

DECODING BEE COLONY STATES: ANALYSIS OF FEATURE EXTRACTION METHODS IN ACOUSTIC SIGNATURE CLASSIFICATION FOR QUEENLESS DETECTION

Punith Kumar^{1*}, Champa H N², Akshay Prasad³

^{1*}Research Scholar, Department of Computer Science and Engineering, University Visvesvaraya College of Engineering, Bangalore, India-560001. E-mail: punithkumar87@gmail.com

²Professor, Department of Computer Science and Engineering, University Visvesvaraya College of Engineering, Bangalore, India-560001. E-mail: champahn@yahoo.co.in

³PG Student, Department of Computer Science and Engineering, University of Visvesvaraya College of Engineering, Bengaluru, India-560001. Email: akshaysprasad28@gmail.com

Manuscript received February 15, 2025; revised XXXX X, 20XX.

Abstract

In the domain of apicultural colony management, the existence of queen in a beehive is indispensable for sustenance of the hive's overall health and productivity. Nonetheless, conventional methodologies for evaluating the status of the queen are labor-intensive and may intrude upon the colony. This research endeavors to explore the application of acoustic pattern classification to differentiate between queenright (the residence of a queen) and queenless (the non-residence of a queen) conditions within bee colonies. Through the monitoring and examining the sound signals pour out by bees, this work seek to discern distinct acoustic characteristics that are associated with each condition. Sophisticated algorithms in machine learning are utilized to effectively sort these sound patterns with remarkable precision, especially KNN and RF models with 94.7% and 93.6% mean accuracy respectively using traditional MFCC's which is widely used feature extraction method. Further this work extended by employing Gabor-MFCC method, a new approach which is a hybrid method demonstrated improved accuracy and noise resilience by 2% to 3% across the models. Then GFCC and MFCC spectral fusion methods produce further improvements in accuracy. The findings indicate that specific sound frequencies and patterns serve as reliable indicators of the queen's status, thereby offering a non-invasive and automated solution for apiarists. This methodology presents substantial prospects for the enhancement of hive management practices by facilitating real-time, continuous surveillance of queen status, ultimately contributing to the sustainability and productivity of apicultural operations.

Keywords: Bee acoustics, Queenless state, Gabor-Mfcc, Sensitivity analysis, MFCC spectral fusion, Gammatone filterbank

1. INTRODUCTION

Honey bees are arguably among the most vital insects on earth. Their significance extends

beyond the production of honey, bee's wax, its jelly, and its propolis, as they play a crucial role in pollination for both wild and cultivated plants [1].

In recent times, there has been a significant rise in bee mortality, leading to severe ecological and economic impacts [2]. This decline is largely attributed to collapse disorder, which is a condition marked by the sudden and unexplained disappearance of honeybees from their hives [3],[4]. In this context, the need for an intensive monitoring process for honey bees is evident to better understand the issues and underlying causes of bees mortality. Monitoring bee colonies involves a detailed examination of their health and behavior, especially regarding *Apis mellifera*, to protect their welfare and elevate their productivity. This methodology entails the employment of a variety of tools and techniques to collect data on critical parameters such as population density, hive conditions, disease incidence, and environmental factors. Beekeepers and researchers engage in colony monitoring to facilitate informed management strategies, identify potential problems at an early stage, and encourage sustainable apicultural practices. Beehives are frequently situated in isolated rural areas or even within urban settings, where the practice of rooftop urban beekeeping has experienced a notable increase in popularity. All these locations are remote and pose challenges for human monitoring efforts. Regardless of their geographic setting, the phenomenon of bee swarming remains a pivotal concern for beekeepers. Consequently, a bee activity monitoring system can provide alerts to beekeepers regarding swarming events. Various indicators can be employed for the monitoring of bee activity, with sound being one of the notable metrics. This auditory data can facilitate the management of beehives situated in remote locations. It is feasible to classify bee swarm activity from audio signals by utilizing such methodologies.

The queen bee constitutes the most pivotal member of a bee colony, serving a crucial function within the hive as the sole individual responsible for oviposition, producing approximately 2000 eggs daily. The queen bee sustains the structural integrity of the colony predominantly through the secretion of pheromones, which facilitate communication among the colony's constituents. It is not unusual to encounter colonies devoid of a queen, as the queen bee may inadvertently perish due to disease or substandard apicultural practices. Upon the loss of their queen, the behavior of the remaining bees typically shifts from a state of organized activity to one characterized by disorder. The absence of the queen bee during a phase when no drones are present can precipitate the demise

of the colony. In addition to conventional methodologies, such as beehive inspections, there exist no alternative strategies for the identification of a queenless state.

Identifying the queen bee constitutes a pivotal element of apicultural colony surveillance, given that the queen fulfills a fundamental function in the overall productivity and well-being of the hive. The automated detection of the queen bee within apiary colonies significantly enhances the efficiency, precision, and non-invasive characteristics of monitoring and management methodologies. In the field of precision apiculture, researchers face a significant challenge in the automated identification of hive conditions to analyze the health status of bee populations using specialized hardware [5]. Researchers are especially interested in the sound signatures of bee colonies because they may be used to efficiently examine and predict key stages within

the colonies [17],[18].

The swarming process, in which a colony modifies its regular pursuits and generates a distinctive buzz prior to the queen departs with little swarm, is one example of an auditory pattern seen in bee colonies [20]. On the other hand, when a colony is queenless, it makes a distinctive sound because the stress of not having a queen causes the colony members to act erratically. Furthermore, high temperatures inside the hive may adversely affect the colony's sound transmissions [21], [22]. Acoustic pattern classification for the identification of queenless conditions within bee colonies represents an avant-garde methodology that employs sound analysis to discern distinct acoustic signatures indicative of the queen bee's absence.

In this study, five distinct classification ML algorithms are trained with the data in which features are extracted using mfcc, gabor-mfcc, gfcc and mfcc spectral fusion methods separately and subjected to comparative analysis to ascertain a model that exhibits optimal performance. Different criteria are assessed to judge the capability of the models. Further performed sensitivity analysis to identify features which are most influential in prediction, in-turn guides in feature selection.

2. LITERATURE REVIEW

Numerous scholarly investigations indicate that machine learning algorithms, alongside traditional methodologies employed for sound recognition, are proficiently modified for the classification of acoustic patterns within bee hives; such approaches are currently being applied with significant efficacy to discern the states of bee colonies, resulting in elevated rates of correct classifications. Moreover, the structure of convolutional neural networks has revealed impressive capability in audio classification and has served to identify acoustic patterns among bees. The subsequent paragraphs will critically examine the most salient research pertaining to the identification of acoustic patterns within the context of precision beekeeping.

Nolasco et al. conducted a comparative analysis of two methodologies for the identification of bee vocalizations, specifically utilizing a CNN model in combination with a SVM [6]. This dataset comprises a curated selection of beehive buzz recorded under a variety of conditions across two distinct projects. The models under scrutiny possess the capability to discern bee sounds amidst extraneous auditory stimuli, such as vehicular traffic or avian calls. MFCCs and Mel spectra were employed for the purpose of feature extraction. These are combined by ingenious feature extraction technique based on the Hilbert Huang transform [23]. Moreover, the audio signals were segmented into varying sizes to facilitate a comprehensive performance analysis. In the findings of this investigation, the SVM model demonstrated superior performance relative to the CNN.

In the domain of sound classification pertaining to bees and external auditory stimuli, Kim et al. conducted a comparative analysis of conventional machine learning models, specifically Random Forest, Support Vector Machine, and Extreme Gradient Boosting, juxtaposed with CNN architectures, VGG-13 and the Shallow CNN. The methods for feature extraction includes Mel spectrograms, MFCCs, and constant Q transformation. The findings indicate that the VGG-13 architecture exhibits superior performance, attaining an accuracy rate of 91% [7].

To differentiate swarming behavior from normal activity within bee colonies, Zgank et al. devised an approach founded on Hidden Markov Models (HMM), which are extensively utilized in the realm of human voice recognition, employing MFCCs for extracting the features [8]. This model demonstrated a classification accuracy of over 80% in classifying swarming activity of bees. Zgank et al. further refined the investigation by comparing the efficacy of HMM and Gaussian Mixture Models (GMM), additionally scrutinized the MFCC values and Linear method of predictive coding as a feature extraction methodologies in this endeavor [9]. A range of performance metrics was provided, revealing that the remarkable accuracy achieved using the HMM model with MFCC features. In a subsequent study, Zgank et al. advanced previous findings by implementing deep neural networks alongside MFCCs [10]. An examination led by Dimitrios et al. [11] and others analyzed different machine learning models to discern swarming activity from non-swarming activity. This work evaluates the performances of KNN, SVM, and U-Net CNN by integrating buzz samples with temperature measurements inside the hive, as well as humidity and temperature readings external to the hive. These auditory signals were subjected to low pass filtering. The results elucidate that SVM outperforms CNN in terms of accuracy.

Support Vector Machine (SVM's) and MFCC's have been employed to ascertain circadian rhythms within bee colonies by the authors Cejrowski et al., with the objective of classifying bee sound activity during diurnal and nocturnal periods, as well as identifying the timeframe of minimal activity throughout the day [12]. The tracking system features a Raspberry Pi (RPI) computer associated with an analog microphone; audio snippets were documented at the rate of 3000 Hz and 12 bit resolution every fifteen minutes. The period of least activity observed within the colony was determined to occur between 11 pm and 4 am. The SVM model, along with MFCCs were used to categorize the sound. This work demonstrated a high level of accuracy in distinguishing between flight sounds along with environmental noises, such as avian sounds and back ground interference [13]. Notably, some researchers have used accelerometers to capture the sounds of bees within the colony and applied discriminant functions to successfully predict swarming events [14]. In another study, machine learning techniques were employed to differentiate between bee buzz, chirping and ambient noise [15]. Additionally, beehive sounds are combined with ML methods to make systems capable of identifying variety of colony states [16]. Acoustics based classification of bee's activity, offering a preliminary comparison of extracted features as inputs to a CNN [19]. Specifically, the scenario of a lonely hive was analyzed using the dataset collected from a real-world situation. Multiple experiments with different setups were conducted to evaluate the system's performance, and the results have confirmed its potential effectiveness. Using self organizing maps, Howard et al. used an alternative method to categorize queenless states in bee colonies [24]. They used the S-Transform and power spectral density for feature extraction. The features of the dataset that was employed may have contributed to the model's inability to correctly categorize the beehive state, as the results showed. Singular value decomposition (SVD) was used by Robles-Guerrero et al. to find commonalities in queenless situations across strong and weak hives in a multi-label classification issue [25].

To develop personalized anomaly detection methods at the colony level, assumptions are made

that a colony's sound during different events carries unique characteristics [26]. Peng et al. [27] offered an alternative method for enhancing the recognition of queenless conditions using Wiener filter in order to lower the signal-to-noise ratio. Comparing the filtered signal to the unfiltered one, it is found that filter improved the classification accuracy by 12%. To enhance the detection of queenless states using CNNs which are also efficient disease identification deep learning methods[30], Orłowska et al. proposed a simple transformation aimed at improving classification performance. A two-step dimension reduction applied to the spectrograms, is said to offer better generalization. As a result, the CNN achieved an accuracy of 96% [28]. Cejrowski et al. conducted another study to detect queen's existence. In their series of experiments aimed at simulating the absence of queen's, they used Linear predictive coding for extracting features and employed a support vector machine (SVM) as the learning algorithm. Their classification approach successfully identified acoustic patterns distinguishing healthy colonies from queenless ones [29]. A self-tuning convolutional kernel (STCK) algorithm has been introduced to improve feature extraction in CNN-based human-computer interaction systems, incorporating Bark-spectrum-based input and novel data augmentation techniques to enhance model generalization [32]. Experimental results demonstrated that the proposed method achieved the lowest training loss and outperformed baseline CNN models without STCK or data augmentation, reducing test recognition error rates by 3.9% and 1%, respectively.

Acoustic event classification (AEC) is essential for various modern applications, including security monitoring, multimedia retrieval, and smart city infrastructure. Two significant challenges in AEC are the limited availability of labeled training data and the inconsistent durations of input audio signals. To tackle these issues, deep learning framework that integrates a pre-trained DenseNet-121 model with a Multilayer Perceptron (MLP) for classifying audio events is introduced [33]. To enhance the diversity of the training data, two augmentation techniques such as time stretching and pitch shifting are applied. These methods help generate more varied training samples from the original audio data. Each audio signal is then transformed into a spectrogram, converting the time-domain information into a consistent image format that retains key features such as energy variation, frequency shifts, and amplitude patterns. The DenseNet-121 model, utilized through transfer learning, extracts high-level features from these spectrogram images, improving classification accuracy while minimizing the need for extensive computational resources. Experimental validation was conducted using three benchmark datasets: Soundscapes1, Soundscapes2, and UrbanSound8K. The proposed model achieved notable performance gains, with F1-scores of 80.7%, 87.3%, and 69.6% on these datasets, respectively, outperforming existing baseline methods. Another study addressed chord recognition in noisy musical environments by introducing a time-correlation support vector machine (TCSVM) to capture temporal relationships between adjacent chords [34]. To minimize interference from vocals and background noise, the accompaniment is separated prior to feature extraction. The method was evaluated using robust logarithmic pitch class profile (RLPCP) features and compared with traditional SVM across various noise levels. Experiments conducted on the MIREX'09 dataset demonstrate that the TCSVM approach, in combination with RLPCP, offers improved noise robustness and higher chord identification accuracy

compared to baseline methods.

3. METHOD

A. Dataset Description

This study utilize the dataset in <https://data.mendeley.com/datasets/t9prmbmdfnthe>. This dataset consists of sound recordings from five different Carniola honeybee (*Apis Mellifera Carnica*) colonies that were chosen at random. The characteristics that are chosen, varied significantly; two of the colonies had high population (about 60,000 bees) and were in good health, while the other two colonies had a moderate population (about 40,000 bees) and one less populated queenless colony (30 thousand bees). Queens are removed from a healthy colony and moderately populated colony, following the artificial induction of the queenless condition. Table 1 delineates a summary of the pertinent information regarding the bee hives.

Table 1: Assessment of the bee colony population sizes

	Hive 1	Hive 2	Hive 3	Hive 4	Hive 5
Bees(In Thousand)	60	60	30	40	40
Hive States	QR	QL	LPQL	QL	QR

B. Extracting Features and Data Preprocessing

1. Traditional MFCC's: For extracting features, a conventional approach utilizing Mel-frequency cepstral coefficients (MFCCs) was employed. MFCCs have proven effective in the initial processing of acoustic signals from bee colonies. The traditional method involves several steps: Pre-emphasis filtering, windowing the input signal, performing rapid fourier transform, mapping the frequencies to Mel scale, creating a triangular filter bank, and then executing the inverse of Discrete cosine transform. MATLAB's mfcc function was used to compute the MFCCs with appropriate parameters. Following feature extraction, the MFCC mean values were calculated to in order to decrease the size of dataset, aiming to reduce model training time and enable execution on low-resource computing systems. The data was then normalized using the Z-scores. This resulting dataset contains 12 features with 720 samples, 144 samples per colony. The output categories are Queenright (QR), Queenless (QL), and Less populated Queenless colonies (LPQL). This is extended by computing the first order derivative (delta) and second order derivative (delta-delta) to capture temporal dynamics. This will help in improving classification by incorporating how MFCC change over time.

For a given MFCC coefficient sequence c_t :

$$\Delta c_t = \frac{\sum_{n=1}^{\{N\}} n (c_{\{t+n\}} - c_{\{t-n\}})}{2 \sum_{n=1}^{\{N\}} n^2} \quad (1)$$

The same structure applies to delta-delta also.

2. MFCC Spectral fusion: Traditional approaches use the MFCC, which is a compact representation of spectral information. That being said, the MFCCs alone are not always discriminative, notably in complex practices. The spectral features are the centroid, bandwidth, rolloff, and flatness that further add the characteristics of frequency distribution that, in turn, enhances the robustness of classification models. For MFCC computation, The input signal undergoes pre-emphasis to enhance high-frequency components. Short-time Fourier Transform (STFT) is applied to obtain frequency-domain representations. A mel-scale filterbank is used to extract perceptually relevant frequency components. Discrete Cosine Transform (DCT) is applied to obtain decorrelated cepstral coefficients. Spectral features calculated are, Spectral Centroid which represents the center of mass of the spectrum and indicates the brightness of the sound. Spectral bandwidth Measures the spread of frequencies around the centroid, indicating the complexity of the spectrum. Spectral rolloff determines the frequency below which a specified percentage (e.g., 85%) of the total spectral energy is contained, providing insight into signal sharpness. Spectral flatness quantifies the uniformity of the spectrum, distinguishing between tonal and noisy signals. Extracted features are concatenated to form a comprehensive feature vector. z-score normalization are applied to standardize feature distributions. PCA dimensionality reduction techniques is employed to remove redundancy and enhance discriminative power.

3. GFCC (Gammatone Frequency Cepstral Coefficients): With Mel-Frequency Cepstral Coefficients (MFCC) being a widely used approach, traditional audio analysis techniques are also limited in their effectiveness when there is noise. GFCC combines the advantages of Gammatone filterbank with Logarithmic transformation, tracking the perceptually relevant peaks and modal responses in the acoustic signal akin to the human auditory system. The computation of GFCC involves a sequence of signal processing stages, simulating the acoustic filtering functions of the human cochlea. Pre-emphasis and framing is applied to the input signal to combat noise and to maintain temporal continuity. Gammatone filterbank is applied to decompose the signal into its frequency components. The filterbank equation is given by,

$$G(f) = t^{(n-1)} e^{(-2\pi bt)} \cos(2\pi f_c t) \quad (2)$$

where f_c represents the center frequency, b is the bandwidth factor, and n is the order of the filter.

The output for each filter is then processed by a nonlinear function (power-law) which simulates the compression characteristics of the human auditory system. The log filterbank energies are subjected to Discrete Cosine Transform (DCT) to extract the GFCC features. The nonlinear filterbank structure of GFCCs ensures stability in the presence of background noise. The gammatone filterbank models the cochlear response more faithfully than the mel filterbank.

4. Gabor-MFCC's: A robust hybrid feature extraction method combines Gabor filters [31] with cepstral coefficients which is an advancement to traditional MFCC's which incorporates several advanced techniques. In this study, multi-scale and multidirectional gabor filters are used to

analyze the audio signals. These will capture frequency along with spatial information, which facilitates a balanced representation of intra and interclass features. Then these filters are integrated with mfcc to enhance the robustness of extraction process. To reduce noise in the signals, a combination of hard and soft threshold wavelet denoising techniques are used in preprocessing to improve the quality of features. Table 2 provides a brief overview of the hyperparameters for the examined machine learning models.

Table 2: Hyperparameters of the ML models

NB	GuassainNBclassier Alpha : 0.1, Fitprior : true, Classprior : None
KNN	KNeighborsclassifier Weight : distance, Neighbors : 25, Algorithm : ball tree, Pvalue : minkowski
LSTM	LSTMclassifier Units : 50, Batchsize : 32, Dropout rate : 0.3, Activation function : ReLU Learning rate : 0.001, Optimizer : Adam, Sequence length : 10,
RNN	RNNclassifier Learning rate : 0.001, Batchsize : 32, Dropout rate : 0.3, Sequence length : 10, Activation function : ReLU, Regularization : L2 Optimizer : Adam,
RF	Randomforestclassifier Number of estimators : 50, Criterion : gini Bootstrap : true, Max depth : 10

4. RESULTS

A Sensitivity Analysis

A comprehensive sensitivity analysis was undertaken to assess the influence of individual features on the performance of the machine learning models. This analysis employed a

permutation method, utilizing the `permutation_importance` function from the scikit-learn library. By permuting the values of each feature, we were able to quantify the impact of these features on the model's predictive accuracy. The analysis on the features importance for each machine learning model are visually represented in the plots presented in Figure 1. The findings reveal that, for the majority of models, most significant features include mean values of the Mel coefficients 1, 2, 3, and 9. These features displayed the highest contribution to model performance. In contrast, Mel coefficients 6, 10, and 12 also demonstrated importance, albeit to a lesser extent. On the other hand, certain features had a minimal effect on the model's performance. Specifically, mean values of Mel coefficients 4, 5, 7, 8, and 11 were found to contribute little to the predictive capability of the models. This differentiation highlights the importance of feature selection in improving model efficiency and interpretability.

To illustrate the effectiveness of statistical analysis on Mel-frequency cepstral coefficients (MFCCs), Among the 12 MFCCs computed, coefficients 1, 2, 3, and 9 demonstrated the highest sensitivity contributions and were therefore selected for further statistical characterization. For each selected MFCC, four statistical descriptors—mean, variance, skewness, and kurtosis—were computed to capture both central tendency and distributional shape across the temporal domain of the audio clip. The results are summarized in table 3 as follows:

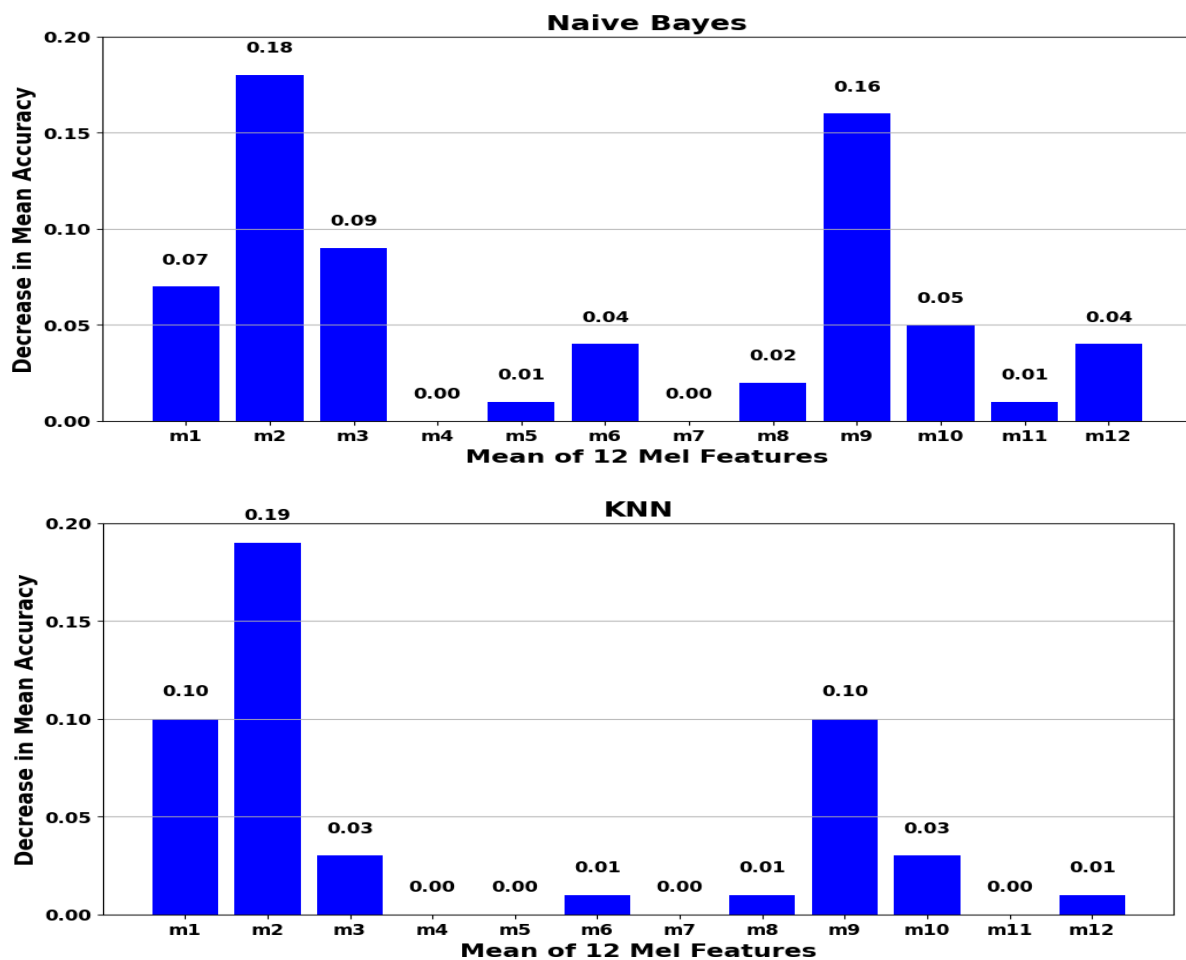
Table 3: Illustration of statistical analysis

MFCC Index	Mean	Variance	Skewness	Kurtosis
MFCC 1	-185.24	145.67	0.42	2.13
MFCC 2	34.82	120.33	-0.18	2.91
MFCC 3	12.46	89.10	0.75	3.58
MFCC 9	-5.89	15.42	-0.44	1.85

MFCC 1 exhibited a mean of -185.24, a variance of 145.67, skewness of 0.42, and kurtosis of 2.13, indicating a relatively stable low-frequency energy pattern. MFCC 2 had a mean of 34.82 and a slightly platykurtic distribution, while MFCC 3 showed a higher skewness of 0.75 and kurtosis of 3.58, reflecting the presence of spectral fluctuations potentially caused by transient acoustic events. MFCC 9, capturing higher-frequency details, showed lower overall magnitude and variation, with a mean of -5.89 and variance of 15.42. These statistics provide a compact yet informative representation of the spectral characteristics over time and are well-suited for use in classifier models. Incorporating such descriptors into the feature pipeline can enhance model interpretability and improve performance in tasks such as sound classification or acoustic event detection.

B. Performance Evaluation

This multiclass classification problem concerns three different outputs: queenright colony which are healthy, queenless colony in which their queens were pulled-out manually, and a less populated queenless colony. Additionally, a number of metrics such as accuracy, precision, recall, and F1-score, have been calculated to evaluate the prediction performance of the machine learning models and provide a comprehensive picture. True positive (TP), false positive (FP), true negative (TN), and false negative (FN) predictions are displayed on the Confusion Matrix (CM), which functions as a graphic that illustrates the accurate and inaccurate predictions produced by a classification model. A clear foundation for visualizing model performance is provided by the CM, from which various metrics can be obtained. The percentage of correctly classified samples relative to the total number of samples is known as accuracy. The precision measures the percentage of correctly categorized as positive samples that were predicted to be positive. Recall, commonly labeled sensitivity, analyzes the percentage of positive cases that are successfully recognized as positive. F1-score is computed as a mean of Precision and Recall. The K-fold cross validation approach is used to achieve validation.



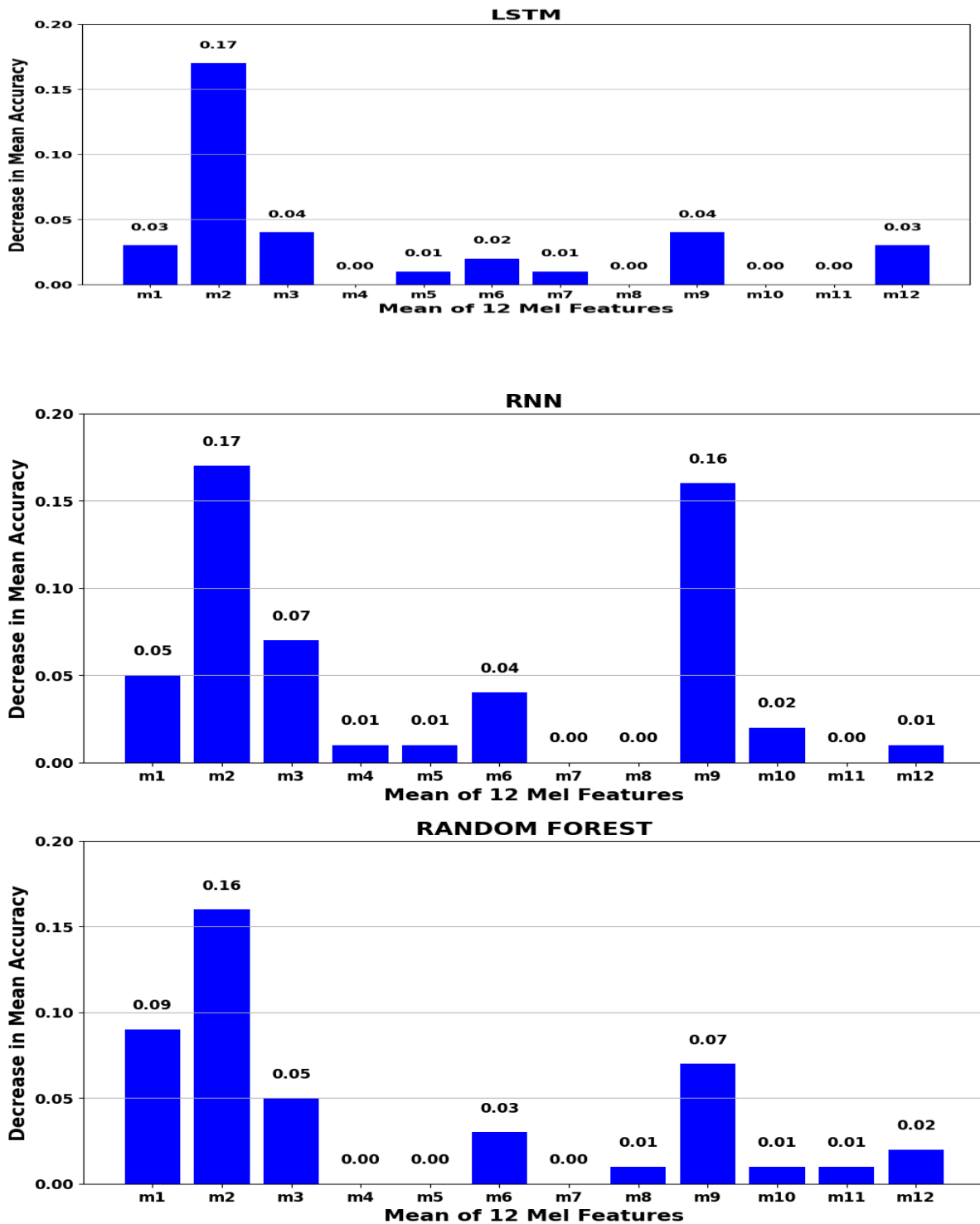


Figure 1: Sensitivity analysis to examine the feature importance for each ML model.

This method is based on splitting the data set into k different groups or subsets. All samples except the first subset which is used as a test set to assess the model performance, are used to train the model. After that, the withheld subset is added back in, and the process is repeated by taking out a second group. The final step usually involves combining all k estimations into a

single mean number. With 720 samples in the dataset, an eight-fold cross validation approach is used. During the eight-fold cross validation procedure, the CM and related performance metrics were calculated. There are 90 samples total in each fold, with 36 samples belonging to QR colonies, and 36 samples to QL colonies, and remaining 18 samples to LPQL colonies. There are 288 samples in total for QR colonies, and 288 samples in total for QL colonies, and remaining 144 samples in total for LPQL colonies in the eight-fold cross-validation. As shown in Figure 4, the outcomes from each fold are combined into a single confusion matrix for every machine learning model. The components on the main diagonal specifies the TP values that are relevant to each of their classes. In summary, the matrix's value distribution defines the TP, FN, FP, and TN values in the designated cells. Reallocating the proper TN, FN, and FP values for the other two classes such as QL and LPQL classes is necessary.

When the primary diagonals in the confusion matrices shown in Figure 2 are examined, it becomes clear that the KNN and RF models attained the highest rates of true classification, while the NB and RNN classifiers performed the worst, and the LSTM model showed intermediate values. Table 4 and table 5 displays the performance characteristics for each machine learning model with mean value of the eight-fold cross validation process for MFCC's and Gabor-MFCC's respectively. Table 6 and table 7 summarizes the performance metrics for GFCC's and MFCC spectral fusion method of feature extraction. In all the methods it is evident that RNN, LSTM, and NB models had the lowest values and the KNN and RF combinations produced the best results. The mean value of the measurements is displayed in Figure 3 to 6 to make it simple to compare the outcomes. The KNN and RF outcomes in this comparison are rather comparable; NB, RNN, and LSTM had the lowest performance. Nonetheless, the majority of the results are higher than 90% and are therefore thought to be suitable for differentiating the colonies' conditions. Figure 7 shows the accuracy's of different ML methods for different feature extraction processes.

	NB			KNN		
QR	267	11	10	285	3	0
QL	7	276	5	3	283	2
LPQL	11	9	124	0	1	143
	QR	QL	LPQL	QR	QL	LPQL

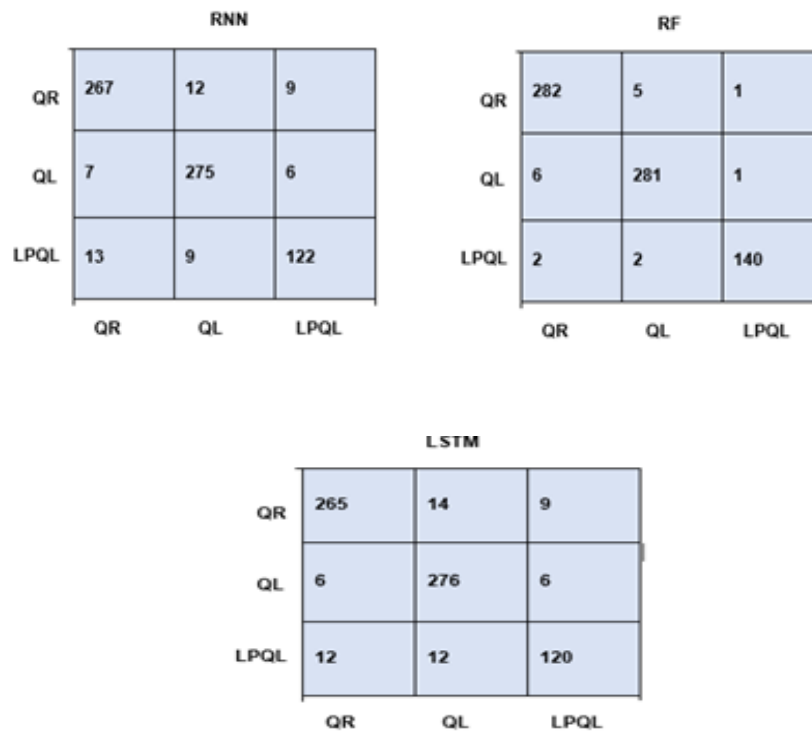


Figure 2: Summary of confusion matrix for eight-fold cross validation

Table 4: Overview of performance metrics for MFCC’s

Metric s	Accurac y	Precisio n	Recal l	F1 Score
NB	0.926	0.916	0.919	0.918
KNN	0.947	0.948	0.948	0.947
LSTM	0.918	0.917	0.916	0.919
RNN	0.922	0.922	0.925	0.930
RF	0.936	0.937	0.939	0.941

Table 5: Overview of performance metrics for Gabor-MFCC’s

Metric s	Accurac y	Precisio n	Recal l	F1 Score
NB	0.954	0.944	0.947	0.946
KNN	0.963	0.964	0.964	0.963
LSTM	0.946	0.945	0.944	0.947

RNN	0.950	0.950	0.953	0.958
RF	0.966	0.967	0.969	0.971

Table 6: Overview of performance metrics for GFCC's

Metric	Accuracy	Precision	Recall	F1 Score
NB	0.953	0.947	0.947	0.948
KNN	0.966	0.964	0.967	0.973
LSTM	0.946	0.946	0.948	0.947
RNN	0.951	0.950	0.952	0.958
RF	0.966	0.968	0.972	0.973

Table 7: Overview of performance metrics for MFCC Spectral fusion

Metric	Accuracy	Precision	Recall	F1 Score
NB	0.964	0.964	0.967	0.966
KNN	0.993	0.994	0.994	0.993
LSTM	0.966	0.965	0.964	0.967
RNN	0.970	0.970	0.973	0.978
RF	0.986	0.987	0.989	0.991

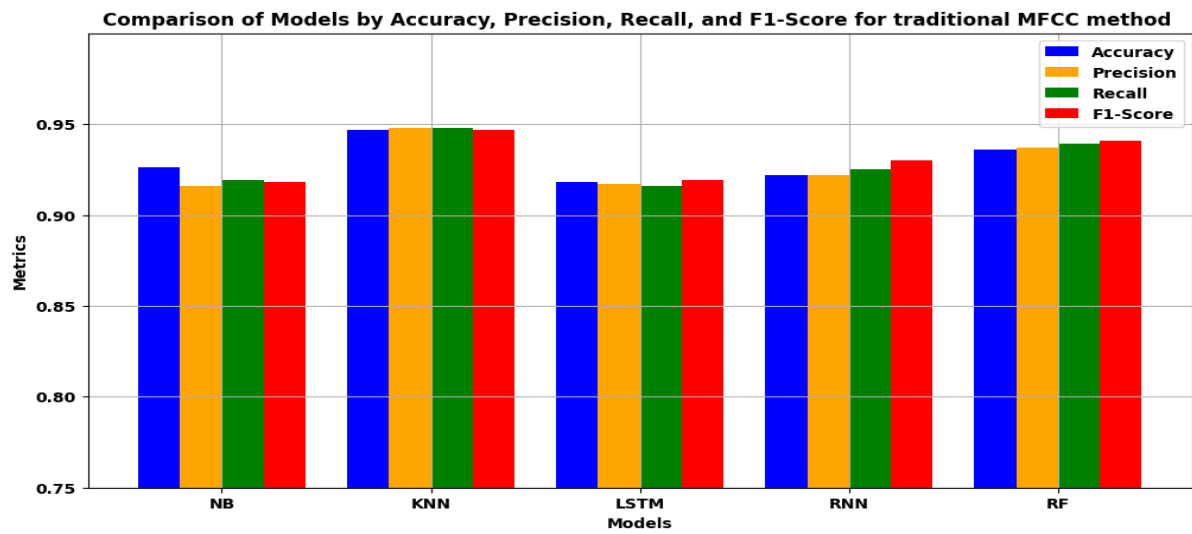


Figure 3: Comparison of performance metrics of ML models for traditional MFCC method

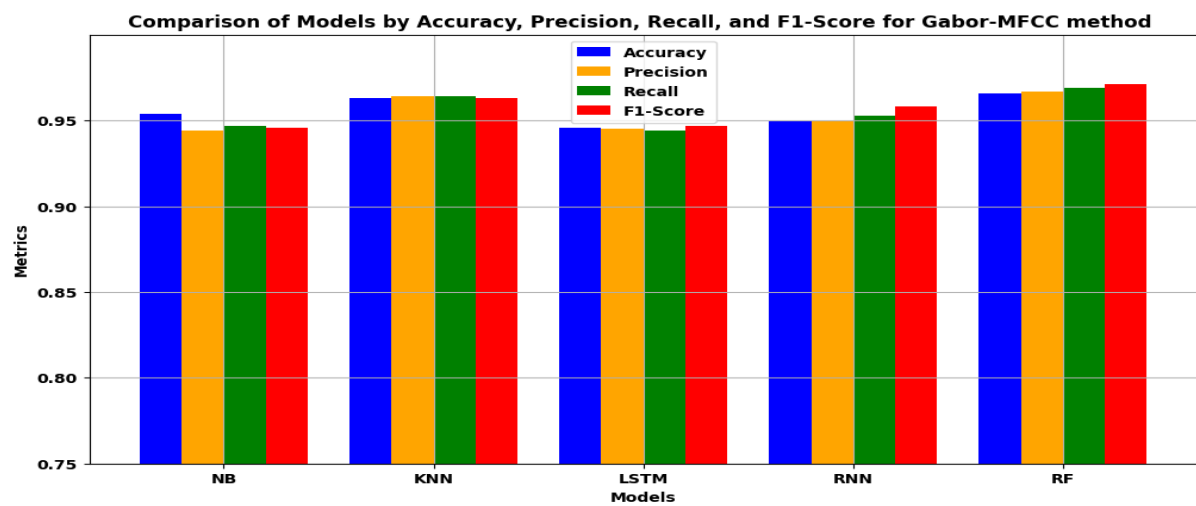


Figure 4: Comparison of performance metrics of ML models for Gabor-MFCC method

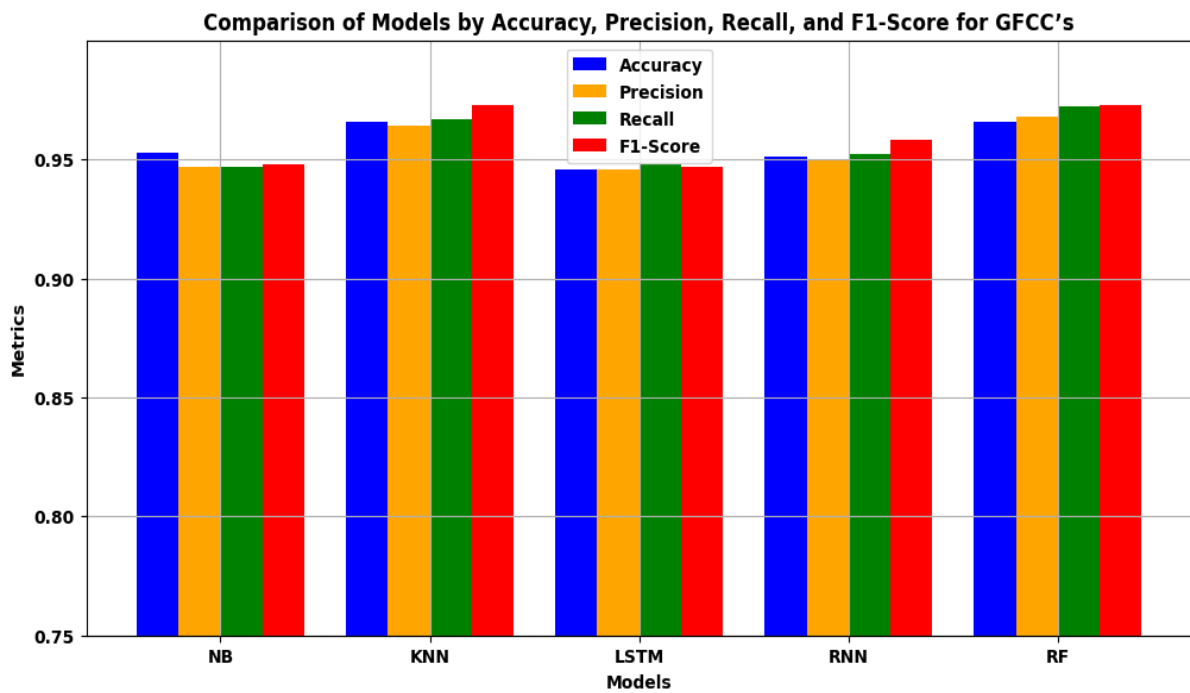


Figure 5: Comparison of performance metrics of ML models for GFCC method

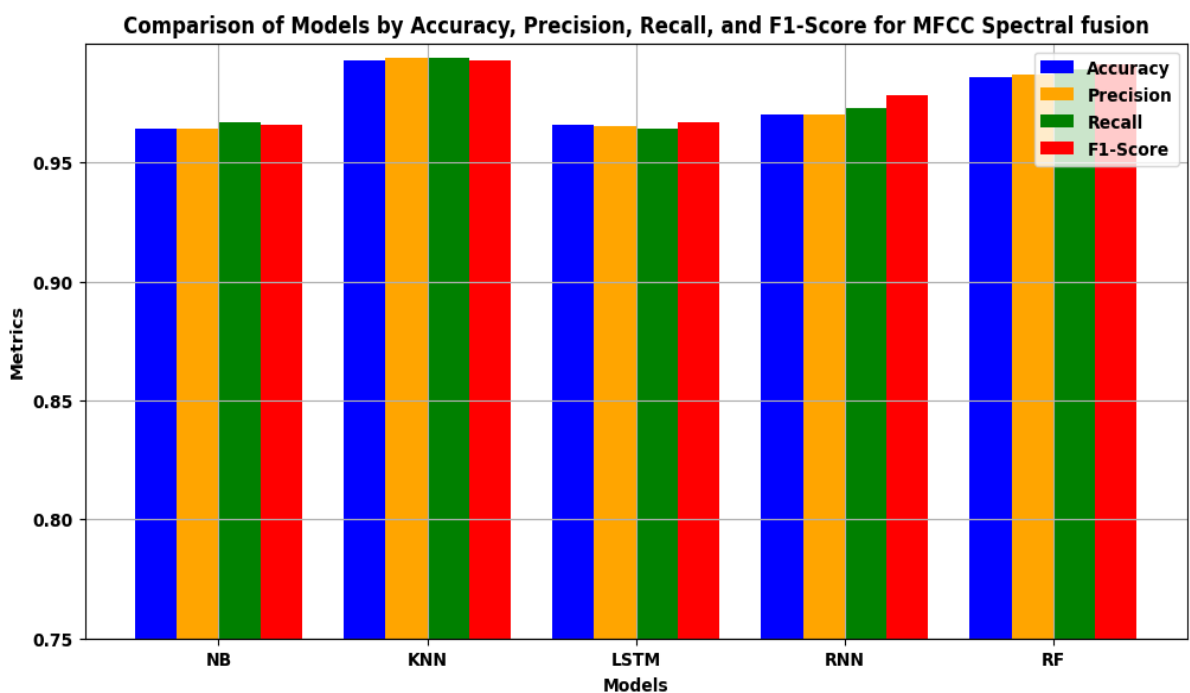


Figure 6: Comparison of performance metrics of ML models for MFCC spectral fusion method

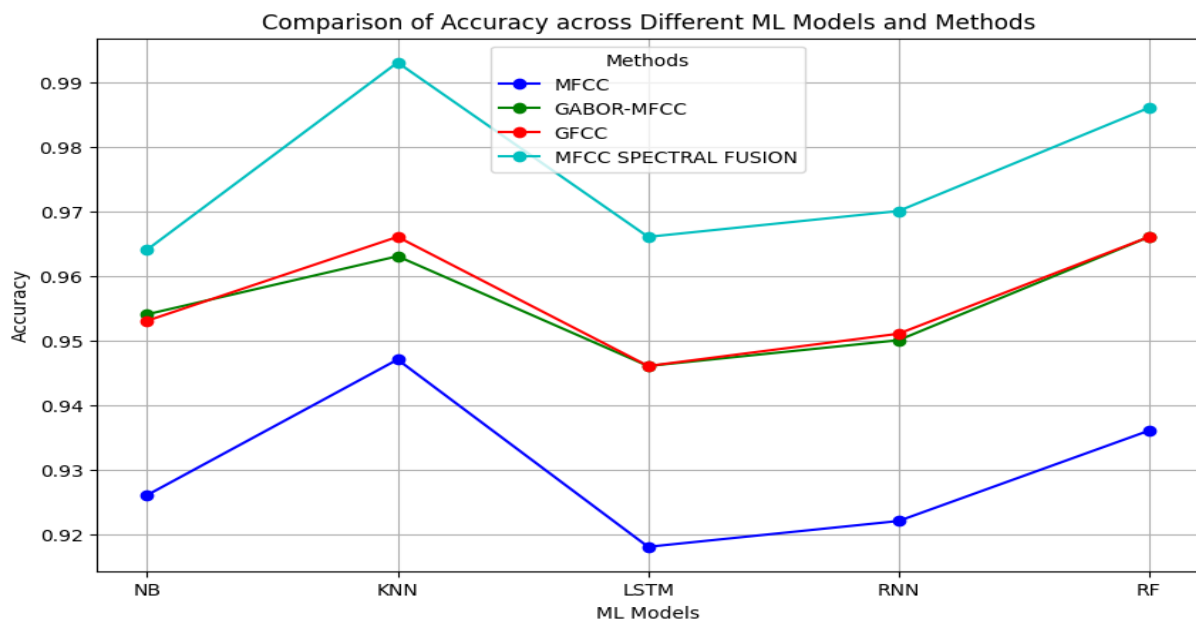


Figure 7: Comparison of Accuracy of ML models for different feature extraction methods

C. Discussions:

This study conducted an examination of the effectiveness of diverse machine learning algorithms in the classification of acoustic signals produced by bees. The assessment of these models yielded significant revelations regarding their efficacy in differentiating between the various states of bee colonies based on acoustic information. The results suggest that, despite its computational efficiency, Naive Bayes (NB) displayed an accuracy of 92.6%, which stemmed from its fundamental assumptions that limited its capacity to process intricate acoustic patterns. The K-Nearest Neighbors (KNN) framework revealed impressive effectiveness, realizing an accuracy of 94.7%, which is linked to its non-parametric nature. However, its effectiveness was contingent upon the selection of the distance metric and p-value utilized. Random Forest (RF) demonstrated commendable efficacy due to its ensemble methodology, proficiently addressing non-linear relationships and interactions inherent within the dataset. It accomplished an accuracy of 93.6%, representing a noteworthy equilibrium between precision and interpretability.

Remarkable results were achieved by Recurrent Neural Networks (RNNs), securing a 92.2% accuracy by skillfully identifying the temporal links present in the dataset; still, their potential was hindered by the difficulties associated with gradient vanishing and considerable computational overhead. With a notable accuracy of 91.8%, LSTM networks proved to be highly effective in modeling long-term dependencies. Their adeptness at managing the temporal characteristics of acoustic signals contributed to this classification performance. There is a noticeable trend where models like LSTM and RNN, which are generally known for their ability to capture temporal dependencies in data, did not perform as well as KNN and RF. This could indicate that the nature of the data may not require complex temporal modeling or that the features extracted via Traditional MFCCs are not optimal for these models.

Similar to the Traditional MFCCs, KNN again leads in performance with an accuracy of 96.3%, indicating that Gabor-MFCCs may provide richer feature representations that enhance KNN's effectiveness. The Random Forest model continues to perform well, achieving an accuracy of 96.6% and an F1 score of 97.1%, which reinforces the reliability of RF in handling complex datasets. The Naive Bayes model shows a significant improvement with Gabor-MFCCs, reaching an accuracy of 95.4%. This suggests that the Gabor features may provide a better representation of the data, making it more suitable for simpler models. The improvement in performance metrics across all models when using Gabor-MFCCs compared to Traditional MFCCs indicates that Gabor features may capture more relevant information from the data. This could be due to their ability to analyze frequency and spatial information more effectively, which is particularly beneficial in audio and speech processing tasks. Further when subjected to the Gammatone filterbanks, The KNN, RF and other models as well maintains their accuracy similar to gabor-mfcc method. Though the results are not significantly improved, GFCC can be considered over gabor-mfcc due to the modeling the cochlear response more faithfully than the mel filterbank. Finally when MFCC spectral features are applied over MFCC, the results are much more efficient and reliable, especially the KNN and RF achieving more 99% accuracy. Thus to enhance the robustness of the classification models this fusion method can be considered over traditional MFCC's in complex patterns also.

5. CONCLUSIONS

The overall trend shows that KNN and Random Forest consistently outperform other models in both feature extraction scenarios. This suggests that these algorithms are particularly well-suited for the classification tasks at hand, possibly due to their ability to handle high-dimensional data effectively. The choice between Traditional MFCCs, Fusion MFCC's, GFCC's and Gabor-MFCCs should consider the specific application and the computational resources available. While Gabor-MFCCs provide better performance, the increased complexity may not be necessary for all applications. For complex practices fusion method suits well where GFCC's are similar to Gabor-MFCCs. Overall, the findings underscore the effectiveness of Gabor-MFCCs and fusion MFCCs in enhancing model performance, particularly for KNN and RF, suggesting that feature extraction methods play a crucial role in the success of machine learning applications in this domain. Further research could explore hybrid approaches or alternative algorithms to improve the performance of the less effective models. Such investigations may lead to innovative solutions that leverage the strengths of various algorithms, ultimately contributing to more robust and accurate predictive models in complex datasets. This could involve integrating ensemble methods or incorporating additional feature selection techniques to refine the input data, thereby enhancing the learning capabilities of LSTM and RNN architectures. Exploring the potential of deep learning techniques in conjunction with these traditional methods could also yield significant improvements, as neural networks are adept at capturing intricate patterns that simpler models might overlook. The combination of these advanced methodologies may not only boost predictive accuracy but also provide deeper insights into the underlying structures within the data, paving the way for more informed decision-making across various applications. Implementing these strategies will require careful experimentation and validation to ensure that the chosen methods effectively

address the specific challenges posed by each dataset, ultimately leading to more reliable outcomes. This iterative process of refinement and adaptation will be crucial in harnessing the full potential of LSTM and RNN architectures, allowing researchers and practitioners to tailor their approaches based on empirical results and emerging trends in data science.

REFERENCES

- [1] Hoshiya, H., & Sasaki, M. (2008). Perspectives of multi-modal contribution of honeybee resources to our life. *Entomological Research*, 38. <https://doi.org/10.1111/J.1748-5967.2008.00170.X>.
- [2] A.-M. Klein *et al.*, “Importance of pollinators in changing landscapes for world crops,” in *Proc. Roy. Soc. B: Biol. Sci.*, 2007, vol. 274, no. 1608, pp. 303–313. [Online]. Available: <https://royalsocietypublishing.org/doi/abs/10.1098/rspb.2006.3721>
- [3] J. Fauconet *et al.*, “Honey bee winter mortality in France in 1999 and 000,” *Bee World*, vol. 83, pp. 14–23, 2002.
- [4] Oldroyd, “What’s killing american honey bees?,” *PLOS Biol.*, vol. 5, 2007, Art. no. e168.
- [5] Robles-Guerrero, A., Saucedo-Anaya, T., Guerrero-Mendez, C., Gómez Jiménez, S., & Navarro-Solis, D. J. (2023). Comparative Study of Machine Learning Models for Bee Colony Acoustic Pattern Classification on Low Computational Resources. *Sensors*, 23(1), 460. <https://doi.org/10.3390/s23010460>
- [6] Nolasco, A. Terenzi, S. Cecchi, S. Orcioni, H. L. Bear, and E. Benetos, “Audio-based identification of beehive states,” in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process.*, 2019, pp. 8256–8260.
- [7] Kim, J.; Oh, J.; Heo, T.Y. Acoustic Scene Classification and Visualization of Beehive Sounds Using Machine Learning Algorithms and Grad-CAM. *Math. Probl. Eng.* **2021**, 2021, 5594498.
- [8] Zgank, A. Acoustic monitoring and classification of bee swarm activity using MFCC feature extraction and HMM acoustic modeling. In *Proceedings of the 2018 ELEKTRO*, Mikulov, Czech Republic, 21–23 May 2018; pp. 1–4.
- [9] Zgank, A. Bee Swarm Activity Acoustic Classification for an IoT-Based Farm Service. *Sensors* **2019**, 20, 21.
- [10] Zgank, A. IoT-Based Bee Swarm Activity Acoustic Classification Using Deep Neural Networks. *Sensors* **2021**, 21, 676.
- [11] Dimitrios, K., Bellos, C. V., Stefanou, K., Stergios, G., Andrikos, I. O., Katsantas, T., & Kontogiannis, S. (2022). Performance Evaluation of Classification Algorithms to Detect Bee Swarming Events Using Sound. *Signals*, 3(4), 807–822. <https://doi.org/10.3390/signals3040048>
- [12] Cejrowski, T., Szymański, J., & Logofătu, D. (2020). Buzz-based recognition of the honeybee colony circadian rhythm. *Computers and Electronics in Agriculture*, 175, 105586. <https://doi.org/10.1016/J.COMPAG.2020.105586>
- [13] Kawakita, S., & Ichikawa, K. (2019). Automated classification of bees and hornet using acoustic analysis of their flight sounds. *Apidologie*, 50(1), 71–79. <https://doi.org/10.1007/S13592-018-0619-6>

- [14] M.-T. Ramsey *et al.*, “The prediction of swarming in honeybee colonies using vibrational spectra,” *Sci. Rep.*, vol. 10, no. 1, pp. 1–17, 2020.
- [15] V. Kulyukin, S. Mukherjee, and P. Amlathe, “Toward audio beehive monitoring: Deep learning vs. standard machine learning in classifying beehive audio samples,” *Appl. Sci.*, vol. 8, no. 9, pp. 1–33, 2018.
- [16] Zgank, “Bee swarm activity acoustic classification for an IoT-based farm service,” *Sensors*, vol. 20, no. 1, pp. 1–14, 2020.
- [17] Terenzi, A.; Cecchi, S.; Spinsante, S. On the Importance of the Sound Emitted by Honey Bee Hives. *Vet. Sci.* **2020**, 7, 168.
- [18] Abdollahi, M.; Giovenazzo, P.; Falk, T.H. Automated Beehive Acoustics Monitoring: A Comprehensive Review of the Literature and Recommendations for Future Work. *Appl. Sci.* **2022**, 12, 3920
- [19] Terenzi, A., Ortolani, N., Nolasco, I., Benetos, E., & Cecchi, S. (2022). Comparison of Feature Extraction Methods for Sound-Based Classification of Honey Bee Activity. *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, 30, 112–122. <https://doi.org/10.1109/taslp.2021.3133194>
- [20] Zlatkova, A.; Kokolanski, Z.; Tashkovski, D. Honeybees swarming detection approach by sound signal processing. In Proceedings of the 2020 XXIX International Scientific Conference Electronics (ET), Sozopol, Bulgaria, 16–18 September 2020; pp. 1–3.
- [21] Lima, M.V.; Queiroz, J.P.A.F.D.; Pascoal, L.A.F.; Saraiva, E.P.; Soares, K.O.; Evangelista-Rodrigues, A. Smartphone-based sound level meter application for monitoring thermal comfort of honeybees *Apis mellifera* L. *Biol. Rhythm. Res.* **2021**, 52, 1073–1086.
- [22] Tlac̃baba, J.; Černý, M.; Dostál, P.; Pridal, A. The acoustic emission in the nest of the honey bee depending on the extreme weather conditions. *Acta Univ. Agric. Silvic. Mendel. Brun.* **2014**, 62, 245–254
- [23] N. Huang, “Introduction to the Hilbert-Huang transform and its related mathematical problems,” in *Hilbert-Huang Transform Applications*, Singapore: World Scientific, 2005, pp. 1–26.
- [24] Howard, D.; Duran, O.; Hunter, G.; Stebel, K. Signal Processing the acoustics of honeybees (APIS MELLIFERA) to identify the “queenless” state in Hives. *Proc. Inst. Acoust.* **2013**, 35, 290–297.
- [25] Robles-Guerrero, A., Saucedo-Anaya, T., González-Ramírez, E., & De la Rosa-Vargas, J.I. (2019). Analysis of a multiclass classification problem by lasso logistic regression and singular value decomposition to identify sound patterns in queenless bee colonies. *Computers and Electronics in Agriculture*, 159, 69–74.
- [26] Cejrowski, T., & Szymański, J. (2021). Buzz-based honeybee colony fingerprint. *Computers and Electronics in Agriculture*, 191, Article 106489
- [27] Peng, R.; Ardekani, I.; Sharifzadeh, H. An Acoustic Signal Processing System for Identification of Queen-less Beehives. In Proceedings of the 2020 Asia-Pacific Signal and Information Processing Association Annual Summit and Conference (APSIPA ASC), Auckland, New Zealand, 7–10 December 2020; pp. 57–63.

- [28] Orłowska, A.; Fourer, D.; Gavini, J.P.; Cassou-Ribehart, D. Honey Bee Queen Presence Detection from Audio Field Recordings using Summarized Spectrogram and Convolutional Neural Networks. In *Lecture Notes in Networks and Systems, Proceedings of the 21st International Conference on Intelligent Systems Design and Applications (ISDA 2021)*, Online, 13–15 December 2021; Springer: Cham, Switzerland, 2021; Volume 418, pp. 83–92.
- [29] Cejrowski, T.; Szymański, J.; Mora, H.; Gil, D. Detection of the Bee Queen Presence Using Sound Analysis. In *Intelligent Information and Database Systems*; Springer International Publishing: Basel, Switzerland, 2018; pp. 297–306.
- [30] Kumar, P., & Champa, H. N. (2023). Efficient Disease Identification Method for Crop Leaf using Deep Learning Techniques. *International Journal on Recent and Innovation Trends in Computing and Communication*, 11(11s), 238–247. <https://doi.org/10.17762/ijritcc.v11i11s.8096>
- [31] L. Gong, S. Xie, Y. Zhang, Y. Xiong, X. Wang and J. Li, "A Robust Feature Extraction Method for Sound Signals Based on Gabor and MFCC," *2022 6th International Conference on Communication and Information Systems (ICCIS)*, Chongqing, China, 2022, pp. 49-55, doi:10.1109/ICCIS56375.2022.9998146.
- [32] Jia Xiao, and Sun Xiaolin, "Research on a Deep Learning Method for Speech Recognition," *IAENG International Journal of Computer Science*, vol. 51, no. 9, pp1272-1280, 2024
- [33] Pooi Shiang Tan, Kian Ming Lim, Cheah Heng Tan, and Chin Poo Lee, "Pre-trained DenseNet-121 with Multilayer Perceptron for Acoustic Event Classification," *IAENG International Journal of Computer Science*, vol. 50, no.1, pp51-62, 2023
- [34] Zhongyang Rao, and Chunyuan Feng, "Automatic Identification of Chords in Noisy Music Using Temporal Correlation Support Vector Machine," *IAENG International Journal of Computer Science*, vol. 50, no.2, pp813-819, 2023