

# STUDY ON WARNING OF QUEENLESS BEE HIVE BASED ON SOUND CLASSIFICATION

Khoi Nguyen Dao, Hoang Le Nguyen, Anh Tuan Hoang Nguyen, Son Trung Nguyen

Hanoi University of Science and Technology, Vietnam

## Abstract

The beekeeping industry in Vietnam is thriving, contributing significantly to the agricultural economy through products like honey, pollen, and beeswax. However, the current manual management methods face many challenges in timely detecting abnormalities, especially the loss of the queen bee - an event that can cause significant disruption to the colony, reducing productivity and increasing the risk of colony collapse. Applying modern technologies such as Artificial Intelligence (AI) is crucial for automating data collection and analysis on the bee colony's environment and health. Our research team has applied an AI model capable of using the results of sound feature extraction from bee colonies using MFCCs and making anomaly predictions using machine learning models such as CNN, LSTM+CNN. This technological solution not only helps reduce risks but also improves productivity and product quality, while reducing the manual inspection burden for beekeepers. The application of AI in beekeeping will bring greater accuracy, efficiency, and better management capabilities to the entire industry.

**Keywords:** sound classification, MFCCs, CNN, LSTM+CNN, AI in Agriculture.

## 1 Introduction

In beekeeping, the queen bee plays a vital role in maintaining the stability and development of the bee colony. The queen is the reproductive center of the colony, and she also regulates the activities and social structure within the hive [1]. The loss of the queen bee often leads to serious consequences such as reduced reproductive capacity, an imbalance in colony activities, and can even lead to the complete collapse of the bee colony. Therefore, early detection and warning of abnormal queen bee loss in beehives is crucial to ensure production efficiency and maintain the health of the bee colony.

If the loss of the queen bee in the colony is not detected promptly, it can cause serious problems such as an imbalance in the activities of the hive, leading to disorientation and chaos, resulting in a decline in the bee population, leading to the weakening and eventual death of the entire bee colony [5]. Therefore, the need for a device to provide early warning of queen loss is very important to help beekeepers intervene promptly. A warning system can quickly detect the loss of the queen, allowing beekeepers to intervene by replacing the queen with a new one or stimulating the colony to create a new queen before the colony weakens, thus reducing the damage caused by the loss of the queen.

## 2 Related Work

The development of Artificial Intelligence (AI) is opening up new opportunities for the beekeeping industry. AI technology allows for efficient remote monitoring, collecting continuous data on various aspects of bee colony life. AI algorithms are applied to analyze and process data, providing accurate information about the health and activity of bee colonies. This enables beekeepers to monitor their colonies 24/7 without direct intervention, detecting and addressing problems such as diseases or other harmful factors in a timely manner. Consequently, they can optimize beekeeping, improve product yield and quality, and save time and costs in colony management. Although this technology has been widely researched and applied in many countries worldwide, the application of AI in beekeeping in Vietnam is still limited. Research related to the detection of anomalies in beekeeping can be mentioned as follows:

Several systems introduced by Schurischuster et al. (2016) [12], Zacepins, Kvišis, and Stalidzans (2016) [13], Antonio et al. (2017) [2], Crawford (2017) [4] have used a multi-sensor beehive monitoring system called BeePi, including a Raspberry Pi computer, a miniature camera, 4 microphones connected to a splitter, a solar panel, a temperature sensor, a battery, and a clock. In the research of Kulyukin, Mukherjee, and Amlathe [7], microphones were placed to collect sound samples of bees, crickets, and ambient noise. The authors then used machine learning models on bee sound datasets collected from different locations to train and classify the collected sounds. The experimental results achieved high accuracy, so it is entirely possible to use sound to monitor the hive status.

Besides the above research on bee monitoring, in 2019, research by Ruvinga and colleagues used the MFCC feature extraction method along with a CNN network to predict queen bee loss sounds with an accuracy rate of up to 99% on the Arnia Ltd. dataset ([www.arnia.co.uk](http://www.arnia.co.uk)) [10], which shows that beehive sound analysis technology has been used as an effective tool for early detection of problems related to the queen bee, especially the loss of the queen. One of the reasons why sound analysis is an effective tool for early detection of problems related to the queen bee is that the transformation and change of sound are obvious when the queen bee of the bee colony has a problem, for example: When the queen bee dies or leaves, the hive sound can change from a quiet and rhythmic state to abnormal sounds such as long and repetitive buzzing of worker bees. This characteristic sound is a clear sign of instability in the hive. Compared to other inspection methods such as temperature or humidity sensors, although temperature or humidity sensors can detect changes

in the bee colony environment, they cannot provide specific information about the status of the queen bee or the bee colony. Sound analysis will have clearer data on queen loss based on the activity and behavior of worker bees.

Based on the analysis of research on detection monitoring in the beekeeping process, and then providing analysis to warn of queen loss, the research team chooses to analyze the sound collected from the beehive to issue an early warning of queen loss for deployment. This method provides high accuracy and is easy to automate, helping beekeepers monitor their bee colonies continuously and reduce risks. Compared to other monitoring methods, sound analysis not only detects early queen loss but also helps protect the health of the bee colony without direct intervention in the hive.

In essence, this text discusses the application of AI, particularly sound analysis, in beekeeping to detect problems like queen bee loss. It highlights the advantages of using AI, such as remote monitoring, early detection, and improved efficiency. The text also provides examples of existing research and the benefits of using sound analysis over other methods.

### 3 Theoretical Foundation for Anomaly Detection in Sound Classification

According to the chosen problem-solving approach, our research team will extract sound features and make predictions following the process shown in the image:

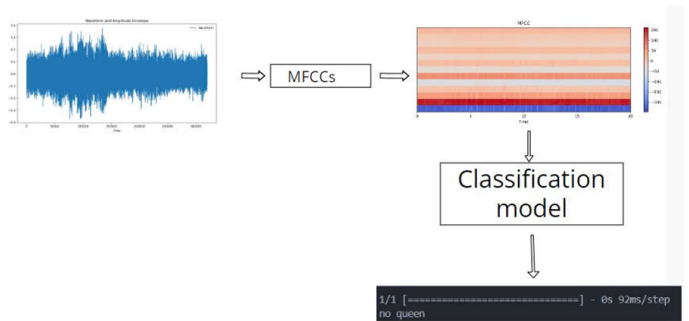


Figure 1: Audio Classification Process

In this process, the audio file will go through the MFCC feature extraction step to transform from the time domain to the frequency domain. Afterward, the classification model will output the result indicating whether the audio file contains a queen bee or not. The two classification models we used are the Convolutional Neural Network (CNN) and the Long Short-Term Memory (LSTM) combined with CNN (LSTM+CNN).

### 4 Feature Extraction in Audio

To extract audio features, we'll employ the MFCC (Mel-Frequency Cepstral Coefficients) technique and spectrogram for audio visualization. The feature extraction process comprises steps such as pre-emphasis, windowing, DFT (Discrete Fourier Transform), and more [6] as illustrated below:

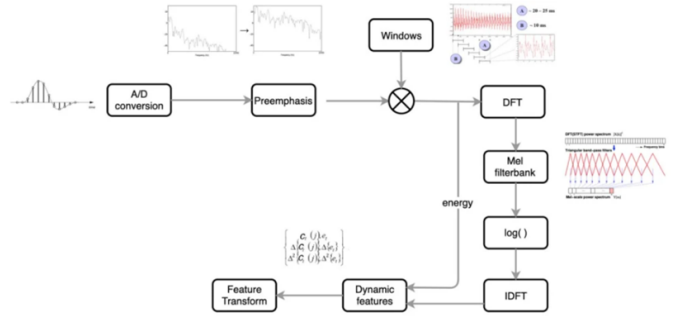


Figure 2: Steps to extract audio features

Among the steps outlined in Figure 2, the Discrete Fourier Transform (DFT) is a crucial step in MFCC. It facilitates the transformation of the input signal from the time domain to independent frequencies in the frequency domain, enabling visualization through a spectrogram. This makes subsequent machine learning steps much easier.

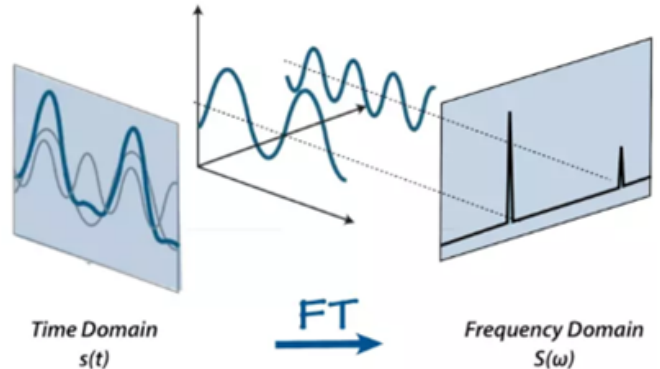


Figure 3: Audio is converted from the time domain to the frequency domain using the Fourier transform. [3]

Following the Fourier transform, we obtain an amplitude-frequency plot. However, we lose temporal information, meaning the system doesn't know the sequence in which the sounds occurred. Here, we need a plot that displays both frequency and amplitude over time. A spectrogram, as shown in Figure 4, can achieve this. The spectrogram's x-axis represents time, the y-axis represents frequency, and the color intensity represents the amplitude of the sound. Brighter colors indicate stronger frequencies.

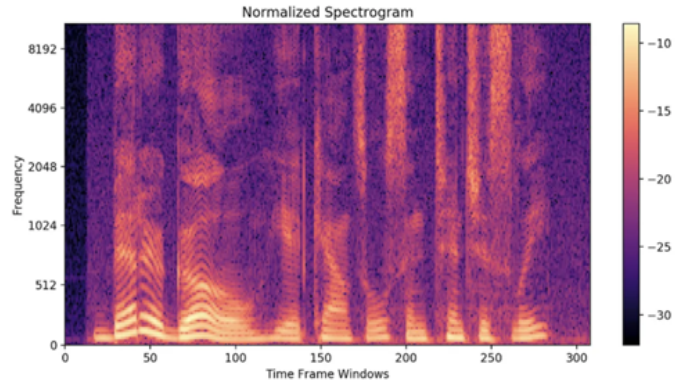


Figure 4: Spectrogram after applying Fourier transform

## 5 Proposed Machine Learning Models

Our research team proposes using two popular architectures: Convolutional Neural Networks (CNNs) and a combination of Long Short-Term Memory networks with CNNs. Each model has its own advantages and disadvantages that need to be considered.

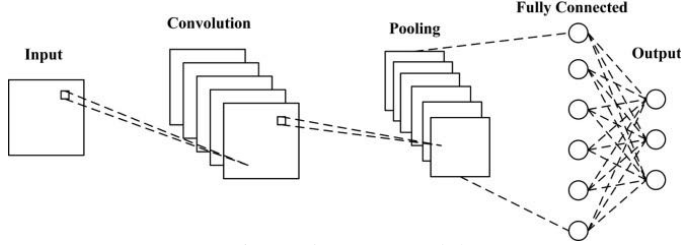


Figure 5: CNN Model

Convolutional Neural Networks (CNNs) (Figure 5) are a powerful tool for extracting spatial features from audio data. When audio is converted into spectrograms, which visually represent sound, CNNs can identify local patterns like edges and textures in the data [9]. One of the biggest advantages of CNNs is their translational invariance, which makes the model resistant to shifts and distortions in the input data. Moreover, CNNs significantly reduce the number of parameters by sharing weights, making the training process more efficient. This model is also easily scalable by adding more layers and filters, allowing for the extraction of increasingly complex features.

Research group proposed a CNN model begins with an input layer designed to accept an input image. It first applies a 2D convolutional layer with 32 filters, a 3x3 kernel size, and ReLU activation, followed by a max pooling layer with a 2x2 pool size and strides of 2x2 to downsample the feature maps. The process is repeated with a second convolutional layer, this time with 64 filters and the same kernel size and activation, followed by another max pooling layer with identical parameters. The feature maps are then flattened into a 1D vector, which is passed through a series of fully connected (dense) layers. The first dense layer has 64 units with ReLU activation, followed by a second dense layer with 250 units and ReLU activation, and a third dense layer with 100 units and ReLU activation. This model concludes with an output dense layer containing a single unit with a sigmoid activation function for binary classification tasks.

However, CNNs have some limitations. The model primarily captures spatial features and may not be effective in capturing the temporal dependencies inherent in sequential audio data [8]. CNNs also require a fixed input size, which can be a limitation when processing audio segments of varying lengths. Additionally, audio data needs to be converted into spectrograms or other image-like representations before being fed into a CNN, requiring an additional preprocessing step.

Research group proposed a CNN model begins with an input layer designed to accept an input image. It first applies a 2D convolutional layer with 32 filters, a 3x3 kernel size, and ReLU activation, followed by a max pooling layer with a 2x2 pool

size and strides of 2x2 to downsample the feature maps. The process is repeated with a second convolutional layer, this time with 64 filters and the same kernel size and activation, followed by another max pooling layer with identical parameters. The feature maps are then flattened into a 1D vector, which is passed through a series of fully connected (dense) layers. The first dense layer has 64 units with ReLU activation, followed by a second dense layer with 250 units and ReLU activation, and a third dense layer with 100 units and ReLU activation. This model concludes with an output dense layer containing a single unit with a sigmoid activation function for binary classification tasks.

In summary, we have outlined the advantages and disadvantages of our models, CNN and LSTM+CNN, based on the theoretical foundation. The evaluation of the performance of the two models, based on accuracy and computational speed, will be conducted through practical experiments on the bee audio dataset collected through IoT devices installed by the group at the bee farm.

## 6 Experimental Results

In this experiment, we conducted tests using the MFCC (Mel-frequency cepstral coefficients) feature extraction method combined with two neural network models: CNN (Convolutional Neural Network) and LSTM+CNN (Long Short-Term Memory + Convolutional Neural Network) to identify bee sounds. The test dataset consisted of 500 untrained audio files collected using IoT devices installed on *Apis mellifera* beehives at the Bee Center of the Vietnam National University of Agriculture, including 405 audio files of bee colonies with queens and 110 files without queen bees. The experimental results showed that the CNN model achieved an accuracy of 96.29% (390/405) with a prediction time of 50.97 seconds, while the LSTM+CNN model only achieved 80.24% (325/405) accuracy and took 123.78 seconds to process. The results of the two models are presented in detail in Table 1:

Model	Accuracy	Processing time
CNN	96.29%	50.97 seconds
LSTM+CNN	80.24%	123.78 seconds

Table 1: Comparison of accuracy and processing time between models

Analysis of the results shows that the CNN model not only outperforms in terms of accuracy but also in processing speed, which is nearly 2.5 times faster than LSTM+CNN. The research results also confirm the effectiveness of combining the MFCC feature extraction method with the CNN model in the problem of anomaly detection through bee sound classification. Although the LSTM+CNN model has a more complex structure, in this particular case, the model does not yield higher performance compared to CNN.

From the experimental results, we can conclude that the combination of the MFCC feature extraction method and the CNN network is an effective method for identifying bee sounds, ensuring not only high accuracy but also fast processing speed, making it a good choice for practical applications in this field.

## 7 Conclusion and Recommendations

In this paper, we concluded that the AI sound classification system to support early warning of queen loss will achieve an accuracy of more than 16% higher and a speed 2.5 times faster when using the MFCC feature extraction technique with the CNN machine learning model compared to LSTM+CNN. The system uses sound from beehives, training AI to predict anomalies with MFCCs (to represent sound as images) and CNN (to extract important features from MFCCs), thereby optimizing bee care, improving productivity and product quality, and saving time and costs in bee management. The research group aims to continue researching, designing devices and developing equipment to analyze and evaluate more abnormal cases in beehives to warn beekeepers such as: warning of swarm division, insect attacks, diseases, etc.

## 8 Appendix

Combining CNNs and LSTMs (Figure 6) offers several benefits. The LSTM+CNN model leverages the strengths of both models, capturing both spatial features (through CNN) and temporal dependencies (through LSTM) [11]. This makes them suitable for audio classification where temporal patterns are crucial. This combination often leads to better performance in tasks that require understanding both local features and global sequence patterns, such as speech recognition and music genre classification. Moreover, LSTMs can handle variable-length sequences [8], making the combined model more flexible in processing audio segments of different lengths.

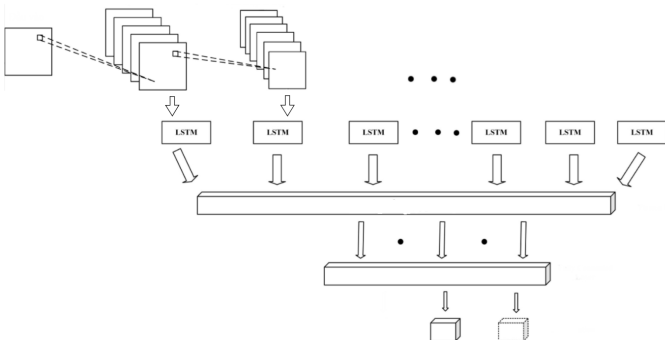


Figure 6: LSTM+CNN Model

However, combining CNNs and LSTMs results in more complex models with more parameters, requiring more computational resources, and training LSTM+CNN models takes longer than training a standalone CNN.

## References

- [1] K. Anton and C. Grozinger. *An introduction to queen honey bee development*. <https://extension.psu.edu/an-introduction-to-queen-honey-bee-development>. Penn State Extension. Dec. 2022.
- [2] R. G. Antonio et al. "Frequency analysis of honey bee buzz for automatic recognition of health status: A preliminary study". In: *Research in Computing Science* (2017).

- [3] D. Biswas. *Fourier transform*. <https://dibyendu-biswas.medium.com/fourier-transform-70aeb7ec252>. June 2021.
- [4] M. Crawford. "Automated collection of honey bee hive data using the Raspberry Pi". Doctoral dissertation. Appalachian State University, 2017.
- [5] "Honey bee colony losses: Why are honey bees disappearing?" In: *Sociobiology* 68.1 (2021), pp. 1–13. DOI: 10.13102/sociobiology.v68i1.5851.
- [6] M. A. Hossan, S. Memon, and M. A. Gregory. "A novel approach for MFCC feature extraction". In: *2010 4th International Conference on Signal Processing and Communication Systems (ICSPCS)*. IEEE, 2011. DOI: 10.1109/ICSPCS.2010.5709752.
- [7] V. Kulyukin, S. Mukherjee, and P. Amlathe. "Toward audio beehive monitoring: Deep learning vs. standard machine learning in classifying beehive audio samples". In: *Applied Sciences* 8 (2018), p. 1573.
- [8] B. Lim and S. Zohren. "Time series forecasting with deep learning: A survey". In: *Department of Engineering Science, University of Oxford* (2020). URL: <https://arxiv.org/abs/2004.14055>.
- [9] K. O'Shea and R. Nash. "An introduction to convolutional neural networks (Version 2)". In: *arXiv* (2015). URL: <https://arxiv.org/abs/1511.06422>.
- [10] S. Ruvinga et al. "Identifying queenlessness in honeybee hives from audio signals using machine learning". In: (Mar. 2023). Published. URL: [www.arnia.co.uk](http://www.arnia.co.uk).
- [11] T. N. Sainath et al. "Convolutional, long short-term memory, fully connected deep neural networks". In: *Proceedings of the 16th Annual Conference of the International Speech Communication Association, INTERSPEECH 2015*. 2015, pp. 1–5. URL: [https://www.isca-speech.org/archive/interspeech\\_2015/saintath15\\_interspeech.html](https://www.isca-speech.org/archive/interspeech_2015/saintath15_interspeech.html).
- [12] S. Schurischuster et al. "Sensor study for monitoring varroa mites on honey bees (*apis mellifera*)". In: *23rd International Conference on Pattern Recognition (ICPR 2016)*. Cancun, Mexico, Dec. 2016.
- [13] A. Zacepins, A. Kviesis, and E. Stalidzans. "Remote detection of the swarming of honey bee colonies by single-point temperature monitoring". In: *Biosystems Engineering* 148 (2016), pp. 76–80.