hörrn T et al (2017) More than 75 percent decline over 27 years in total flying insect biomass in protected areas. *PloS One* 12(10):e0185809. <https://doi.org/10.1371/journal.pone.0185809>
- Harris JE, Rodenhous NL, Holmes RT (2019) Decline in beetle abundance and diversity in an intact temperate forest linked to climate warming. *Biol Conserv* 240:108219. <https://doi.org/10.1016/j.biocon.2019.108219>
- Heidari AA, Faris H, Mirjalili S, Aljarah I, Mafarja M (2020) Ant lion optimizer: theory, literature review, and application in multi-layer perceptron neural networks, pp 23–46. Cham: Springer International Publishing
- Heller KG, Ingris S, Liu CX, Shi FM, Hemp C, Warchałowska-Śliwa E, Rentz DC (2017) Complex songs and cryptic ethospecies: the case of the ducetia japonica group (orthoptera: Tettigoniidae: Phaneropteridae: Phaneropterinae). *Zool J Linn Soc* 181(2):286–307. <https://doi.org/10.1093/zoolinnean/zlw019>
- Hermosilla G, Pizarro F, Fingerhuth S, Lazcano F, Santibanez F, Baker N, Castro D, Yáñez C (2021) Real-time remote sensing of the lobesia botrana moth using a wireless acoustic detection sensor. *Appl Sci* 11(24). <https://doi.org/10.3390/app112411889>
- Hershey S, Chaudhuri S, Ellis DP, Gemmeke JF, Jansen A, Moore RC, Plakal M, Platt D, Saurous RA, Seybold B, et al (2017) CNN architectures for large-scale audio classification. In: 2017 IEEE international conference on acoustics, speech and signal processing (ICASSP), pp 131–135. IEEE
- Hibino S, Suzuki C, Nishino T (2021) Classification of singing insect sounds with convolutional neural network. *Acoust Sci Technol* 42(6):354–356
- Hodzic A, Hoang D (2021) Detection of deviations in beehives based on sound analysis and machine learning
- Horwood M, Nahrung HF, Fitzgerald C, Carnegie A (2022) Insect pests of timber-in-service: an Australian review. *Aust For* 85(4):199–210
- Jiuxiang G, Zhenhua W, Jason K, Lianyang M, Amir S, Bing S, Ting L, Xingxing W, Li W, Gang W, Jianfei C, Tsuhan C (2018) Recent advances in convolutional neural networks. *Pattern Recognit* 77
- Kampichler C, Wieland R, Calmé S, Weissenberger H, Arriaga-Weiss S (2010) Classification in conservation biology: a comparison of five machine-learning methods. *Ecol Inform* 5(6):441–450. <https://doi.org/10.1016/j.ecoinf.2010.06.003>
- Karar ME, Reyad O, Abdel-Aty AH, Owyed S, Hassan MF (2021) Intelligent IOT-aided early sound detection of red palmweevils. *Comput Mater Contin* 69(3):4095–4111
- Karl W, Taghi KM, Ding Ding W (2016) A survey of transfer learning. *J Big Data* 3. <https://doi.org/10.1186/s40537-016-0043-6>
- Karunamoorthi K, Ilango K, Murugan K (2010) Laboratory evaluation of traditionally used plant-based insect repellent against the malaria vector anopheles arabiensis patton (diptera: Culicidae). *Parasitol Res* 106:1217–1223. <https://doi.org/10.1007/s00436-010-1797-y>

- Kathy H, David M (2023) Insect singers - cicadas of the United States and Canada east of the 100th meridian
- Kattel M, Nepal A, Shah A, Shrestha D (2019) Chroma feature extraction. In: Conference: chroma feature extraction using fourier transform, Number 20
- Kaur N, Kaleka AS (2022) Diversity, importance and decline of pollinating insects in present era. In: El-Shafie HAF (Ed) Global decline of insects, Chapter 5. Rijeka: IntechOpen. <https://doi.org/10.5772/intechopen.100316>
- Kavitha S, Manikandan J (2021) Improved methodology of svm to classify acoustic signal by spectral centroid. *J Trends Comput Sci Smart Technol* 3(4):294–304
- Kawakita S, Ichikawa K (2019) Automated classification of bees and hornet using acoustic analysis of their flight sounds. *Apidologie* 50(1):71–79
- Kevan PG, Clark EA, Thomas VG (1990) Insect pollinators and sustainable agriculture. *Am J Altern Agric* 5(1):13–22. <https://doi.org/10.1017/S0889189300003179>
- Khalifa SAM, Elshafiey EH, Shetaia AA, El-Wahed AAA, Algethami AF, Musharraf SG, AlAjmi MF, Zhao C, Masry SHD, Abdel-Daim MM, Halabi MF, Kai G, Al Naggar Y, Bishr M, Diab MAM, El-Seedi HR (2021) Overview of bee pollination and its economic value for crop production. *Insects* 12(8). <https://doi.org/10.3390/insects12080688>
- Kim J, Oh J, Heo TY (2021) 05. Acoustic scene classification and visualization of beehive sounds using machine learning algorithms and grad-cam. *Math Probl Eng* 2021:1–13
- Kim J, Oh J, Heo TY (2021) Acoustic classification of mosquitoes using convolutional neural networks combined with activity circadian rhythm information. *Int J Interact Multimed Artif Intell*
- Kiskin, I., A.D. Cobb, L. Wang, and S. Roberts 2020. Humbug zooniverse: a crowd-sourced acoustic mosquito dataset. In *ICASSP 2020-2020 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pp. 916–920. IEEE
- Kiskin I, Perez Orozco B, Windebank T, Zilli D, Sinka M, Willis K (2017) Mosquito detection with neural networks: The buzz of deep learning. Cornell University library
- Kora Venu S (2021) An ensemble-based approach by fine-tuning the deep transfer learning models to classify pneumonia from chest x-ray images. In: Proceedings of the 13th international conference on agents and artificial intelligence. SCITEPRESS - Science and Technology Publications
- Kremen C, Chaplin-Kramer R (2007) Insects as providers of ecosystem services: crop pollination and pest control. CABI Books 349–382
- Lange KW, Nakamura Y (2021) Edible insects as future food: chances and challenges. *J Future Foods* 1(1):38–46. <https://doi.org/10.1016/j.jfutfo.2021.10.001>
- Lazzarini V (2001) Sound processing with the sndobj library: an overview. In: Proceedings of the 4th international conference on digital audio effects (DAFX '01)
- Le-Qing Z (2011) Insect sound recognition based on mfcc and pnn. In: 2011 International conference on multimedia and signal processing
- Lin CC, Hedwig B (2021) Wing movements underlying sound production in calling, rivalry, and courtship songs of the cricket *gryllus bimaculatus* (degeer). *J Insect Physiol* 134:104299
- Lunau K, Ren ZX, Fan XQ, Trunschke J, Pyke GH, Wang H (2020) Nectar mimicry: a new phenomenon. *Sci Rep* 10(1):7039. <https://doi.org/10.1038/s41598-020-63997-3>
- Malek A (2023) Spafe: Simplified python audio features extraction. *J Open Sourc Softw* 8(81):4739. <https://doi.org/10.21105/joss.04739>
- Mankin R (2023) United states department of agriculture. Accessed 19-Jan-2023
- Mankin RW (2012) Applications of acoustics in insect pest management. *CABI Rev* 2012:1–7. <https://doi.org/10.1079/PAVS.NR20127001>
- Manosathiyadevan M, Bhuvaneshwari V, Latha R (2017) Impact of insects and pests in loss of crop production: a review. *Sustain Agric Towards Food Secur* 1:57–67
- Marinko J, Ivanovska A, Marzidovšek M, Ramsden M, Debeljak M (2023) Incentives and barriers to adoption of decision support systems in integrated pest management among farmers and farm advisors in europe. *Int J Pest Manag* 1–18. <https://doi.org/10.1080/09670874.2023.2244912>
- McFee B, Raffel C, Liang D, Ellis DP, McVicar M, Battenberg E, Nieto O (2015) Librosa: audio and music signal analysis in python. In: Proceedings of the 14th python in science conference pp 18–25
- Ramsey MT, Bencsik M, Newton MI, Reyes M, Pioz M, Crauser D, Delso NS, Le Conte Y (2020) The prediction of swarming in honeybee colonies using vibrational spectra. *Sci Rep* 10. <https://doi.org/10.1038/s41598-020-66115-5>
- Mitchell T (1997) Machine learning. In: Proceedings of the 21st ACM international conference on multimedia vol 1 no 9
- Miyashita A, Kizaki H, Sekimizu K, Kaito C (2016) No effect of body size on the frequency of calling and courtship song in the two-spotted cricket, *gryllus bimaculatus*. *PloS One* 11:e0146999. <https://doi.org/10.1371/journal.pone.0146999>
- Montalvão J, Araujo MRR (2012) Is masking a relevant aspect lacking in mfcc? a speaker verification perspective. *Pattern Recognit Lett* 33(16):2156–2165
- Muniz-Junior G, de Oliveira Roque F, Pires AP, Guariento RD (2023) Are lower pesticide doses better? an evolutionary perspective on integrated pest management. *Ecol Model* 482:110408. <https://doi.org/10.1016/j.ecolmodel.2023.110408>
- Noda JJ, Travieso CM, Sánchez-Rodríguez D, Dutta MK, Singh A (2016) Using bioacoustic signals and support vector machine for automatic classification of insects. In: 2016 3rd international conference on signal processing and integrated networks (SPIN), pp 656–659. IEEE
- Nolasco I, Terenzi A, Cecchi S, Orcioni S, Bear HL, Benetos E (2019) Audio-based identification of beehive states. In: ICASSP 2019-2019 IEEE International conference on acoustics, speech and signal processing (ICASSP), pp 8256–8260. IEEE
- OSBeehives (2020) Open source beehive
- O'Shea K, Nash R (2015) An introduction to convolutional neural networks. <https://doi.org/10.48550/arXiv.1511.08458arXiv:1511.08458>
- Owens AC, Cochard P, Durrant J, Farnworth B, Perkin EK, Seymoure B (2020) Light pollution is a driver of insect declines. *Biol Conserv* 241:108259. <https://doi.org/10.1016/j.biocon.2019.108259>
- Paim KO, Rohweder R, Recamonde-Mendoza M, Mansilha RB, Cord-eiro W (2023) Acoustic identification of ae. aegypti mosquitoes using smartphone apps and residual convolutional neural networks. <https://doi.org/10.48550/arXiv.2306.10091arXiv:2306.10091>
- Park J, Lee SH, Lee HU, Kim Y (2013) Effect of stress sound on the development of the black soldier fly, *hermetia illucens*. *Korean J Appl Entomol*. <https://doi.org/10.5656/KSAE.2013.06.0.025>
- Parvathy S, Jayan DP, Pathrose N, Rajesh K (2021) Convolutional autoencoder based deep learning model for identification of red palm weevil signals. In: 2021 Asia-pacific signal and information processing association annual summit and conference (APSIPA ASC), pp 1987–1992. IEEE
- Phung QV, Ahmad I, Habibi D, Hinckley S (2017) Automated insect detection using acoustic features based on sound generated from insect activities. *Acoust Aust* 45(2):445–451
- Ponce EJS (2014) Acoustic communication in insects. *Quehacer Científico en Chiapas* 9(2):48–56
- Potamitis I (2023) Mosquito wingbeats recordings. Accessed 10-April-2023
- Potamitis I, Ganchev T, Fakotakis N (2007) Automatic acoustic identification of crickets and cicadas. In: 2007 9th International

- symposium on signal processing and its applications, pp 1–4. IEEE
- Prasad S, Malav LC, Choudhary J, Kannojiya S, Kundu M, Kumar S, Yadav AN (2021) Soil microbiomes for healthy nutrient recycling, pp 1–21. Singapore: Springer Singapore
- Qassim H, Verma A, Feinzimer D (2018) Compressed residual-vgg16 cnn model for big data places image recognition. In: 2018 IEEE 8th Annual computing and communication workshop and conference (CCWC), pp 169–175
- Ribeiro AP, da Silva NFF, Mesquita FN, Araújo PdCS, Rosa TC, Mesquita-Neto JN (2021) Machine learning approach for automatic recognition of tomato-pollinating bees based on their buzzing-sounds. *PLoS Comput Biol* 17(9):e1009426. <https://doi.org/10.1371/journal.pcbi.1009426>
- Riyaz M, Shah RA, Packiam SM (2022) Insect conservation and management: a need of the hour. IntechOpen
- Robles-Guerrero A, Saucedo-Anaya T, Guerrero-Mendez CA, Gómez-Jiménez S, Navarro-Solis DJ (2023) Comparative study of machine learning models for bee colony acoustic pattern classification on low computational resources. *Sensors* 23(1):460
- Sánchez-Bayo F, Wyckhuys KA (2019) Worldwide decline of the entomofauna: A review of its drivers. *Biol Conserv* 232:8–27. <https://doi.org/10.1016/j.biocon.2019.01.020>
- Sharma S, Kumar S (2022) The xception model: A potential feature extractor in breast cancer histology images classification. *ICT Express* 8(1):101–108. <https://doi.org/10.1016/j.ict.2021.11.010>
- Shi L, Li C, Tian L (2019) Music genre classification based on chroma features and deep learning. In: 2019 Tenth international conference on intelligent control and information processing (ICICIP), pp 81–86. IEEE
- Shrestha B, Lee Y (2020) Cellular and molecular mechanisms of deet toxicity and disease-carrying insect vectors: A review. *Genes & genomics* 42:1131–1144. <https://doi.org/10.1007/s13258-020-00991-z>
- Silva DF, De Souza VM, Batista GE, Keogh E, Ellis DP (2013) Applying machine learning and audio analysis techniques to insect recognition in intelligent traps. In: 2013 12th International conference on machine learning and applications, vol 1, pp 99–104. IEEE
- Simonyan K, Zisserman A (2015) Very deep convolutional networks for large-scale image recognition
- Soroye P, Newbold T, Kerr J (2020) Climate change contributes to widespread declines among bumble bees across continents. *Science* 367(6478):685–688. <https://doi.org/10.1126/science.aax8591>
- Srivastava C, Subramanian S (2016) Storage insect pests and their damage symptoms: an overview. *Indian J Entomol* 78(special):53–58. <https://doi.org/10.5958/0974-8172.2016.00025.0>
- Stowell D (2022) Computational bioacoustics with deep learning: a review and roadmap. *PeerJ* 10:e13152
- Terenzi A, Ortolani N, Nolasco I, Benetos E, Cecchi S (2021) Comparison of feature extraction methods for sound-based classification of honey bee activity. *IEEE/ACM Trans Audio Speech Lang* 30:112–122
- Terenzi A, Cecchi S, Spinsante S (2020) On the importance of the sound emitted by honey bee hives. *Vet Sci* 7(4). <https://doi.org/10.3390/vetsci7040168>
- Thomas W (2022) Singing insects of north america (sina). <https://orths.oc.org/sina/cricklist.htm>. Accessed 19-April-2023
- Torky M, Dahy G, Hassanien AE (2023) Recognizing sounds of red palm weevils (rpw) based on the vggish model: Transfer learning methodology. *Comput Electron Agri* 212:108079. <https://doi.org/10.1016/j.compag.2023.108079>
- Towsey MW (2017) In: The calculation of acoustic indices derived from long-duration recordings of the natural environment
- Trivedi K, Shroff H (2021) Identification of deadliest mosquitoes using wing beats sound classification on tiny embedded system using machine learning and edge impulse platform. In: 2021 ITU kaleidoscope: connecting physical and virtual worlds (ITU K), pp 1–6. IEEE
- Uthoff C, Homs MN, von Bergen M (2023) Acoustic and vibration monitoring of honeybee colonies for beekeeping-relevant aspects of presence of queen bee and swarming. *Comput Electron Agri* 205:107589
- Valletta JJ, Torney C, Kings M, Thornton A, Madden J (2017) Applications of machine learning in animal behaviour studies. *Anim Behav* 124:203–220. <https://doi.org/10.1016/j.anbehav.2016.12.005>
- Várkonyi DT, Junior JLS, Horváth T (2023) Dynamic noise filtering for multi-class classification of beehive audio data. *Expert Syst Appl* 213:118850
- Varma ALS, Bateshwar V, Rathi A, Singh A (2021) Acoustic classification of insects using signal processing and deep learning approaches. In: 2021 8th International conference on signal processing and integrated networks (SPIN), pp 1048–1052. IEEE
- Wang Y, Zhao H, Zhang X, Ren B (2016) The complete mitochondrial genome and song evolution of the monotypic genus *u. tarbinsky*, 1932 (orthoptera: Tettigoniidae). *Environ Entomol* 45(3):737–746
- Wei X, Hossain MZ, Ahmed KA (2022) A resnet attention model for classifying mosquitoes from wing-beating sounds. *Sci Rep* 12(1):10334. <https://doi.org/10.1038/s41598-022-14372-x>
- Weiping Z, Jiantao Y, Xiaotao X, Xiangtao L, Shaohu P (2017) Acoustic scene classification using deep convolutional neural network and multiple spectrograms fusion. In: Detection and classification of acoustic scenes and events (DCASE)
- Xie J, Bertram SM (2019) Using machine learning techniques to classify cricket sound. In: Eleventh international conference on signal processing systems, vol 11384, pp 141–148. SPIE
- Xu Y, Zhou Y, Sekula P, Ding L (2021) Machine learning in construction: From shallow to deep learning. *Dev Built Environ* 6:100045. <https://doi.org/10.1016/j.dibe.2021.100045>
- Yayin X, Ying Z, Przemyslaw S, Lieyun D (2021) Machine learning in construction: From shallow to deep learning. *Dev Built Environ* 6
- Yazgac BG, Kirci M, Kivan M (2016) Detection of sunn pests using sound signal processing methods. In: 2016 Fifth international conference on agro-geoinformatics (Agro-Geoinformatics), pp 1–6. IEEE
- Zamanian H, Pourghassem H (2017) Insect identification based on bioacoustic signal using spectral and temporal features. In: 2017 Iranian conference on electrical engineering (ICEE), pp 1785–1790. IEEE
- Zgank A (2022) Reduced mfcc feature extraction dimension for acoustic classification of bee swarm activity. In: 2022 ELEKTRO (ELEKTRO), pp 1–4. IEEE
- Zgank A (2020) Bee swarm activity acoustic classification for an iot-based farm service. *Sensors* 20(1):1. <https://doi.org/10.3390/s20010021>
- Zhang, N. and M. Guo 2010. Recognition of fruit fly wings vibration sound based on hmm. In *2010 2nd International Conference on Information Engineering and Computer Science*, pp. 1–4. IEEE
- Zhang C, Wang P, Guo H, Fan G, Chen K, Kamarainen J (2017) Turning wingbeat sounds into spectrum images for acoustic insect classification. *Electron Lett* 53
- Zhang M, Yan L, Luo G, Li G, Liu W, Zhang L (2021) A novel insect sound recognition algorithm based on mfcc and cnn. In: 2021 6th International conference on communication, image and signal processing (CCISP), pp 289–294. IEEE
- Zhao Y, Deng G, Zhang L, Di N, Jiang X, Li Z (2021) Based investigate of beehive sound to detect air pollutants by machine learning. *Ecol Inf* 61:101246. <https://doi.org/10.1016/j.ecoinf.2021.101246>
- Zhuang F, Qi Z, Duan K, Xi D, Zhu Y, Zhu H, Xiong H, He Q (2021) A comprehensive survey on transfer learning. *Proc IEEE* 109(1):43–76. <https://doi.org/10.1109/JPROC.2020.3004555>