Multimedia Tools and Applications (2024) 83:23269–23296
https://doi.org/10.1007/s11042-023-15192-5

Bee detection in bee hives using selective features from acoustic data

Furqan Rustam[1] · Muhammad Zahid Sharif[2,3] · Wajdi Aljedaani[4] · Ernesto Lee[5] · Imran Ashraf[6]

Received: 19 May 2022 / Revised: 11 October 2022 / Accepted: 30 March 2023 /
Published online: 14 August 2023

Abstract
Honeybees, a key pollinator of the world's most cultivated crops, are experiencing colony collapses due to a variety of factors. The existence of honeybees and queens is critical to the sociality of a colony, and the presence of bees in agricultural settings is vital to the ecological balance. Moreover, beehives without a queen may lead to the decline of an entire colony therefore finding them through effective and an accurate approach is a critical task. In this scenario, we analyzed acoustic/sound data of various classes (i.e. Bee, NoBee, and NoQueen) from beehive colonies. This study examines five distinct features including spectral centroid, zero-crossing rate, Mel-frequency cepstral coefficients (MFCC), chromagram, and constant Q-transform characteristics for their suitability in detecting bees using the acoustic data. In addition, selective features using principal component analysis (PCA), Chi-square analysis (Chi2), and singular value decomposition (SVD) are used. Moreover,

Furqan Rustam and Muhammad Zahid Sharif contributed equally in this research.

✉ Imran Ashraf
imranashraf@ynu.ac.kr

Furqan Rustam
furqan.rustam1@gmail.com

Muhammad Zahid Sharif
2011ag4010@uaf.edu.pk

Wajdi Aljedaani
wajdi.j1@gmail.com

Ernesto Lee
elee@mdc.edu

1 School of Computer Science, University College Dublin, D04 V1W8 Dublin, Ireland

2 Hefei Institutes of Physical Science, Chinese Academy of Sciences, Hefei 230031, People's Republic of China

3 University of Science and Technology of China, Hefei 230026, People's Republic of China

4 University of North Texas, 155 Union Cir, Denton, TX 76203, USA

5 College of Engineering and Technology, Miami Dade College, Miami, FL, USA

6 Information and Communication Engineering, Yeungnam University, Gyeongsan 38541, Korea

the study proposes hybrid features where selective PCA, Chi2, and SVD characteristics are integrated to create a suitable feature set. Experimental results exhibit the suitability of the hybrid feature set which outperformed individual features for “Bee”, “NoBee” and “NoQueen” classes prediction. Cross-validation and T-test results also confirm the superior performance of hybrid MFCC features. The results indicate that RF and KNN show better performance than other machine learning models with maximum accuracy scores of 0.82 and 0.83, respectively.

Keywords Bee detection · Acoustic analysis · Feature engineering · Mel-frequency cepstral features

1 Introduction

Bees are very significant pollinators because they are reported to pollinate about 73% of the world’s cultivated crops [1]. Being specific to honeybees, they deliver very effective pollination services for an extensive variety of cultivated crops [9]. Apis mellifera out of other honeybee species is reported to be the most frequent foragers of flowers in natural settings on a global scale [28]. Monitoring bees’ presence in the field and within the hives has become a very important task because they are facing a decline due to various factors including pest infection, environmental pollution, pesticide applications, or possibly the phenomenon designated as colony collapse disorder (CCD) [16, 20, 25].

Apart from the above-mentioned factors, the role of the queen in a honeybee colony is very significant because she controls worker bees by releasing pheromones and producing eggs. Literature testifies that the queen-less state may result dying of an entire colony within merely a few months until an alternative or new queen is introduced [40]. That is why a colony holding a queen bee is considered a healthy colony while a queen-less colony is unhealthy. Because unhealthy queen and queen-less states both are extremely unfavorable therefore should be recognized as soon as possible to adopt proper management strategies [11].

To monitor honeybee colonies, various monitoring systems have been employed: most of those systems relied on the measurement of multiple hive parameters including humidity, carbon dioxide, as well as weight [2, 8]. But in recent years, previous systems are updated and shifted to web-based monitoring systems assembled upon sensors as well as cloud architecture, to keep an eye on bees’ behavior [14]. But a non-invasive approach to observing bees’ status mainly relies on sound-based investigations because bees use vibroacoustic signals to communicate within the colony [21, 33, 36]. Sound patterns of honeybee colonies can be used as an indicator of the health of the bee community. They are considered efficient markers to estimate abnormal changes occurring in the beehive because their measure can be simple and low-cost, moreover can be used in multiple hives for their continuous monitoring [17, 19, 32]. In climate-smart agriculture, a myriad of data on hives’ acoustics can be used to extract novel information [47] regarding colony health, colony strength, swarming, weather condition, pests and disease infection, and pesticide impact [10, 11, 38, 40, 46, 57].

Cejrowski, et al. [11] used the Support vector machine (SVM) approach on two data sources, the first is input data in the form of an n-dimensional vector, and the second is the output of t-SNE. The t-SNE (t-distributed stochastic neighbor embedding) is a technique to work on high dimensional data. Based on the output of the t-SNE algorithm they were certain that honeybee coonies with a queen as well as without a queen act differently. Laurens van der Maaten created the method in 2008 as a variant of the previously known SNE algorithm established by Geoffrey Hinton and Sam Roweis in 2002. T-SNE transforms a multidimensional collection of data $x = x1, x2, ..., xn$ into 2D or 3D vectors $Y = y1, y2, ..., yn$. This approach works by comparing the density distribution of multivariate variables to the distribution of their projection on a two or three-dimensional plane [51]. Moreover, in another study researchers investigated to know whether healthy (with a queen) and unhealthy (queenless or with a missing queen) colonies can be recognized through frequency analysis of the honeybee sound. They characterized sound samples of both colonies via Mel Frequency Cepstral Coefficient (MFCC) and a statistical descriptor was acquired instead of each Mel coefficient. After exploratory analysis, they revealed that both hives show different characteristics [40].

In another study, Ruvinga, et al. [43] applied various models to characterize sound samples (collected from four separate honeybee hives of Apis mellifera colonies) to discriminate between colonies which are having a queen and those with the absence of a queen bee. For this purpose, they employed the long short-term memory (LSTM) model, multi-layer perception (MLP) neural networks model, and logistic regression model. After analyzing data, results exhibited that LSTM is the higher performance model than the other models in terms of attaining the best accuracy of 0.92 while MLP, as well as logistic regression, also presented the best accuracy of 0.90 and 0.87 respectively. They also expressed their plan about testing above mentioned models on data accumulated from multiple hives and encompassing this task by employing Mel spectrograms as features on a CNN (Convolutional Neural Network) in order to further generate perfection or genuineness in the results.

Sound-based methodologies including spectrographic analysis as well as some other approaches in human speech analysis are being employed to distinguish the "queenright" and "queenless" states of different colonies 24. In this scenario, Howard, et al. [27] compared the spectrograms, FFT (Fast Fourier transform), and S-transform of the honeybee colony audio recordings, and to evaluate various approaches, they classified outcomes of the frequency examination through a Kohonen Self-Organising map (SOM) artificial neural network. An FFT is an algorithm that computes the discrete Fourier transform of a sequence, or its inverse. Fourier analysis converts a signal from its original domain to a representation in the frequency domain and vice versa. They found that the results obtained by using a SOM, to sort or categorize the data from variant hives, are less successful while histogram analysis yielded good potential. Research shows that whenever a queen bee is present or absent, in both situations FFT of the sound signal varies (400 Hz with the queen while different for the queenless colony) [40]. Another work shows that colonies with a queen presence produce a frequency range of 300-350 Hz [38] but we couldn't see any clear information about the frequency range in a queenless colony.

In order to recognize beehive sound, Nolasco and Benetos [35] explored the perspective of the machine learning approach. In this approach, they used SVM as well as a CNN classifier, and the results of each experiment were examined employing the area under the curve (AUC). The scores accomplished via the CNN implementation fail to achieve the level of the SVM approach (which performed better in comparison). They consider this work as a first step to further updating the honeybee monitoring system to detect hive bees through their buzzing sound. All the relevant work done to recognize bees and queen status is summarized in Table 1.

Self-Learning network-based segmentation has been published in medical diagnosis for assessing disease or damaged regions of the brain using a heuristic approach for real-time

Table 1 Summary of sound-based analysis of beehive in related work

Study	Models	Dataset	No. of classes (name of classes)	Evaluation metrics	Results
[40]	Mel Frequency Cepstral Coefficient (MFCC), Logistic Regression Model	Sound samples were gathered from a couple of colonies of Apis mellifera carnica. One colony was with a queen and the other was with a missing queen. Duration for honeybee colony recording was about 216 hours.	2 (with a queen, and missing queen)	Machine learning algorithms	Mean decrease accuracy of both Mel coefficient 4 and 10 is more than 15 Receiver operative characteristic (ROC) curve showing the maximum number of correctly classified samples
[11]	SVM Extracted from two honeybee colonies, one data to classify colonies with a queen and without a queen while other data were used to characterize colonies with the old queen and new queen.	4 (colony with a queen, missing queen, old queen, and new queen). Data were divided into two sets, one obtained from normal bees' work and the other from abnormal. Data were acquired after every 15 minutes while the colony was monitored between February 2017–August 2017.	Linear predictive coding (LPC), Classification potential – t-SNE, Learning - SVM	SVM algorithms with C-classification result in correct characterization and the model is very accurate.	Error for test data of old queen = 5.55% Error for test data of new queen = 9.28%
[43]	LSTM model, Multi-MLP Neural Networks model, and Logistic regression model	Waveform audio files (each 1 min duration) collected on an hourly basis for a whole week. Data were collected from four different honeybee colonies. One-minute Waveform audio files were sampled on an hourly basis from the 3rd to the 9th of August 2012 which makes a total duration of about 168 minutes.	2 (with a queen, and missing queen)	MFCCs as input feature, LSTM, energy computed in MATLAB, ANOVA test for MFCC plus log energy averages	With queen and queen-less colonies are significantly different at P = 0.001, Confusion matrix shows high accuracy for LSTM as 0.92 = 92%, while 0.90, and 0.87 for MLP and Logistic regression, correspondingly.

Table 1 (continued)

Study	Models	Dataset	No. of classes (name of classes)	Evaluation metrics	Results
[27]	ANN model	Data was collected from two subspecies of the honeybee, Apis mellifera ligustica and Apis mellifera carnica. 9 days of recording from four separate hives of equal size. Data was collected from two subspecies of the honeybee, Apis mellifera ligustica, and Apis mellifera carnica. Seven days of recording from four separate hives of equal size. Total recording duration is about 168 minutes.	2 (Queen right and Queenless)	Spectrogram, FFT, S-transform, Kohonen Self-Organising Map (SOM) artificial neural network	Classification by SOM is less successful, the histogram shows good potential
[35]	Machine learning	Data collection was made via Open source beehive (OSBeehive) project, AND NU-Hive project/ Total duration of 78 recordings of several lengths were as 12 hours. 25% of this is annotated as noBee event. 60% recordings were from NU-Hive dataset representing two hives while other recordings were produced from OSBH dataset representing 6 different hives.	2 (Recognizing hive bees [i.e., bee, no bee])	SVM classifier CNN classifier The area under the curve score (AUC)	SVM is a better classifier

image segmentation (HARIS). The same can be applied in several other fields. The benefit of employing this strategy is that it is productive with little training and has a 77% accuracy rate [49]. A recent study on honeybee colony monitoring outlines how an integrated camera module powered by a deep learning algorithm can be used to early detect Varroa infestations (varroa is one of the major pests and a primary cause of colony collapse in honeybees) [52].

Various novel technological tools naming as BuzzBox mini, Arnia, and ApisProtect monitor multiple situations occurring in the hive, such as temperature and humidity variation, brood state, queen status, foraging events, nectar flow, swarming, as well as notify some suggestions to keep colonies vigorous and prevent them from decline. In a recent scenario, we feel that there is a dire need to develop a more updated tool with efficient application to more precisely detect different colony classes like the presence of bees, no bees, queen-right, queen missing, etc. Considering the significance of these situations we are addressing this issue through the introduction of a novel updated application (which works based on a machine learning approach) in the present paper.

1.1 Contributions and significance

This study proposed an approach for beehive acoustic analysis using a machine learning approach. We deployed our proposed approach for bee detection and perform experiments on a publicly available dataset. These are key contributions of the study

- A comprehensive analysis of different features is performed to investigate the suitability of features for distinguishing three classes of bees: 'Bee', 'NoBee', and 'NoQueen'. In this perspective, spectral centroid (SC), zero-crossing-rate (ZCR), Mel-frequency cepstral coefficients (MFCC), chromagram (CG), and constant Q-transform features are studied. The feasibility of these features is analyzed for bee detection from the acoustic data.
- Performance of three feature selection approaches has been investigated including Chi-square (Chi2), principal component analysis (PCA), and singular value decomposition (SVD) for each of the five features.
- A hybrid feature engineering approach is proposed to boost the performance of learning models by combining MFCC, PCA, SVD, and Chi2. Results are validated by cross-validation and a statistical t-test.

The rest of the paper is structured into four sections. A discussion of the proposed methodology is presented in Section 2 followed by a detailed evaluation of various feature extraction and feature selection algorithms regarding the acoustic data in Section 3. Experimental results are analyzed in Section 4 while Section 5 concludes this study.

2 Materials and methods

We performed experiments for bee detection using a sound dataset and for that we used several machine-learning techniques and methods. For this, we have used a publicly available dataset [5] containing three target classes Bee, NoBee, and NoQueen. We perform feature engineering on the used dataset and for that, we deploy feature extraction and feature selection techniques. First, we perform feature extraction with the best technique for sound datasets such as MFCC, SC, CG, CQT, and ZCR. These techniques extract features from sound segments and then we pass these features to the feature selection technique to select the best features. For that, we used Chi2, PCA, SVD, and a hybrid approach which is a combination of these three. The selected features are then fed into the machine learning models. We split the feature dataset into training and testing sets with an 80:20 ratio where 80% of the data we used for the training of models and 20% for the testing of models. The hybrid approach is used to improve the performance of learning models and we evaluate the performance of all models in terms of accuracy, precision, recall, and F1 score. The flow diagram of the proposed approach is shown in Fig. 1.

2.1 Dataset description

We used a sound segment dataset consisting of three Beehive buzz anomalies Bee, NoBee, and NoQueen. We can say that the "No-Bee" is an ambient sound. The no-Bee intervals identify moments when an outside sound can be heard [29]. It is identified that if the queen is old or diseased (with low pheromonal signal) or it dies (producing no pheromonal signal), workers start rearing new queens from young brood within a period of 12-24 hours. The removal of the queen (when there is no young brood) quickly causes the colony to depopulate, and decline the colony since the workers cease their activities [7]. In this situation

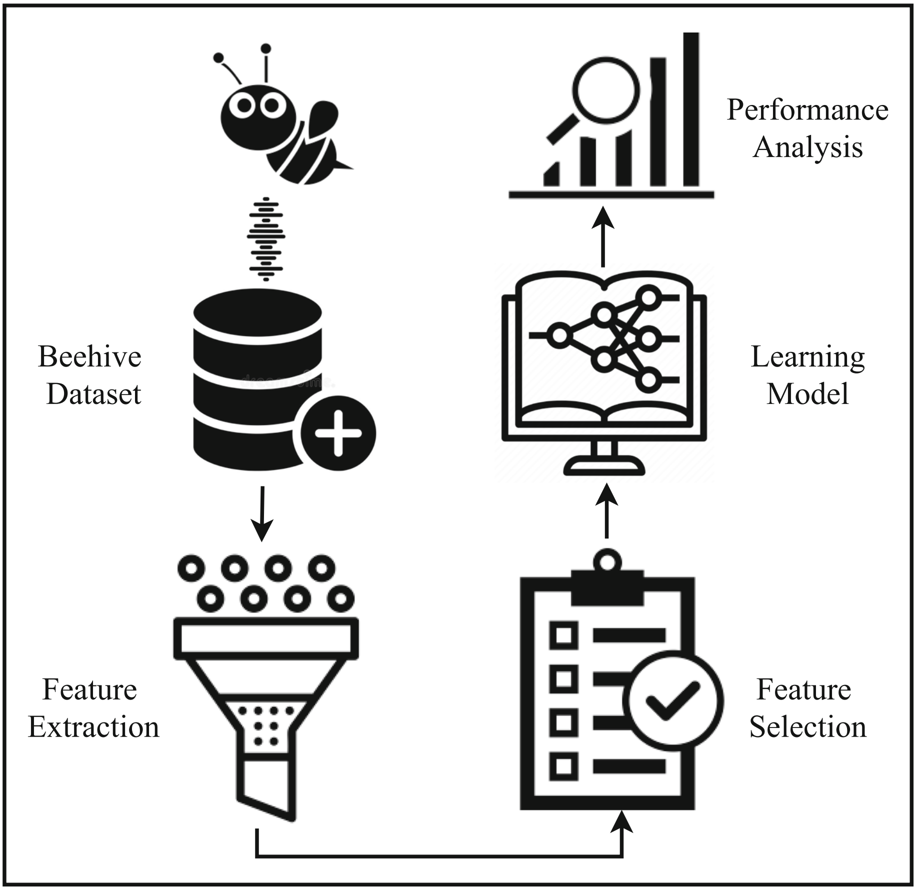


Fig. 1 Proposed approach flow diagram

Table 2 Training and testing count

Target	Training set	Testing set	Total
Bee	4382	1091	5473
NoBee	2752	706	3458
NoQueen	3899	962	4861
Total	11033	2759	13792

when workers cease their activities we listen to no sound inside the hive nonetheless the only background/ambient noise or outside sound. So, detecting this kind of sound means that the colony is collapsed, die, or is in the process of rearing a new queen. The dataset consists of sound records collected through the citizen science initiative and contains a total of 13792 sound samples [5]. This dataset contains 5473 Bee samples, 3458 NoBee samples, and 4861 NoQueen samples. We split this dataset into training and testing sets with an 80:20 ratio. The count for training and testing sets is shown in Table 2.

3 Feature engineering

This study uses feature extraction and feature selection techniques to help learning models improve bee detection accuracy.

3.1 Feature extraction

Several feature extraction techniques have been investigated for their suitability and efficacy for bee detection from acoustic data. These feature extraction techniques have been selected based on the findings from the literature and their reported results [15, 23, 34, 44, 54].

3.1.1 Spectral centroid features

SC is used to characterize a spectrum in a digital signal as it indicates where the spectrum center of mass is located. We used the 'librosa' library to extract the SC features. The high value of SC indicates the high energy of the signal with higher frequency [22]. The Center of gravity of the spectrum is SC and the value of SC of i^{th} frame in audio can be defined as in (1) [22].

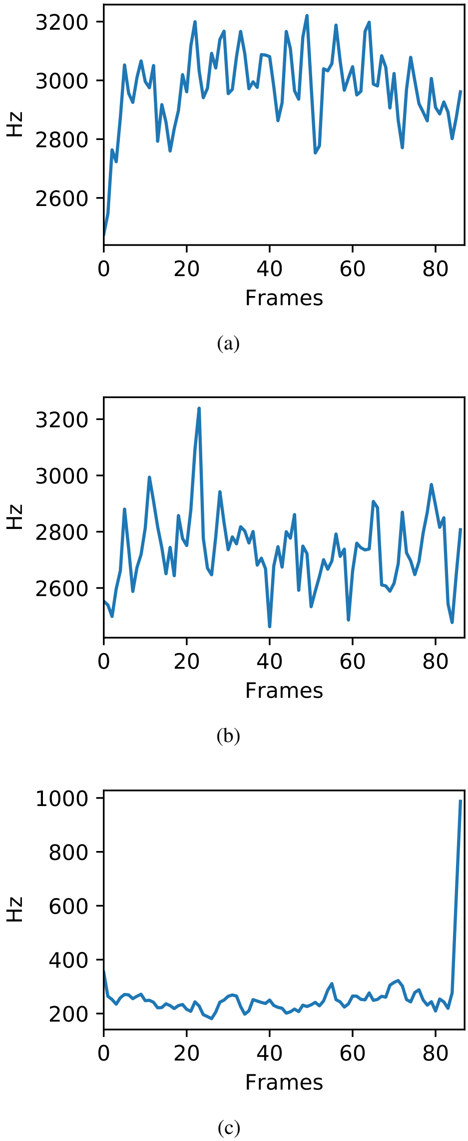
$$C_i = \frac{\sum k = 1^{Wf_L} k X_i(k)}{\sum k = 1^{Wf_L} X_i(k)} \tag{1}$$

Figure 2 shows the SC for the sample sound of Bee, NoBee, and NoQueen categories. It can be observed that the frequency varies for each target class. For a better understanding, Fig. 2 is placed where the difference in frequency of 'Bee', 'NoBee', and 'NoQueen' can be seen. The visual representation indicates that the SC feature can be used to distinguish these classes.

3.1.2 Zero-crossing-rate features

ZCR is the number of time a signal goes up and down from zero point or a signal change its sign from positive to zero and zero to negative. Simply put, we can define ZCR as the change in a signal in a frame [4]. The smoothness of the signal can be calculated using the

Fig. 2 SC of all three categorize of sound, (a) Bee (b) NoBee and (c) NoQueen



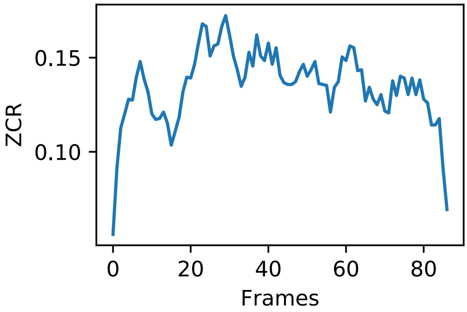
ZCR [22]. The measure of nosiness of a signal can be interpreted as ZCR and a high value indicates the noise in the signal. ZCR can be calculated using the (13) [22]

$$Z(i) = \frac{1}{2W_L}\sum_{n=1}^{W_L} |sgn[x_i(n)] - sgn[x_i(n-1)]| \tag{2}$$

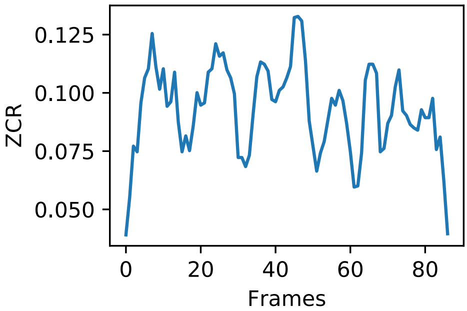
where $sgn(.)$ is a sign function and we can define this sign function as [22]

$$sgn[x_i(n)] = \begin{cases} 1 & x_i(n) \geq 0 \\ -1 & x_i(n) < 0 \end{cases} \tag{3}$$

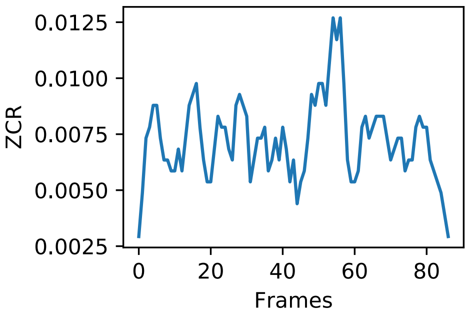
Fig. 3 ZCR of all three categorize of sound, (a) Bee (b) NoBee and (c) NoQueen



(a)



(b)



(c)

Figure 3 shows the ZCR for the sample sound of the Bee, NoBee, and NoQueen categories. It is observed that the ZCR for the three target classes is different whereas the ZCR for ’NoQueen’ class is substantially different from those for ’Bee’ and ’NoBee’. Similarly, the ZCR for Bee and NoBee classes are different and can be used to discriminate the sound signals for these classes correspondingly.

3.1.3 Mel-frequency cepstral coefficients features

MFCC is one of the most used techniques for feature extraction from audio data. MFCC represents the sound power spectrum based on linear cosine transform. Learning models can perform significantly with MFCC features as compared to the original signal as the input. MFCC feature extraction is different from other cepstral features which are on the Mel scale. MFCC technique normalizes the values between 0 and 1 in the presence of adaptive noise [26]. Mels for any frequency can be calculated using the formula given in (4) [12, 24] where $mel(f)$ is the frequency (mels) and f is the frequency (Hz).

$$mel(f) = 2595\ x\ log_{10}(1 + \frac{f}{700}) \tag{4}$$

While we can calculate MFCC feature using (5) according to study [12].

$$\hat{C}_n = \sum_{n=1}^{k} \left(log\hat{S}_k\right) cos\left[n(k - \frac{1}{2})\frac{\pi}{k}\right] \tag{5}$$

Figure 4 shows the MFCC interpretation for the sample sound of Bee, NoBee, and NoQueen categories. Although apparently looking similar, MFCC patterns are different for each of the three target classes. Especially, the MFCC for NoBee class looks significantly different from the other two classes.

3.1.4 Chromagram features

CG features are a very powerful representation of audio data and represent the entire spectrum in 12 bins. Its also known as pitch class profile which is a time-frequency distribution variation. CG is a descriptor that represents the audio signal in a condensed form [6]. We used 'librosa' library to calculate the CG and set the hop length equal to 512.

Figure 5 shows the CG interpretation for the sample sounds of the Bee, NoBee, and NoQueen categories. It can be observed that the patterns across pitch class and time are very different for each target class indicating the suitability of the CG features for bee detection.

3.1.5 Constant Q-transform features

CQT is related to the Fourier transform which transforms the data series into the frequency domain. It is more useful when frequencies span several octaves. Its frequency resolution depends on the frequencies from the center of the window [55]. We used 'librosa' library to calculate the CQT and set hop length equal to 512 and frequency minimum equal to 36 Hz. Figure 6 shows the CQT interpretation for the sample sound of Bee, NoBee, and NoQueen categories. Figure 6 reveals that the CQT features are different for each of the target classes, especially, CQT features for NoQueen class are significantly different from other classes indicating the potential of CQT features for distinguishing between the target classes.

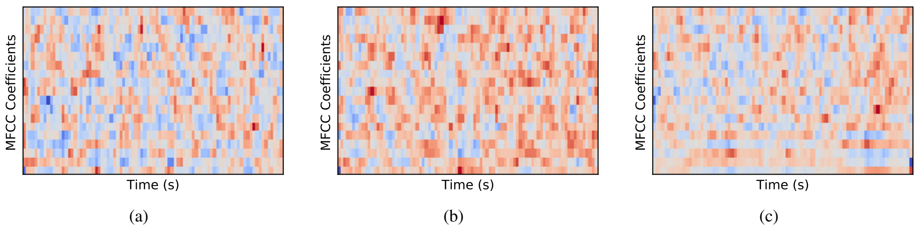
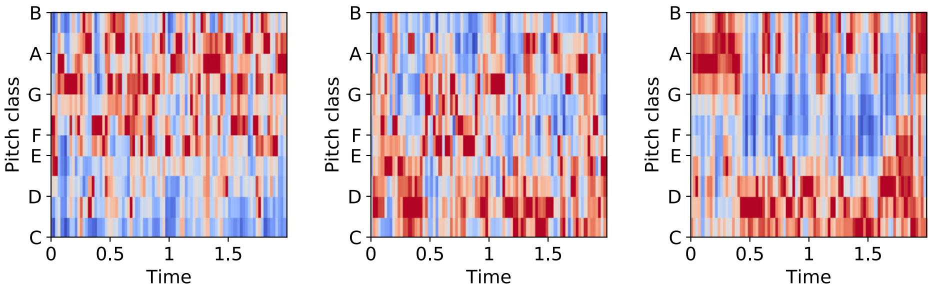


Fig. 4 MFCC of all three categorize of sound, (a) Bee (b) NoBee and (c) NoQueen



(a) (b) (c)

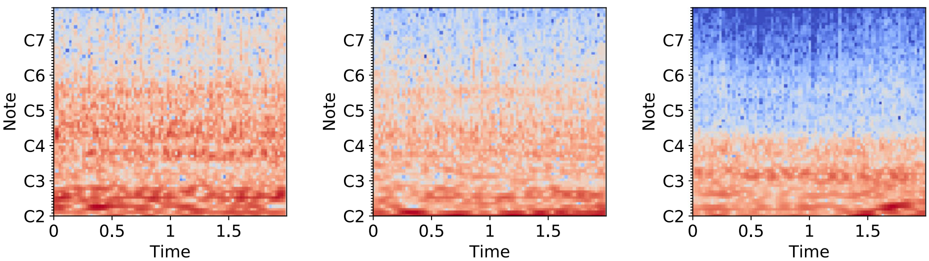
Fig. 5 CG of all three categorize of sound, (a) Bee (b) NoBee and (c) NoQueen

3.2 Feature selection

This study performs feature selection to obtain better performance for beehive acoustic analysis. We used PCA, SVD, and Chi2 techniques for this purpose. In addition, the features from these techniques are combined for experiments.

3.2.1 Principal component analysis

PCA is a linear unsupervised feature selection technique based on eigenvectors analysis [31]. The eigenvector helps to identify the critical features and generates a new feature set named principal components. PCA selects the features according to their magnitude and reduces the feature set containing the best features for the training of learning models. An investigation of PCA-important features is performed using SC, ZCR, MFCC, CG, and CQR. Figure 7 shows the feature space for all three categories using PCA features for SC, ZCR, MFCC, CG, and ZCR. Visualization of PCA-based feature space reveals that the selected features have different patterns for feature space for each of the selected feature types and can distinguish between Bee, NoBee, and NoQueen classes. For example, threshold values across x and y axes, as well as the distribution of feature space are different in terms of the number of samples and sparsity.



(a) (b) (c)

Fig. 6 CQT of all three categorize of sound, (a) Bee (b) NoBee and (c) NoQueen

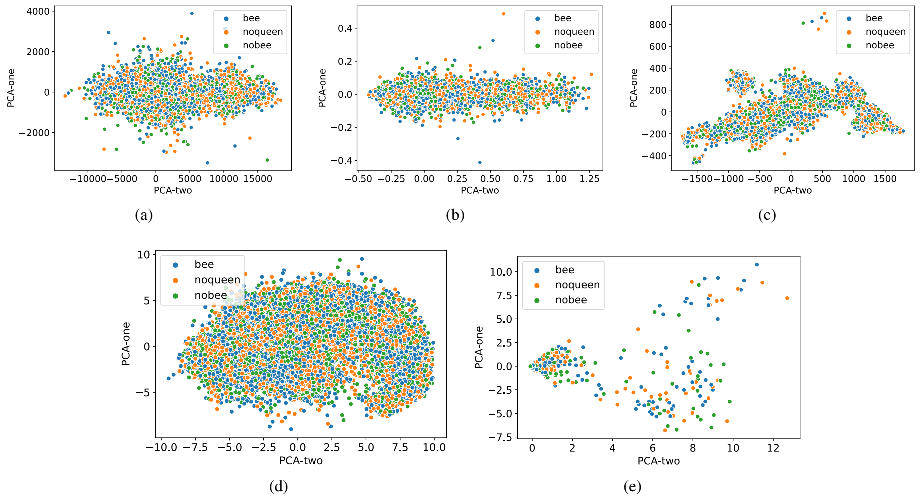


Fig. 7 Feature space representation using PCA (a) Spectral Centroid (b) Zero-crossing-rate, (c) MFCC, (d) Chromagram, and (e) Constant Q-transform

3.2.2 Chi square

Chi2 is a feature selection technique that finds the relationship between two independent variables [30]. It uses the concepts of supervised learning and statistical approach and generates null hypothesis (H_0). If the null hypothesis is accepted it indicates that the null hypothesis is independent else dependent. The Chi2 can be estimated using (6) [13]

$$x_c^2 = \sum \frac{(O_i - E_i)^2}{E_i} \tag{6}$$

where c is the degree of freedom, O is the observed count, and E is the expected count.

Figure 8 shows the Chi2-based feature space representation for five feature extraction techniques, previously described. Chi2-based feature space is different for each of the selected features in this study indicating the suitability of these features for the task at hand. Moreover, it also shows that Chi2 can be used for better training of the models as the Chi2-selected features vary significantly which can help to obtain better performance.

3.2.3 Singular value decomposition

Similar to PCA and Chi2, this study also used the SVD feature selection technique which is very similar to PCA. SVD technique has been used to factorize a matrix that is more correlated to the target class as compared to the original feature set [18]. SVD is a classic linear algebra approach to reduce dimensionality and feature ranking. Figure 9 shows the visual representation of SVD-based feature space for the Bee, NoBee, and NoQueen classes. It reveals that the thresholds o both planes and distributions are different for different classes indicating the feasibility of using SVD-based selected features for bee detection.

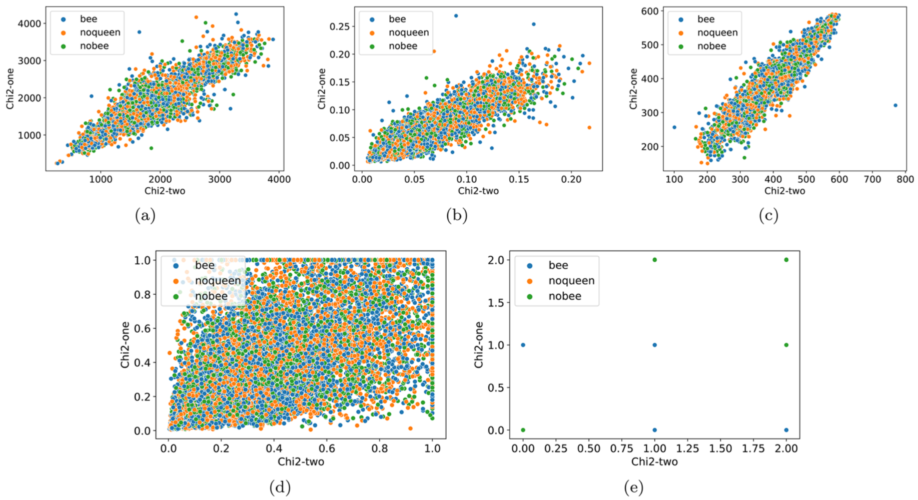


Fig. 8 Feature space representation using Chi2 (a) Spectral Centroid (b) Zero-crossing-rate, (c) MFCC, (d) Chromagram, and (e) Constant Q-transform

3.2.4 Hybrid features

This study also utilizes a hybrid feature set which is generated by combining the selected features from the PCA, SVD, and Chi2 for MFCC, CQT, CG, SC, and ZCR. We deploy the feature selection after features are extracted from the acoustic data, as shown in Fig. 10.

For making a hybrid feature set, the selected features from PCA, Chi2, and SVD are combined as follows

$$HF_{MFCC} = PCA(MFCC) + SVD(MFCC) + Chi2(MFCC) \tag{7}$$

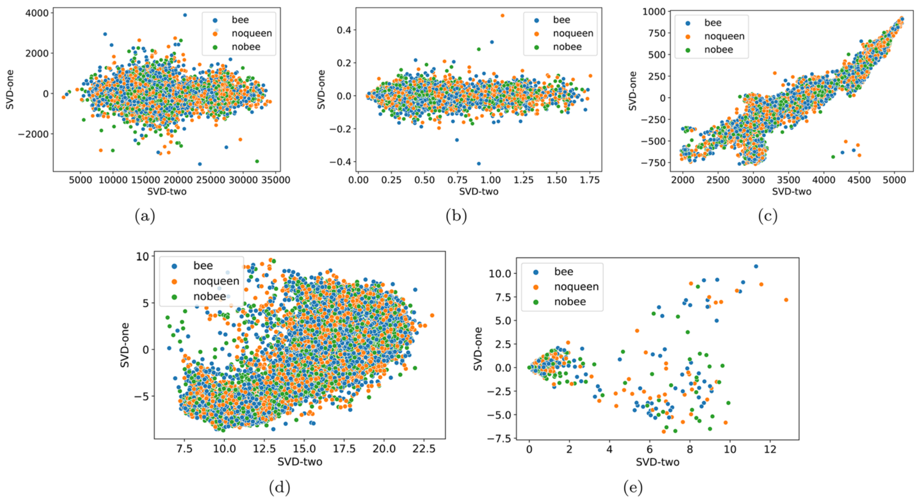


Fig. 9 Feature space representation using SVD (a) Spectral Centroid (b) Zero-crossing-rate, (c) MFCC, (d) Chromagram, and (e) Constant Q-transform

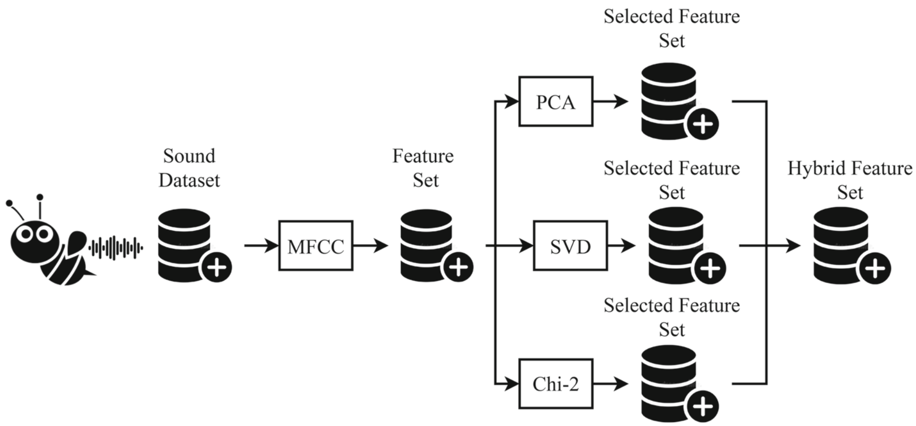


Fig. 10 Hybrid feature approach. The same procedure is used for every feature like CG, ZCR, etc

$$HF_{CQT} = PCA(CQT) + SVD(CQT) + Chi2(CQT) \tag{8}$$

$$HF_{CG} = PCA(CG) + SVD(CG) + Chi2(CG) \tag{9}$$

$$HF_{CS} = PCA(CS) + SVD(CS) + Chi2(CS) \tag{10}$$

$$HF_{ZCRS} = PCA(ZCRS) + SVD(ZCRS) + Chi2(ZCRS) \tag{11}$$

Table 3 shows the number of features for each MFCC, CQT, etc. for the hybrid feature set. The hybrid feature (HF) set is the best for all types of feature extraction techniques as it reduces the size of the feature set and combines only the important features highly correlated with the target class. Also, SC and ZCR generate a small feature set, and combining them increases the size of features which helps to obtain better results.

3.3 Machine learning models

We deploy four state-of-the-art machine learning models with all used feature extraction techniques. We used these models with their best hyperparameter settings selected by analyzing the performance of each model for parameter range. We used KNN, LR, SVC, and RF, and their hyperparameters and values are shown in Table 4.

KNN is a simple machine learning model which predicts by matching the new data with the existing data and considers neighbors for matching the similarity. The class with the

Table 3 Number of samples for hybrid feature set

Feature	Total features	PCA	SVD	Chi2	HF
MFCC	1740	50	50	50	150
CQT	6264	50	50	50	150
CG	1044	50	50	50	150
SC	87	50	50	50	150
ZCR	87	50	50	50	150

Table 4 Hyperparameters used for optimizing the performance of models

Model	Hyperparameters	Tuning range
KNN	n_neighbors= 100	n_neighbors={1 to 200}
LR	solver = saga, C = 5.0	solver = {sag, saga} C = {1.0 to 10.0}
SVC	kernel = 'poly', C = 5.0	kernel= {'linear', 'poly'} , C = {1.0 to 10.0}
RF	n_estimators = 300, max_depth = 200	n_estimators = {50 to 500}, max_depth = {50 to 500}

closest samples to the new data will be the predicted class for new data [56]. The similarity is measured using specific distance metrics like the Euclidean, and Manhattan distances, etc. The Euclidean distance is the most widely used distance estimation metric for KNN and can be calculated as

$$Euclidean\ distance = \sqrt{\sum_{i=1}^{k}(x_i - y_i)^2} \tag{12}$$

We used the KNN with a single parameter n_neighbours with value 100 which considers 100 neighbors for classifying the new data sample.

LR is a statistical model used for classification problems. LR finds the relationship between dependent and independent variables [48]. LR uses the Sigmoid function to calculate the relationship which can be defined as

$$sigmoid\ function = \frac{1}{1 + e^{-(\beta_o + \beta_i)}} \tag{13}$$

We used LR with three hyperparameters such as solver with value 'saga' and multi_class with value multinomial because of multi-class data and C with value 5.0.

RF is an ensemble model that uses several decision trees for training and combines their output under majority voting criteria [39]. RF trains multiple weak learners (decision trees) and uses weak learners' predictions to make the final prediction which can be defined as

$$rf = mode\{\sum_{i=1}^{N} tree_n\} \tag{14}$$

We used RF with two hyperparameters such as

n_estimators with value 300 which defines the number of decision trees to be 300 and used max_depth with value 200 which restricts each decision tree to a maximum 200-level depth.

SVC is a linear model used for the classification of data. It uses hyperplanes to classify the data [50]. Basically, SVC puts multiple hyperplanes in the feature space, and the task is to find the hyperplanes with the best margin [45]. SVC is used with two hyperparameters such as kernel with value 'poly' and C with value 5.

3.4 Evaluation metrics

This study uses accuracy, precision, recall, and F1 score to evaluate the performance of machine learning and deep learning models. These metrics are based on different sections of a confusion matrix. The confusion matrix is a table that illustrates true positives (TP), true negatives (TN), false positives (FP), as well as false negatives (FN) events. When the model detects the event as positive (or 1) and the actual result is positive, the output is termed

TP. When the model detects the case as negative (or 0), and the actual output is shown as negative, then the output is termed TN. Once the model detects the instance as positive (or 1) but the actual result is negative, the output is termed FP. So when the model detects the instance as negative (or 0), but the actual output is positive, the output is termed FN. The more TPs and TNs there are (or fewer FPs and FNs), the more accurate the model is [53].

4 Results and discussions

This section contains the result of machine learning models with each feature set. Experiments are performed on an Intel Core i7 11th generation machine with Windows operating system. We used Jupyter notebook and Python language for experiments.

4.1 Results using CQT features

We deployed machine learning models for Bee, NoBee, and NoQueen classes using acoustic data. Table 5 shows the results with original CQT features, as well as, selective features using PCA, SVD, and Chi2. The performance of models with CQT features is not significant because the feature set generated by CQT has a low correlation with the target class. In comparison to the original CQT features, selective features show slightly better performance. The highest performance is achieved by RF and LR with a 0.45 accuracy score using the Chi2 feature selection. Despite little improvement in the performance of LR, SVC, and RF when used with selected features, the performance of machine learning models is poor.

4.2 Results with CG features

Table 6 shows the performance of learning models with CG features and similarly to CQT features, the performance of models with CG features is also not good. The highest accuracy

Table 5 Machine learning models performance with CQT features

Feature	Model	Acc.	Prec.	Recall	F1
Original CQT	KNN	0.41	0.46	0.35	0.23
	LR	0.40	0.43	0.35	0.23
	SVC	0.39	0.36	0.34	0.19
	RF	0.43	0.48	0.38	0.27
PCA + CQT	KNN	0.21	0.41	0.34	0.21
	LR	0.42	0.46	0.37	0.27
	SVC	0.41	0.48	0.36	0.24
	RF	0.43	0.54	0.38	0.29
Chi-2 + CQT	KNN	0.27	0.28	0.34	0.18
	LR	0.45	0.50	0.37	0.28
	SVC	0.44	0.49	0.36	0.26
	RF	0.45	0.56	0.38	0.29
SVD + CQT	KNN	0.27	0.28	0.34	0.18
	LR	0.42	0.45	0.37	0.26
	SVC	0.40	0.47	0.35	0.23
	RF	0.43	0.55	0.38	0.29

Table 6 Machine learning models performance with CG features

Feature	Model	Acc.	Prec.	Recall	F1
Original CG	KNN	0.55	0.56	0.55	0.54
	LR	0.47	0.46	0.45	0.45
	SVC	0.60	0.57	0.57	0.57
	RF	0.65	0.63	0.62	0.61
PCA + CG	KNN	0.56	0.56	0.56	0.55
	LR	0.52	0.50	0.49	0.49
	SVC	0.64	0.61	0.61	0.60
	RF	0.61	0.62	0.55	0.53
Chi-2 + CG	KNN	0.37	0.35	0.35	0.35
	LR	0.47	0.46	0.44	0.43
	SVC	0.52	0.51	0.48	0.47
	RF	0.56	0.54	0.52	0.51
SVD + CG	KNN	0.57	0.57	0.56	0.56
	LR	0.53	0.50	0.50	0.49
	SVC	0.65	0.62	0.61	0.60
	RF	0.61	0.61	0.55	0.52

with CG is 0.65 with RF using the original feature set and with SVC using the SVD technique. The results with CG are much better as compared to results with CQT features.

4.3 Performance of models with SC features

Table 7 shows the results of machine learning models with SC features, both original and selective. Models' performance is better than CQT using both original features, as well as, selective features. RF achieves the highest accuracy score of 0.58 with original SC features.

Table 7 Machine learning models performance with SC features

Feature	Model	Acc.	Prec.	Recall	F1
Original SC	KNN	0.51	0.49	0.48	0.48
	LR	0.42	0.45	0.38	0.33
	SVC	0.46	0.45	0.42	0.40
	RF	0.58	0.58	0.54	0.54
PCA + SC	KNN	0.52	0.51	0.49	0.49
	LR	0.44	0.46	0.46	0.44
	SVC	0.44	0.63	0.49	0.44
	RF	0.56	0.62	0.54	0.51
Chi-2 + SC	KNN	0.51	0.50	0.49	0.49
	LR	0.40	0.26	0.34	0.32
	SVC	0.47	0.48	0.43	0.38
	RF	0.57	0.59	0.54	0.53
SVD + SC	KNN	0.53	0.51	0.50	0.50
	LR	0.43	0.45	0.37	0.32
	SVC	0.48	0.43	0.44	0.35
	RF	0.55	0.61	0.53	0.50

In comparison with CG, models do not perform well with SC as CG helps to achieve a 0.65 accuracy score while SC it achieves only a 0.58 accuracy score. Both CQT and SC are not well compared to SC because these techniques are not significant to generate a correlated feature set.

4.4 ZCR features and machine learning models

The performance of machine learning models using ZCR features is shown in Table 8. Similar to CQT, and SP features, the performance of ZCR features is poor. The highest accuracy using ZCR is 0.61 from RF when trained using PCA-based ZCR features. On average, the performance of all models is poor.

4.5 Models' results using MFCC features

The results of machine learning models are not significant with CQT, SC, CG, and ZCR but are more improved with MFCC features. Table 9 shows the results of machine learning models with MFCC features indicating that RF and KNN both achieve 0.83 accuracy scores each. Overall MFCC features show significantly better results with all used models and become more significant with Chi2 feature selection. These results show that learning models with MFCC features can be used for bee detection.

Figure 11 shows the comparison with all features for original features and selective features using PCA, Chi2, and SVD. It indicates the performance of the models is better when models are used with MFCC features.

4.6 Results using hybrid feature set

For analyzing the performance of machine learning models, we combined all features used in this study whereby the combination means that selective features from each feature

Table 8 Machine learning models performance with ZCR features

Feature	Model	Acc.	Prec.	Recall	F1
Original ZCR	KNN	0.51	0.50	0.48	0.48
	LR	0.42	0.45	0.39	0.37
	SVC	0.44	0.49	0.41	0.32
	RF	0.59	0.57	0.57	0.56
PCA + ZCR	KNN	0.52	0.52	0.49	0.49
	LR	0.44	0.45	0.41	0.40
	SVC	0.48	0.44	0.43	0.34
	RF	0.61	0.59	0.57	0.56
Chi-2 + ZCR	KNN	0.39	0.13	0.33	0.19
	LR	0.39	0.13	0.33	0.19
	SVC	0.39	0.13	0.33	0.19
	RF	0.39	0.13	0.33	0.19
SVD + ZCR	KNN	0.51	0.51	0.49	0.48
	LR	0.42	0.46	0.39	0.36
	SVC	0.46	0.42	0.42	0.33
	RF	0.60	0.59	0.57	0.56

Table 9 Machine learning models performance with MFCC features

Feature	Model	Acc.	Prec.	Recall	F1
Original MFCC	KNN	0.83	0.81	0.79	0.79
	LR	0.70	0.67	0.67	0.66
	SVC	0.71	0.68	0.68	0.67
	RF	0.82	0.82	0.77	0.77
PCA + MFCC	KNN	0.82	0.81	0.79	0.79
	LR	0.70	0.66	0.66	0.66
	SVC	0.73	0.69	0.68	0.68
	RF	0.82	0.81	0.77	0.77
Chi-2 + MFCC	KNN	0.83	0.82	0.80	0.80
	LR	0.67	0.64	0.64	0.64
	SVC	0.72	0.69	0.69	0.68
	RF	0.83	0.82	0.80	0.79
SVD + MFCC	KNN	0.77	0.75	0.73	0.72
	LR	0.68	0.65	0.64	0.64
	SVC	0.69	0.67	0.65	0.64
	RF	0.81	0.81	0.77	0.77

selection approach like SVD, Chi2, and PCA are combined for each feature like CG, MFCC, etc. The objective is to increase the number of features for those techniques which generate a small feature set such as SC, and SCR, and also help to reduce the number of features for those techniques which generate a large feature set such as MFCC, CQT, and CG. Combining the selective features from SVD, Chi2, and PCA help to solve the problem of small and large feature sets at the same time.

Table 10 shows the results of HF with each model. Results show that the MFCC tends to be prudent in providing more accurate results as compared to other features. Combining the features selected by the SVD, PCA, and Chi2, KNN improves accuracy from 0.77 to 0.83, LR from 0.68 to 0.71, SVC from 0.69 to 0.73, and RF from 0.81 to 0.82. Accuracy with Hybrid ZCR is improved from 0.60 to 0.62 and Hybrid CG accuracy is improved from 0.65 to 0.67.

Figure 12 shows the comparison of machine learning models for each hybrid feature. It indicates that the models tend to show better results when hybrid MFCC features are used. This high performance is followed by hybrid CG features. The highest accuracy is shown by the KNN with hybrid MFCC features.

4.7 Results of K-fold cross-validation

To validate the significance of the hybrid feature approach, we deploy 10-fold cross-validation and results are shown in Table 11. In comparison to other features, results using the MFCC features are superior to all machine learning models. RF gives the highest mean accuracy of 0.82 with ± 0.01 standard deviation using hybrid MFCC features while KNN achieves a 0.81 mean accuracy with ± 0.00 standard deviation.

4.8 Deep learning models results

In this section, we present the deep learning model results in comparison to machine learning models. We deployed two deep learning models long short-term memory (LSTM) and

Fig. 11 Models performance comparison with each feature, (a) Original Feature, (b) PCA Feature, (c) Chi-2 Features, and (d) SVD Features

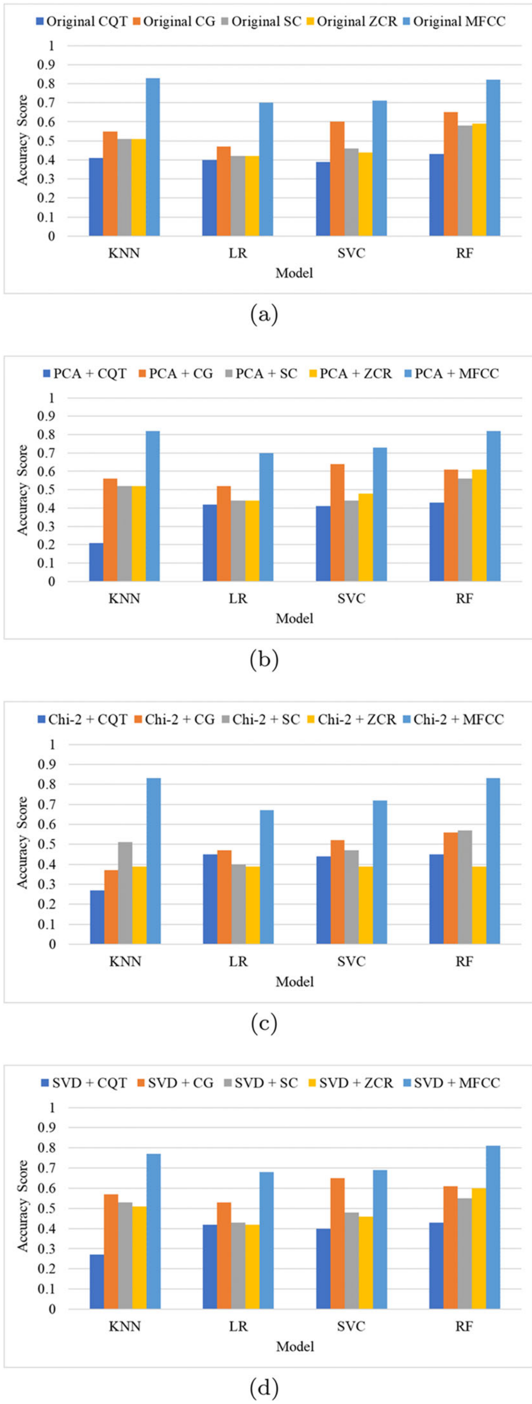


Table 10 Machine learning models performance with hybrid features

Feature	Model	Acc.	Prec.	Recall	F1
Hybrid CQT	KNN	0.42	0.34	0.35	0.22
	LR	0.44	0.52	0.38	0.29
	SVC	0.43	0.55	0.36	0.25
	RF	0.43	0.55	0.36	0.25
Hybrid CG	KNN	0.59	0.58	0.57	0.56
	LR	0.51	0.49	0.49	0.48
	SVC	0.65	0.62	0.61	0.61
	RF	0.67	0.65	0.63	0.62
Hybrid SC	KNN	0.57	0.60	0.53	0.52
	LR	0.41	0.41	0.37	0.32
	SVC	0.48	0.43	0.44	0.46
	RF	0.57	0.59	0.55	0.54
Hybrid ZCR	KNN	0.53	0.53	0.50	0.50
	LR	0.43	0.43	0.40	0.38
	SVC	0.52	0.61	0.51	0.49
	RF	0.62	0.61	0.59	0.59
Hybrid MFCC	KNN	0.83	0.81	0.79	0.79
	LR	0.71	0.66	0.67	0.66
	SVC	0.73	0.69	0.69	0.68
	RF	0.82	0.81	0.78	0.77

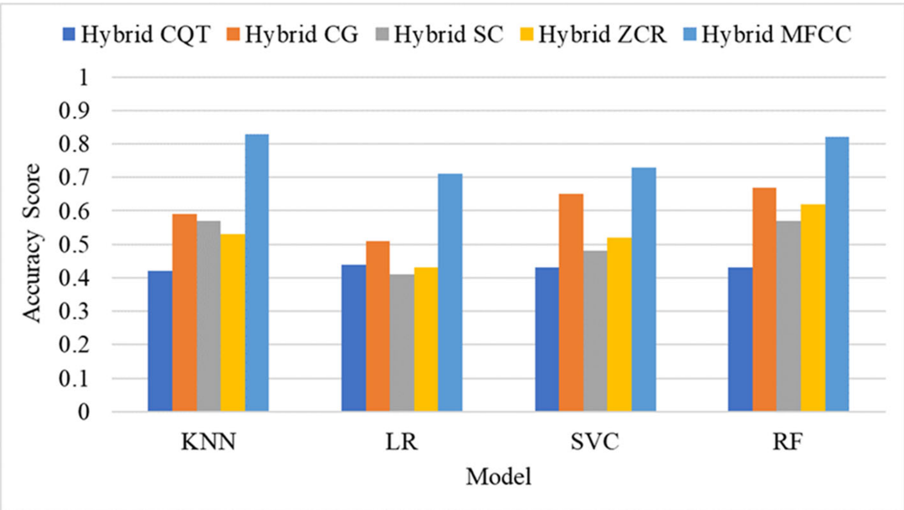


Fig. 12 Models' performance comparison with each hybrid feature set

Table 11 Machine learning models performance with 10-fold cross-validation using each feature

Feature	Model	HF CG	PCA	SVD	Chi2	Original CG
CG	KNN	0.59±0.01	0.49±0.01	0.60±0.01	0.34±0.01	0.60±0.01
	LR	0.51±0.01	0.44±0.00	0.51±0.00	0.47±0.00	0.48±0.00
	SVC	0.65±0.01	0.45±0.00	0.64±0.01	0.51±0.01	0.61±0.01
	RF	0.66±0.01	0.43±0.01	0.61±0.00	0.56±0.00	0.66±0.01
CQT	Model	HF CQT	PCA	SVD	Chi2	Original CQT
	KNN	0.41±0.00	0.39±0.02	0.41±0.00	0.41±0.00	0.41±0.00
	LR	0.44±0.01	0.40±0.00	0.44±0.01	0.44±0.00	0.44±0.01
	SVC	0.42±0.00	0.40±0.00	0.42±0.00	0.42±0.00	0.42±0.00
	RF	0.44±0.00	0.40±0.00	0.45±0.00	0.44±0.00	0.44±0.00
SC	Model	HF SC	PCA	SVD	Chi2	Original SC
	KNN	0.56±0.01	0.54±0.01	0.56±0.01	0.56±0.01	0.56±0.01
	LR	0.42±0.01	0.40±0.00	0.42±0.01	0.40±0.00	0.43±0.01
	SVC	0.48±0.00	0.48±0.00	0.48±0.00	0.47±0.00	0.47±0.00
	RF	0.57±0.00	0.49±0.01	0.56±0.01	0.57±0.00	0.57±0.00
ZCR	Model	HF ZCR	PCA	SVD	Chi2	Original ZCR
	KNN	0.55±0.01	0.55±0.01	0.55±0.01	0.40±0.00	0.55±0.01
	LR	0.43±0.00	0.43±0.00	0.43±0.00	0.32±0.06	0.43±0.00
	SVC	0.50±0.01	0.46±0.00	0.46±0.00	0.40±0.00	0.45±0.00
	RF	0.61±0.01	0.59±0.01	0.59±0.01	0.40±0.00	0.61±0.01
MFCC	Model	HF MFCC	PCA	SVD	Chi2	Original MFCC
	KNN	0.81±0.00	0.80±0.00	0.80±0.01	0.75±0.00	0.79±0.00
	LR	0.72±0.01	0.67±0.01	0.67±0.01	0.67±0.01	0.70±0.01
	SVC	0.72±0.01	0.75±0.01	0.70±0.01	0.67±0.01	0.71±0.01
	RF	0.82±0.01	0.83±0.01	0.83±0.01	0.80±0.00	0.82±0.01

convolutional neural networks (CNN) [42]. These models are deployed with their stat of the art architectures as shown in Table 12.

Results of deep learning models are present in Table 13. The performance of deep learning models is not good as compared to machine learning models. Deep learning models required a large dataset to perform significantly. Overall the performance of LSTM is good with HF MFCC as it achieved a 0.67 accuracy score. While we reduce the feature set size using PCA, SVD, or Chi2 it also reduces the performance accuracy of models.

4.9 Comparison with other studies

In this section, we compare previous studies with the proposed approach. We select recent studies on bee detection and compare their performance with the current study. We implemented the models from these studies using the same dataset which is used for the proposed approach. Selected studies used machine learning and deep learning approaches for bee detection such as study [11] where SVM was used to classify honeybee colonies into one with a queen and another without a queen. SVM algorithm with C-classification resulted in correct characterization. Error for test data of new queen is 9.28%. In another study by [3], MFCC features were used with the LSTM model, and energy was computed in MATLAB. ANOVA test for MFCC plus log energy averages were used to determine the difference

Table 12 Architecture of deep learning models

Model	Hyperparameters
LSTM	Embedding(1000,100, input_length=..)
	LSTM(64)
	Dense(8)
	Dense({3}, activation='softmax')
CNN	Embedding(1000,100, input_length=..)
	Conv1D(64, 3, activation='relu')
	MaxPooling1D(pool_size=3)
	Activation('relu')
	Flatten()
	Dense({3}, activation='softmax')
{loss='categorical_crossentropy}', optimizer='adam', epochs=100, batch_size=8	

between colonies with and without a queen. The result shows that queen and queen-less colonies are significantly different at P=0.001, and the confusion matrix shows high accuracy for LSTM with a 0.92 accuracy score while accuracy scores of 0.90 and 0.87 are obtained for MLP and logistic regression, respectively.

In [27], Spectrogram, FFT, S-transform, Kohonen Self-organizing Map (SOM) artificial neural network were used for queen right and queenless state. The result shows that classification by SOM is less successful while histogram shows good potential. In order to detect bee hives (Bee/No-Bee), an SVM classifier, and CNN classifier were used with the area under the curve (AUC) metric. Results exhibited that SVM is better classifier [35]. MFCC and machine learning algorithms suggest that they can be effectively used for bee status recognition by analyzing sound files obtained from inside the beehives [40]. The study

Table 13 Machine learning models performance with 10-fold cross-validation using each feature

Feature	Model	HF CG	PCA	SVD	Chi2	Original CG
CG	LSTM	0.56	0.42	0.39	0.42	0.45
	CNN	0.49	0.45	0.39	0.44	0.45
CQT	Model	HF CQT	PCA	SVD	Chi2	Original CQT
	LSTM	0.59	0.49	0.45	0.48	0.48
	CNN	0.55	0.48	0.44	0.45	0.45
SC	Model	HF SC	PCA	SVD	Chi2	Original SC
	LSTM	0.45	0.39	0.39	0.39	0.43
	CNN	0.42	0.40	0.41	0.40	0.41
ZCR	Model	HF ZCR	PCA	SVD	Chi2	Original ZCR
	LSTM	0.49	0.43	0.43	0.42	0.41
	CNN	0.42	0.39	0.40	0.40	0.39
MFCC	Model	HF MFCC	PCA	SVD	Chi2	Original MFCC
	LSTM	0.67	0.58	0.58	0.59	0.64
	CNN	0.59	0.52	0.52	0.51	0.54

Table 14 Performance comparison with existing studies

Ref.	Model	Accuracy	Precision	Recall	F1
[27]	ANN	0.66	0.65	0.63	0.64
[40]	LR	0.44	0.50	0.40	0.45
[11]	SVM	0.43	0.55	0.36	0.25
[35]	SVM	0.44	0.44	0.44	0.44
[3]	LSTM	0.56	0.45	0.39	0.42
This study	KNN	0.83	0.81	0.79	0.79

used LR for honey bee recognition with MFCC features. The models from these studies are implemented as per the given architecture and relevant parameters and results are compared in Table 14.

4.10 Statistical T-test

Results indicate that the bee detection using hybrid MFCC features from PCA, Chi2, and SVD shows superior results, and to show its significance over other used techniques, a statistical T-test is performed [37, 41]. The performance of the machine learning model is significant using MFCC in comparison to CQT, CG, SC, and ZCR. To show its statistical significance, the following hypotheses are formulated.

- Null Hypothesis (H_0): The use of hybrid MFCC features is not significant for bee detection as compared to other approaches.
- Alternative Hypothesis(H_a): The use of hybrid MFCC features is significant for bee detection as compared to other approaches.

We perform a statistical significance T-test on hybrid MFCC in comparison to SVD-based MFCC, PCA-based MFCC, and Chi2-based MFCC. Table 15 shows the results of the t-test. T-test gives output as t-stats (t) and critical value (CV), if the cv value is greater than or equal to t value then T-test accepts the null hypothesis (H_0) else it accepts the alternative hypothesis (H_a). In all cases, the T-test accepts the alternative hypothesis and rejects the null hypothesis (H_0) which indicates that the hybrid MFCC features-based approach is statistically significant in comparison with other approaches.

5 Conclusions

When a colony suffers queenlessness, the whole colony collapses or dies within a few months. Detecting the presence and absence of the queen and bee population, therefore,

Table 15 Results of T-test for hybrid MFCC features approach

Comparison	t	CV	(H_0)
Hybrid MFCC vs PCA MFCC	2.611	0	Rejected
Hybrid MFCC vs SVD MFCC	5.490	0	Rejected
Hybrid MFCC vs Chi-2 MFCC	0.426	0	Rejected
Hybrid MFCC vs MFCC	1.576	0	Rejected

indicates the health status of a colony. Moreover, honey bees have a central role in the pollination of cultivated crops and ecological balance. Understanding the importance of honeybees, this study considers the problem of bee classification into 'Bee', 'NoBee', and 'NoQueen' using the acoustic data obtained from inside the beehives. For bee detection, this study investigates the use of SC, ZCR, MFCC, CG, and CQT features with and without feature selection using PCA, SVD, and Chi2. To examine different colony statuses or the aforementioned bee classes, results suggest that using MFCC features with machine learning models tends to outperform other features, both when original features are utilized and when selective features are employed. According to MFCC features, RF and KNN both achieved high accuracy scores of 0.83 when compared to other models such as LR and SVC. The proposed approach where selective features from PCA, Chi2, and SVD are combined shows the highest accuracy of bee detection. Cross-validation and t-test results corroborate the superior performance with hybrid features. Currently, using the deep learning models with the acoustic dataset requires a substantial amount of resources, we intend to deploy deep learning models to obtain better results in the future. Furthermore, we are positive that MFCC features can be used to detect acoustic patterns of different colony conditions such as pre-swarming behavior, varroa mites/pests/ parasitic infections, and size of the population, among others.

Data Availability The used dataset is publicly available on Kaggle. https://www.kaggle.com/yevheniiklymenko/beehive-buzz-anomalies

Declarations

Conflict of Interests The authors declare that they have no conflict of interest.

References

1. Abrol DP (2012) Pollination biology: biodiversity conservation and agricultural production. Springer
2. Andreev N (2017) Lactic acid fermentation of human excreta for agricultural application. CRC Press, Boca Raton
3. Anwar O, Keating A, Cardell-Oliver R, Datta A, Putrino G (2022) Design and development of low-power, long-range data acquisition system for beehives-beedas. Comput Electron Agric 201:107281
4. Bachu R, Kopparthi S, Adapa B, Barkana B (2008) Separation of voiced and unvoiced using zero crossing rate and energy of the speech signal. In: American Society for Engineering Education (ASEE) zone conference proceedings. American Society for Engineering Education, pp 1–7
5. Beehive buzz anomalies dataset https://www.kaggle.com/yevheniiklymenko/beehive-buzz-anomalies. Accessed 20 Feb 2022
6. Birajdar GK, Patil MD (2020) Speech/music classification using visual and spectral chromagram features. J Ambient Intell Humaniz Comput 11(1):329–347
7. Bortolotti L, Costa C (2014) Chemical communication in the honey bee society, Neurobiology of chemical communication
8. Braga AR, Gomes DG, Rogers R, Hassler EE, Freitas BM, Cazier JA (2020) A method for mining combined data from in-hive sensors, weather and apiary inspections to forecast the health status of honey bee colonies. Comput Electron Agric 169:105161
9. Calderone NW (2012) Insect pollinated crops, insect pollinators and us agriculture: trend analysis of aggregate data for the period 1992–2009. PloS one 7(5):e37235
10. Cecchi S, Spinsante S, Terenzi A, Orcioni S (2020) A smart sensor-based measurement system for advanced bee hive monitoring. Sensors 20(9):2726
11. Cejrowski T, Szymański J, Mora H, Gil D (2018) Detection of the bee queen presence using sound analysis. In: Asian conference on intelligent information and database systems. Springer, pp 297–306
12. Chakroborty S, Roy A, Saha G (2006) Fusion of a complementary feature set with mfcc for improved closed set text-independent speaker identification. In: 2006 IEEE International conference on industrial technology. IEEE, pp 387–390

13. Chi-square test for feature selection in machine learning (2022) https://towardsdatascience.com/chi-square-test-for-feature-selection-in-machine-learning-206b1f0b8223. Accessed 20 Feb 2022
14. Debauche O, El Moulat M, Mahmoudi S, Boukraa S, Manneback P, Lebeau F (2018) Web monitoring of bee health for researchers and beekeepers based on the internet of things. Procedia Comput Sci 130:991–998
15. Ellis DP (2007) Classifying music audio with timbral and chroma features
16. Epa (2020) Pesticide issues in the works: Honey bee colony collapse disorder. (environmental protection agency). http://www.epa.gov/pesticides/about/intheworks/honeybee.hlm. Accessed 20 Feb 2022
17. Eskov E, Toboev V (2011) Changes in the structure of sounds generated by bee colonies during sociotomy. Entomol Rev 91(3):347–353
18. Fallucchi F, Zanzotto FM (2009) Svd feature selection for probabilistic taxonomy learning. In: Proceedings of the workshop on geometrical models of natural language semantics, pp 66–73
19. Ferrari S, Silva M, Guarino M, Berckmans D (2008) Monitoring of swarming sounds in bee hives for early detection of the swarming period. Comput Electron Agricul 64(1):72–77
20. Flores JM, Gil-Lebrero S, Gámiz V, Rodríguez MI, Ortiz MA, Quiles FJ (2019) Effect of the climate change on honey bee colonies in a temperate mediterranean zone assessed through remote hive weight monitoring system in conjunction with exhaustive colonies assessment. Sci Total Environ 653:1111–1119
21. Frings H, Little F (1957) Reactions of honey bees in the hive to simple sounds. Science 125(3238):122–122
22. Giannakopoulos T, Pikrakis A (2014) Chapter 4 - audio features. In: Giannakopoulos T, Pikrakis A (eds) Introduction to audio analysis. Academic Press, Oxford, pp 59–103. https://www.sciencedirect.com/science/article/pii/B9780080993881000042
23. Gouyon F, Pachet F, Delerue O et al (2000) On the use of zero-crossing rate for an application of classification of percussive sounds. In: Inproceedings of the COST G-6 conference on Digital Audio Effects (DAFX-00), Verona, Italy, vol 5, p 16. Citeseer
24. Hasan MR, Jamil M, Rahman M et al (2004) Speaker identification using mel frequency cepstral coefficients. Variations 1(4):565–568
25. Henry M, Beguin M, Requier F, Rollin O, Odoux J.-F., Aupinel P, Aptel J, Tchamitchian S, Decourtye A (2012) A common pesticide decreases foraging success and survival in honey bees. Science 336(6079):348–350
26. Hossan MA, Memon S, Gregory MA (2010) A novel approach for mfcc feature extraction. In: 2010 4th International conference on signal processing and communication systems. IEEE, pp 1–5
27. Howard D, Duran O, Hunter G, Stebel K (2013) Signal processing the acoustics of honeybees (apis mellifera) to identify the queenless state in hives. Proc Instit Acoust 35(1):290–297
28. Hung K-LJ, Kingston JM, Albrecht M, Holway DA, Kohn JR (2018) The worldwide importance of honey bees as pollinators in natural habitats. Proc R Soc B: Biol Sci 285(1870):20172140
29. Jafor Sadeek Quaderi S, Afrin Labonno S, Mostafa S, Akhter S (2022) Identify the beehive sound using deep learning, arXiv e-prints, pp arXiv–2209
30. Liu H, Setiono R (1995) Chi2: feature selection and discretization of numeric attributes. In: Proceedings of 7th IEEE international conference on tools with artificial intelligence. IEEE, pp 388–391
31. Malhi A, Gao RX (2004) Pca-based feature selection scheme for machine defect classification. IEEE Trans Instrum Measur 53(6):1517–1525
32. Mezquida DA, Martínez JL (2009) Platform for bee-hives monitoring based on sound analysis. a perpetual warehouse for swarm's daily activity. Span J Agric Res 7(4):824–828
33. Michelsen A, Kirchner WH, Lindauer M (1986) Sound and vibrational signals in the dance language of the honeybee, apis mellifera. Behav Ecol Sociobiol 18(3):207–212
34. Muda L, Begam M, Elamvazuthi I (2010) Voice recognition algorithms using mel frequency cepstral coefficient (mfcc) and dynamic time warping (dtw) techniques, arXiv: 1003.4083
35. Nolasco I, Benetos E (2018) To bee or not to bee: investigating machine learning approaches for beehive sound recognition, arXiv: 1811.06016
36. Ntalampiras S, Potamitis I, Fakotakis N (2012) Acoustic detection of human activities in natural environments. J Audio Eng Soc 60(9):686–695
37. Omar B, Rustam F, Mehmood A, Choi GS et al (2021) Minimizing the overlapping degree to improve class-imbalanced learning under sparse feature selection: application to fraud detection. IEEE Access 9:28101–28110
38. Qandour A, Ahmad I, Habibi D, Leppard M (2014), Remote beehive monitoring using acoustic signals
39. Rigatti SJ (2017) Random forest. J Insur Med 47(1):31–39

40. Robles-Guerrero A, Saucedo-Anaya T, González-Ramérez E, Galván-Tejada CE (2017) Frequency analysis of honey bee buzz for automatic recognition of health status: a preliminary study. Res Comput Sci 142:89–98
41. Rustam F, Ashraf I, Mehmood A, Ullah S, Choi GS (2019) Tweets classification on the base of sentiments for us airline companies. Entropy 21(11):1078
42. Rustam F, Ishaq A, Munir K, Almutairi M, Aslam N, Ashraf I (2022) Incorporating cnn features for optimizing performance of ensemble classifier for cardiovascular disease prediction. Diagnostics 12(6):1474
43. Ruvinga S, Hunter G, Duran O, Nebel J-C (2021) Use of lstm networks to identify a queenlessness in honeybee hives from audio signals. In: 2021 17th International Conference on Intelligent Environments (IE). IEEE, pp 1–4
44. Schörkhuber C, Klapuri A, Sontacchi A (2013) Audio pitch shifting using the constant-q transform. J Aud Eng Soc 61(7/8):562–572
45. Shafique R, Siddiqui H-U-R, Rustam F, Ullah S, Siddique MA, Lee E, Ashraf I, Dudley S et al (2021) A novel approach to railway track faults detection using acoustic analysis. Sensors 21(18):6221
46. Sharif MZ, Wario F, Di N, Xue R, Liu F (2020) Soundscape indices: new features for classifying beehive audio samples. Sociobiology 67(4):566–571
47. Sharif MZ, Di N, Liu F (2021) Monitoring honeybees (apis spp.)(hymenoptera: Apidae) in climate-smart agriculture: a review. Appl Entomol Zool, 1–15
48. Sperandei S (2014) Understanding logistic regression analysis. Biochemia Medica 24(1):12–18
49. Srinivasu PN, Balas VE (2021) Self-learning network-based segmentation for real-time brain mr images through haris. PeerJ Comput Sci 7:e654
50. Suthaharan S (2016) Support vector machine. In: Machine learning models and algorithms for big data classification. Springer, pp 207–235
51. Van der Maaten L, Hinton G (2008) Visualizing data using t-sne. J Mach Learn Res 9:11
52. Voudiotis G, Moraiti A, Kontogiannis S (2022) Deep learning beehive monitoring system for early detection of the varroa mite. Signals 3(3):506–523
53. Vulli A, Srinivasu PN, Sashank MSK, Shafi J, Choi J, Ijaz MF (2022) Fine-tuned densenet-169 for breast cancer metastasis prediction using fastai and 1-cycle policy. Sensors 22(8):2988
54. Wun S, Horner A, Wu B (2014) Effect of spectral centroid manipulation on discrimination and identification of instrument timbres. J Audio Eng Soc 62(9):575–583
55. Yang J, Das RK (2019) Low frequency frame-wise normalization over constant-q transform for playback speech detection. Digital Signal Process 89:30–39
56. Yigit H (2013) A weighting approach for knn classifier. In: 2013 international conference on electronics, computer and computation (ICECCO). IEEE, pp 228–231
57. Zgank A (2020) Bee swarm activity acoustic classification for an iot-based farm service. Sensors 20(1):21