

Development of Wingbeat-Based Acoustic Health Monitoring System for Bee Colonies [†]

Li-Hao Chen ¹, Shi-You Zhou ¹, Jia-Wen He ¹ and Chau-Chung Song ^{2,*}

¹ Department of Aeronautical Engineering, National Formosa University, Yunlin 632, Taiwan; 11378104@nfu.edu.tw (L.-H.C.); 11378122@nfu.edu.tw (S.-Y.Z.); 11378159@nfu.edu.tw (J.-W.H.)

² Department of Aeronautical Engineering and Smart Machine and Intelligent Manufacturing Research Center, National Formosa University, Yunlin 632, Taiwan

* Correspondence: ccsong@nfu.edu.tw; Tel.: +886-5631-5547

[†] Presented at 8th International Conference on Knowledge Innovation and Invention 2025 (ICKII 2025), Fukuoka, Japan, 22–24 August 2025.

Abstract

We developed an intelligent acoustic health monitoring system for honeybee colonies based on wingbeat frequency analysis, offering a practical solution for modernizing apicultural practices. The system employs a three-layer architecture—the Internet of Things, fog, and cloud—to achieve real-time, non-invasive hive condition assessment. At the edge level, a Raspberry Pi and low-noise microphone continuously capture in-hive audio, which is converted into spectrograms using short-time Fourier transform (STFT). These are analyzed by a deep learning classification model deployed on the fog layer to distinguish four critical queen-related states: original queen present, queen absent, new queen rejected, and new queen accepted. The cloud layer supports data storage, visualization, and model refinement through manual annotations. Our results show that both the vision Transformer and CNN models perform effectively in classifying complex hive states, each contributing to the overall classification task, demonstrating the system’s potential for improving colony management and early intervention. This work contributes to precision apiculture by enabling scalable, real-time queen status monitoring through acoustic sensing and deep learning.

Keywords: health monitoring system; wingbeat-based acoustic; IoT layer; queen bee; short-time Fourier transform; transformer-based classification model; convolutional neural networks



Academic Editors: Teen-Hang Meen, Chun-Yen Chang and Cheng-Fu Yang

Published: 16 March 2026

Citation: Chen, L.-H.; Zhou, S.-Y.; He, J.-W.; Song, C.-C. Development of Wingbeat-Based Acoustic Health Monitoring System for Bee Colonies [†]. *Eng. Proc.* **2025**, *120*, 73. <https://doi.org/10.3390/engproc2025120073>

Copyright: © 2026 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Honeybees (*Apis mellifera*) are indispensable pollinators that significantly contribute to global food production and biodiversity. More than 75% of the world’s food crops depend, at least in part, on pollination by insects, with honeybees playing a dominant role in this ecosystem service [1]. In addition to producing honey, beeswax, and royal jelly, honeybee colonies are considered sensitive biosensors for environmental changes. However, colony health is increasingly threatened by environmental stressors, including pesticides, climate change, parasites, and poor queen management [2]. In particular, queen loss and swarming behaviors often precede colony collapse and can severely impact hive productivity and survival.

Traditional hive monitoring methods rely on visual inspections or manual observations by beekeepers, which are labor-intensive, imprecise, and often fail to detect early signs of

behavioral abnormalities [3]. Recently, acoustic monitoring has emerged as a promising non-invasive technique, as bees produce characteristic sounds associated with wing movement, vibration, and hive state changes [4,5]. Prior studies have explored the use of audio features such as Mel spectrograms (MFCCs), combined with machine learning classifiers, to distinguish simple hive conditions (e.g., queen presence vs. absence or swarming detection) [6–9]. However, these systems are mostly limited to binary classification and lack the granularity needed to reflect subtle but critical transitions in queen status.

To address this gap, we propose a practical and deployable acoustic-based monitoring system that classifies four distinct hive states associated with queen dynamics. The system adopts a three-tier architecture: the IoT layer for acoustic sensing using a Raspberry Pi and microphone, the fog layer for real-time audio processing and classification using a deep learning model, and the cloud layer for data visualization and long-term analytics. While previous methods only detect whether the queen is present, the system developed distinguishes among four critical statuses: (1) original queen present, (2) queen not present, (3) rejected new queen, and (4) accepted new queen. To support real-world deployment, the system is implemented on a Raspberry Pi 4B platform, enabling local audio recording and real-time classification directly within the hive environment.

The developed system transforms raw in-hive audio into spectrograms using short-time Fourier transform (STFT) and leverages vision Transformers (ViT) to perform robust classification under complex acoustic conditions. Both CNN- and ViT-based approaches contribute to effective classification of fine-grained queen-related states, each model excelling in different aspects of the classification task. This work not only advances the technical frontier of smart apicultural monitoring but also provides a foundation for real-time decision support in sustainable beekeeping practices.

2. Methodology

2.1. Hardware

To achieve continuous and non-invasive monitoring of bee colony health, the developed system integrates multiple sensing modules and embedded devices, as illustrated in Figure 1.

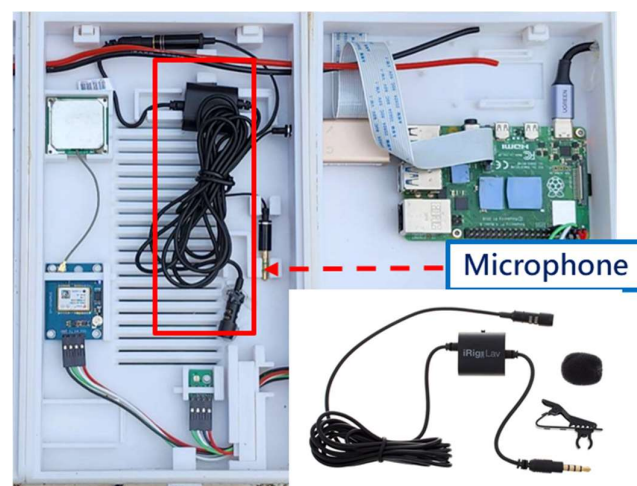


Figure 1. Smart beehive hardware equipment.

The hardware architecture contains a Raspberry Pi 4B, which functions as the local edge computing unit responsible for data acquisition, preprocessing, and inference tasks within the hive. An iRig Mic Lav omnidirectional microphone is embedded inside the beehive to capture wingbeat-generated acoustic signals. These audio recordings reflect

various behavioral and physiological states of the colony, particularly those associated with the queen bee. The microphone is strategically positioned to minimize environmental noise and ensure high-fidelity signal acquisition. Additionally, an infrared (IR) camera is installed to visually monitor bee activity, enabling real-time observation even under low-light conditions.

To support environmental sensing, the system includes an AHT20 temperature and humidity sensor and an HX711-based load cell module. These sensors allow the estimation of internal hive conditions and colony population dynamics. The data collected not only informs health diagnostics but also drives a temperature control subsystem that automatically activates a heater or cooling fan to maintain an optimal environment for the colony. All sensor data is transmitted via Wi-Fi to a cloud database, enabling real-time visualization through a web-based human-machine interface (HMI). The artificial intelligence (AI) classification model, initially trained on an NVIDIA RTX 4060 GPU using labeled audio spectrograms, is deployed back onto the Raspberry Pi for local inference. This allows beekeepers to receive immediate feedback on queen status and act accordingly.

As shown in Figure 1, the complete hardware setup consists of the following.

- Raspberry Pi 4B with microSD and Wi-Fi module
- iRig Mic Lav microphone
- IR camera
- AHT20 sensor
- HX711 load cell module
- Environmental control components (fan/heater)

These components form an integrated smart beehive platform capable of real-time data collection, local analysis, and remote monitoring—supporting proactive and data-driven apiculture management.

2.2. System Architecture

The developed hive monitoring system is structured into three layers: the IoT layer for acoustic sensing, the fog layer for local processing and classification, and the cloud layer for data storage and user interaction (Figure 2). This layered architecture enables distributed processing while ensuring real-time responsiveness and scalability in field deployment. In the IoT layer, a Raspberry Pi device connected to a low-noise omnidirectional microphone is installed inside a beehive and continuously records internal sounds in WAV format at a sampling rate of 22,050 Hz. The device segments recordings into 10 min files, which are locally stored and served via a lightweight web interface, enabling real-time, non-invasive audio acquisition directly from the hive environment without disturbing bee behavior. The fog layer, also deployed on the Raspberry Pi, performs local audio preprocessing and classification. Recorded WAV files are segmented and transformed into spectrograms using STFT (Figure 3), which visually represent the frequency distribution of wingbeat sounds over time and serve as input to the classification model.

We implemented two deep learning architectures: CNN and ViT, both of which were used for classifying the hive states. The CNN model consists of a standard stack of convolutional and pooling layers followed by fully connected layers optimized for image-based classification, while the ViT model is initialized from ImageNet-pretrained weights and fine-tuned using spectrogram images. Both models were trained using the same set of hyperparameters and input data for a fair comparison, specifically employing a batch size of 32, a learning rate of 2×10^{-5} , the AdamW optimizer, and a training duration of 150 epochs. The classification task was formulated to predict one of four queen-related hive states. All training and evaluation processes were implemented in PyTorch and executed either on a graphics processing unit (GPU) workstation or through edge deployment

testing on the Raspberry Pi. The cloud layer supports long-term data management and user interaction, where audio recordings, classification results, and model outputs are periodically synchronized from the fog layer to a central server. A web-based dashboard allows beekeepers to visualize historical hive conditions, monitor system performance, and optionally annotate samples for model retraining. The system supports manual review and future model refinement using updated labeled datasets, enabling continual learning and adaptation to different hive environments.

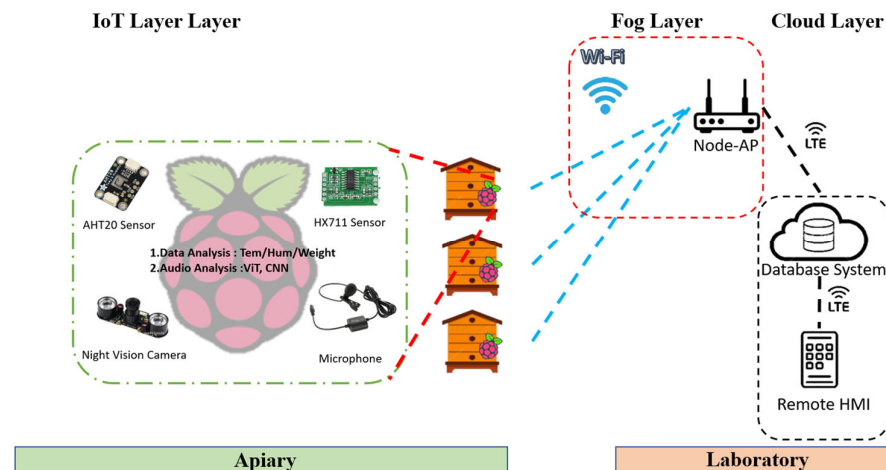


Figure 2. Beehive system architecture.

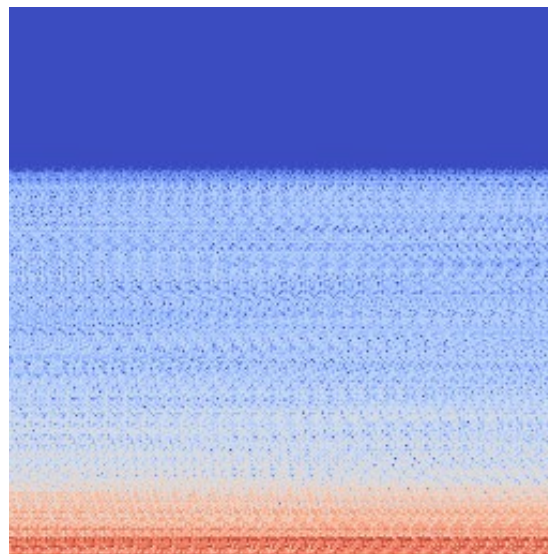


Figure 3. STFT spectrogram of beehive sounds, showing frequency over time. The color gradient represents the magnitude of the acoustic energy, with warmer colors (e.g., orange/red) indicating higher intensity and cooler colors (e.g., blue) indicating lower intensity.

2.3. Dataset

Sound analysis research in beehive monitoring has often been constrained by limited access to annotated audio datasets. The Beehive Sounds dataset, published on Kaggle [5], provides a valuable resource containing audio clips recorded from functioning hives under various environmental and behavioral conditions. The dataset includes annotations specifying one of four queen-related hive states: (1) original queen present, (2) queen not present, (3) rejected new queen, and (4) accepted new queen. This resource facilitates the development of machine learning models for detecting queen presence and enables finer-grained analysis of transitions between queen-related states. Identifying such transitions

is essential because failing to recognize when a newly introduced queen is missing or rejected can cause the colony to become unstable, trigger aggressive worker behavior, or result in repeated episodes of queenlessness. In this study, the dataset was utilized as a foundation for training and evaluating audio classification models aimed at supporting early intervention and informed decision-making in precision apiculture.

3. Results

To evaluate the effectiveness of the proposed acoustic classification system, we compared two deep learning models: a traditional CNN and ViT. Both models were trained and tested on the Kaggle Beehive Sounds dataset, which contains labeled audio clips corresponding to the following four queen-related hive states.

- Original queen present
- Queen absent
- New queen rejected
- New queen accepted

Before training, raw audio recordings were processed using STFT to generate spectrograms, which were used as the input representations for both models. To ensure fair comparison, both models were trained under identical settings: a batch size of 32, a learning rate of 2×10^{-5} , the AdamW optimizer, and 150 training epochs. The models were implemented in PyTorch (version 1.9.0) and trained using an NVIDIA GeForce RTX 4060 GPU.

Both the ViT and CNN models demonstrated strong performance in differentiating subtle acoustic differences between the rejected new queen and accepted new queen categories, with each model excelling in different aspects. After training the models with a maximum of 150 epochs and using a patience of 10, the ViT model stopped converging at epoch 77, while the CNN model converged at epoch 133. Furthermore, a critical difference for edge deployment is the model's final inference file size: the CNN model's file was 121 MB, whereas the ViT model's was 335 MB. The training and validation accuracy progression over the epochs is visualized in Figure 4. The final quantitative performance metrics for both models are detailed in Table 1, which compares the precision, recall, and F1-score values.

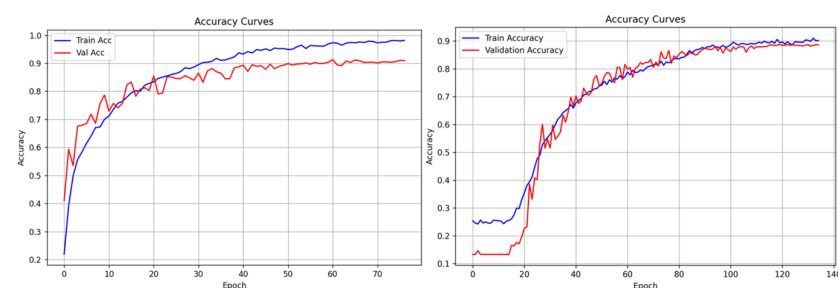


Figure 4. Training and validation accuracy curves for the ViT (left) and CNN (right) models.

Table 1. Comparison of experimental results between CNN and Transformer.

Model	Precision	Recall	F1-Score
CNN	0.89	0.86	0.85
ViT-transformer	0.90	0.86	0.87

Overall, the results validate the feasibility of using both CNN and Transformer-based models for fine-grained acoustic classification in beehive environments. The proposed

system not only provides accurate and real-time monitoring of queen status but also represents a promising tool for early intervention in colony management.

4. Conclusions

We developed a practical and intelligent acoustic health monitoring system tailored for honeybee colonies, using a three-tier architecture and deep learning classification. Through non-invasive wingbeat audio sensing and spectrogram analysis, the system accurately identifies four key queen-related hive states. The experimental results confirm that both the Vision Transformer and CNN models perform well, with each model excelling in different aspects of queen transition classification. By integrating IoT sensing, edge AI inference, and cloud-based data visualization, this system offers a low-cost, scalable, and real-time monitoring solution that enhances decision-making for beekeepers. However, the reliance on a publicly available dataset highlights the need for localized acoustic data to further optimize model performance. Future work will focus on collecting native Taiwanese beehive sounds, enabling model adaptation to regional conditions, and supporting broader deployment in smart apiculture systems.

Although both the ViT and CNN models showed comparable results in terms of classification performance, the ViT model's inference file (335 MB) is nearly 3 times larger than the CNN's (121 MB). In the context of limited computational power and storage on platforms like the Raspberry Pi, the CNN model emerges as the more suitable choice for real-time classification tasks due to its significantly smaller file size and lower resource overhead. This insight opens up a promising research direction focused on optimizing model architectures for edge devices, ensuring high performance without compromising system efficiency.

Author Contributions: Conceptualization and Funding acquisition, C.-C.S.; methodology, L.-H.C.; software, L.-H.C.; validation, S.-Y.Z. and J.-W.H.; Formal analysis and Writing—review and editing, L.-H.C.; Investigation and data curation, S.-Y.Z. and J.-W.H. project administration and writing—original draft, L.-H.C. and C.-C.S.; resources and supervision, C.-C.S.; visualization, S.-Y.Z. and J.-W.H. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the Miaoli District Agricultural Research and Extension Station, Ministry of Agriculture, Taiwan (grant numbers: 114AG27 and 114AG38).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Gallai, N.; Salles, J.-M.; Settele, J.; Vaissière, B.E. Economic valuation of the vulnerability of world agriculture confronted with pollinator decline. *Ecol. Econ.* **2009**, *68*, 810–821. [[CrossRef](#)]
2. Goulson, D.; Nicholls, E.; Botías, C.; Rotheray, E.L. Bee declines driven by combined stress from parasites, pesticides, and lack of flowers. *Science* **2015**, *347*, 1255957. [[CrossRef](#)] [[PubMed](#)]
3. Meikle, W.G.; Holst, N. Application of continuous monitoring of honeybee colonies. *Apidologie* **2014**, *46*, 10–22. [[CrossRef](#)]
4. Di, N.; Zhu, C.; Hu, Z.; Sharif, M.Z.; Yu, B.; Liu, F. Honeybee colony soundscapes: Decoding distance-based cues and environmental stressors. *Ecotoxicol. Environ. Saf.* **2025**, *297*, 118241. [[CrossRef](#)]
5. Abdollahi, M.; Coallier, N.; Giovenazzo, P.; Falk, T.H. Performance Comparison of Voice Activity Detectors for Acoustic Beehive Monitoring. In Proceedings of the 2023 IEEE Canadian Conference on Electrical and Computer Engineering (CCECE), Regina, SK, Canada, 24–27 September 2023; pp. 320–323. [[CrossRef](#)]
6. Yang, A. *Smart Bee Colony Monitor: Clips of Beehive Sounds*; About Dataset; Kaggle: San Francisco, CA, USA, 2022. [[CrossRef](#)]

7. Nolasco, I.; Terenzi, A.; Cecchi, S.; Orcioni, S.; Bear, H.L.; Benetos, E. Audio-based Identification of Beehive States. In Proceedings of the 2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP 2019), Brighton, UK, 12–17 May 2019; pp. 8256–8260. [[CrossRef](#)]
8. Kampelopoulos, D.; Sofianidis, I.; Tananaki, C.; Tsiapali, K.; Nikolaidis, S.; Siozios, K. Analyzing the Beehive's Sound to Monitor the Presence of the Queen Bee. In Proceedings of the 2022 Panhellenic Conference on Electronics & Telecommunications (PACET), Tripolis, Greece, 2–3 December 2022; pp. 1–4. [[CrossRef](#)]
9. Allamaprabhu, B.M.; Sanket, U.G.; CG, R.; Shreyas, A.R.; Mahantesh, M.B. Review on Machine Learning Models for Efficient Queen Bee Identification in Beehives. In Proceedings of the 2025 International Conference on Intelligent Computing and Control Systems (ICICCS), Erode, India, 19–21 March 2025; pp. 1383–1390. [[CrossRef](#)]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.