



Universidad Autónoma de Chihuahua

Facultad de Ingeniería

Detección de la Abeja Reina en una Colmena por Medio del Aprendizaje Profundo

Análisis de Frecuencias Emitidas por el Panal

Student: Leonardo Rafael León Mora.

Thesis director: Dr. Daniel Espinobarro Velázquez.

Advisors:

- M.C. Carlos Hugo Larrinúa Pacheco.
- M.I. Joseph Isaac Ramírez Hernández.

Chihuahua, Chih., 12 de agosto de 2025.

Índice general

RESUMEN	2
1. INTRODUCCIÓN	3
1.1. Planteamiento del Problema	3
1.2. Justificación	3
1.3. Objetivos	4
1.3.1. Objetivo General	4
1.3.2. Objetivos Específicos	4
1.4. Hipótesis de Trabajo	4
2. MARCO TEÓRICO	5
2.1. Definición de Conceptos Clave	5
3. METODOLOGÍA	6
3.1. Procedimiento	6
3.2. Instrumentos	6
3.3. Base de Datos	6
4. RESULTADOS	8
5. DISCUSIÓN	9
6. CONCLUSIONES Y RECOMENDACIONES	10
6.1. Conclusiones	10
BIBLIOGRAFÍA	11
APÉNDICES	12

RESUMEN

1. INTRODUCCIÓN

Los polinizadores, como las abejas melíferas, abejorros, abejas solitarias, mariposas, colibríes y murciélagos, desempeñan un papel fundamental en los ecosistemas, ya que contribuyen a la fertilización de plantas y cultivos. Las abejas melíferas, en particular, han sido ampliamente estudiadas por su destacada participación en la polinización de cultivos que forman parte de nuestra alimentación.

Durante las temporadas de mayor producción de miel, como primavera y verano, las abejas melíferas pueden recorrer distancias de hasta 6 km [1], viajando 3 km de ida y vuelta entre la colmena y fuentes de alimento como flores, para recolectar polen, néctar o agua. Pero, ¿por qué surge este comportamiento? ¿Cuál es la motivación para las abejas en cuanto a regresar a la colmena y surtir la misma?

Según Ruvinga [2], *“La colonia está formada por la reina, miles de abejas obreras estériles, unos pocos cientos de zánganos, huevos, larvas y pupas”*. Las abejas trabajan a diario, realizando tareas que se dividen no solo entre machos y hembras, sino también entre abejas jóvenes y adultas [3]. Un panal puede albergar entre 30,000 y 50,000 abejas, aunque este número puede reducirse hasta 5,000 durante el invierno. Este comportamiento de abastecimiento constante está estrechamente vinculado a la estructura social de la colmena y, en particular, al papel de la abeja reina, que es la única capaz de procrear [2].

1.1. Planteamiento del Problema

La reina puede poner en promedio hasta 2,000 huevos al día. Las obreras se encargan de construir, defender y abastecer la colmena con recursos esenciales como agua, polen y néctar [4]. Los zánganos cumplen funciones reproductivas

y también se cree que ayudan a regular la temperatura del panal, mientras que las abejas más jóvenes se limitan a tareas dentro de la colmena [3]. La pérdida de la abeja reina puede comprometer seriamente la supervivencia de la colonia. Las colmenas están expuestas a múltiples amenazas: insectos invasores, pesticidas, condiciones climáticas y hasta errores humanos. En sistemas apícolas tradicionales donde se utilizan cajas de madera, cada inspección física implica abrir la colmena y romper el sello de propóleo que las abejas han construido para mantener un ambiente limpio y controlado. Esta intervención frecuente no solo representa una pérdida energética para la colonia (que debe reconstruir el sello), sino que además incrementa el riesgo de aplastar accidentalmente a la reina durante la manipulación interna del panal. Esta situación puede conducir al colapso de toda la colonia.

1.2. Justificación

Contar con herramientas que permitan identificar la presencia o ausencia de la abeja reina de manera no invasiva resulta fundamental para garantizar la continuidad y productividad de la colmena. Esta investigación propone un enfoque basado en el análisis acústico del panal, aprovechando técnicas de aprendizaje automático para detectar el estado de la reina. Dado que estudios recientes han demostrado que el sonido dentro de la colmena varía según la presencia o ausencia de la reina [2, 5, 6, 7], se busca validar un sistema que, a partir de grabaciones de audio, identifique estas variaciones de forma automática y eficiente.

1.3. Objetivos

1.3.1. Objetivo General

Desarrollar un programa de monitoreo de frecuencias con el que se pueda determinar si una abeja reina se encuentra en la colmena, mediante el análisis de grabaciones de audio y técnicas de aprendizaje automático, a partir del cual se pueda proponer un sistema no invasivo de seguimiento para determinar el estado de la colmena.

1.3.2. Objetivos Específicos

- ☞ Realizar una revisión bibliográfica sobre la estructura de las colonias de abejas melíferas y los efectos de la ausencia de la reina.
- ☞ Procesar y organizar un conjunto de datos de audio de colmenas etiquetado con el estado de la reina.

- ☞ Extraer características acústicas relevantes, como los coeficientes cepstrales en la escala Mel (MFCC), a partir de las grabaciones.
- ☞ Entrenar y evaluar modelos de aprendizaje automático para clasificar el estado de la reina basado en características extraídas.
- ☞ Validar la efectividad del sistema en términos de precisión, sensibilidad y especificidad.

1.4. Hipótesis de Trabajo

Dado que el comportamiento sonoro de una colmena varía según el estado de su reina, se espera que un sistema basado en el análisis de características acústicas sea capaz de detectar, con un grado aceptable de precisión, si la reina está presente, ausente, rechazada o recién aceptada en la colonia.

2. MARCO TEÓRICO

2.1. Definición de Conceptos Clave

- ☞ **Colmena:** estructura donde viven las abejas melíferas, compuesta por una reina, obreras y zánganos.
- ☞ **Reina:** abeja hembra fértil encargada de la reproducción en la colmena.
- ☞ **Razón de muestreo (Sample Rate):**

número de muestras de audio tomadas por segundo (cuyos valores para el monitoreo de abejas comunmente rondan entre los 16000 a 22050 Hz).

- ☞ **MFCC (Mel-Frequency Cepstral Coefficients):** características acústicas utilizadas para representar el contenido espectral de una señal de audio, especialmente útiles en el reconocimiento de patrones sonoros.

3. METODOLOGÍA

3.1. Procedimiento

Para desarrollar un sistema de detección automática con respecto a la presencia de la abeja reina en una colmena, se llevó a cabo un proceso que incluyó la recolección de datos, el análisis de audio y la implementación de técnicas de aprendizaje automático. El objetivo fue crear un modelo capaz de clasificar el estado de la reina en función de las características acústicas extraídas de grabaciones de audio.

Para ello se debe determinar el formato de los archivos de audio y la razón de muestreo, esto se puede hacer por medio un programa en Python (6.1). En este estudio, se trabajó con audio monofónico (1 canal), con formato PCM de 16 bits y una razón de muestreo de 22050 Hz. Posteriormente, se realizó una prueba de extracción de características acústicas, específicamente los coeficientes cepstrales en la escala Mel (MFCC), con apoyo de la librería `librosa` (6.2). En este programa se utilizó un archivo de audio (2022-06-08--14-52-28_1_ _segment3.wav), a partir del cual se obtuvo una matriz tiempo-frecuencia:

$$[\text{Time Frames}] \times [13 \text{ MFCC Coefficients}] ,$$

correspondiente a 13 coeficientes calculados por cada 25 milisegundos de audio, con un solapamiento del 50 % entre ventanas. Los resultados fueron exportados en formatos `.csv` y `.npy` para su posterior análisis.

3.2. Instrumentos

Para el proceso de recolección de datos, se empleó un dispositivo IoT diseñado a medida, que incluía un módulo Wi-Fi ESP32, un micrófono digital INMP441, y un sensor de temperatura y humedad BME280. Estos sensores se instalaron dentro de la cubierta telescópica de la colmena —por encima de los cuadros pero por debajo

de la cubierta exterior—, y se protegieron con una malla metálica para evitar la interferencia de las abejas con los componentes electrónicos, especialmente el micrófono.

3.3. Base de Datos

El conjunto de datos utilizado en esta investigación fue grabado y publicado por Anna Yang [8]. Contiene datos de audio recolectados de colmenas de abejas melíferas europeas ubicadas en California. Las grabaciones fueron segmentadas en clips de 60 segundos, resultando en un total de 7,100 muestras recolectadas durante un período de cinco semanas, desde el 8 de junio de 2022 (14:52:28) hasta el 15 de julio de 2022 (15:28:21). Cada día contiene aproximadamente entre 23 y 24 muestras.

El conjunto de datos incluye múltiples variables de metadatos que describen el entorno de la colmena y la condición de la colonia. De particular relevancia para este trabajo se encuentran las siguientes etiquetas:

- 🔊 **queen acceptance:** indica si la reina es aceptada por la colonia. Valores: 0 (sin reina), 1 (presente pero no aceptada), 2 (aceptada).
- 🔊 **queen presence:** indica si la reina está físicamente presente en la colmena.
- 🔊 **queen status:** una etiqueta combinada derivada de las dos anteriores, con los siguientes valores: 0 (original/con reina funcional), 1 (no presente), 2 (presente y rechazada), 3 (presente y recién aceptada). Esta es la etiqueta principal utilizada para la clasificación.
- 🕒 **time:** hora del día, escalada entre 0 y 1 para representar un ciclo de 24 horas.

El proceso de etiquetado se basó en inspecciones manuales diarias de las colmenas, durante las cuales se observó el comportamiento de las abejas hacia la reina (por ejemplo, formación

de enjambre o conducta tranquila). En aquellos casos en los que el estado de la reina cambió entre inspecciones, generando ambigüedad sobre

el momento exacto de aceptación, las muestras correspondientes fueron descartadas para mantener la integridad del etiquetado.

4. RESULTADOS

5. DISCUSIÓN

6. CONCLUSIONES Y RECOMENDACIONES

6.1. Conclusiones

Bibliografía

- [1] M. Beekman y F. L. W. Ratnieks. «Long-range foraging by the honey-bee, *Apis mellifera* L». En: *Functional Ecology* 14 (2000), págs. 490-496. DOI: [10.1046/j.1365-2435.2000.00443.x](https://doi.org/10.1046/j.1365-2435.2000.00443.x).
- [2] Stenford Ruvinga et al. «Use of LSTM networks to identify “queenlessness” in honeybee hives from audio signals». En: *2021 17th International Conference on Intelligent Environments (IE)*. IEEE. 2021, págs. 1-4.
- [3] Brian R Johnson. «Division of labor in honeybees: form, function, and proximate mechanisms». En: *Behavioral ecology and sociobiology* 64 (2010), págs. 305-316.
- [4] ML Winston. *Temperate and Tropical Honey Bees in The Biology of the Honey Bee*. 214–224. 1987.
- [5] Hichem Maaref y Dominique Fourer. «Remote Beehive Health Analysis using Embedded System and Relevant Audio Features». En: ().
- [6] D Hunter y K Stebel. «Signal processing the acoustics of honeybees (*Apis Mellifera*) to identify the ‘Queenless’ state in hives». En: *Proceedings of the Institute of Acoustics* 35.1 (2013), pág. 290.
- [7] Stenford Ruvinga et al. «Identifying queenlessness in honeybee hives from audio signals using machine learning». En: *Electronics* 12.7 (2023), pág. 1627.
- [8] Anna Yang. *Smart Bee Colony Monitor: Clips of Beehive Sounds*. <https://www.kaggle.com/datasets/annajyang/beehive-sounds>. [Online; accessed 29-April-2025]. 2023.

APÉNDICE A

```
1 import numpy as np
2 import matplotlib.pyplot as plt
3 from scipy.io import wavfile
4
5 # cargar archivo de audio
6 sr, audio = wavfile.read('2022-06-08--14-52-28_1__segment3.wav')
7
8 # normalizar si es int16
9 if audio.dtype == np.int16:
10     audio = audio.astype(np.float32) / np.iinfo(np.int16).max
11
12 # crear eje de tiempo
13 duration_sec = len(audio) / sr
14 time_axis = np.linspace(0, duration_sec, num=len(audio))
15
16 # graficar la onda
17 plt.figure(figsize=(12, 4))
18 plt.plot(time_axis, audio)
19 plt.title(f'Onda (Sample Rate: {sr} Hz)')
20 plt.xlabel('Tiempo (segundos)')
21 plt.ylabel('Amplitud')
22 plt.grid(True)
23 plt.tight_layout()
24 plt.show()
```

Listing 6.1: Visualización de la forma de onda de un archivo de audio.

APÉNDICE A (continuación)

```
1 import librosa
2 import numpy as np
3 import pandas as pd
4 import os
5
6 # === CONFIG ===
7 audio_path = <<archivo>>
8 sample_rate = 22050
9 n_mfcc = 13
10
11 # === CARGAR ARCHIVO ===
12 y, sr = librosa.load(audio_path, sr=sample_rate)
13
14 # === EXTRAER MFCCs ===
15 mfccs = librosa.feature.mfcc(y=y, sr=sr, n_mfcc=n_mfcc)
16
17 # Transponer de modo que la forma resultante es: [frames x coefficients]
18 mfccs = mfccs.T
19
20 # === Guardar en CSV ===
21 base_name = os.path.splitext(os.path.basename(audio_path))[0]
22 csv_output_path = f'{base_name}_mfcc.csv'
23 np_output_path = f'{base_name}_mfcc.npy'
24
25 pd.DataFrame(mfccs).to_csv(csv_output_path, index=False)
26
27 # 0, guardar como un archivo binario de numpy
28 np.save(np_output_path, mfccs)
29
30 print(f'MFCCs en:\n- CSV: {csv_output_path}\n- NPY: {np_output_path}')
31 print(f'MFCCs forma: {mfccs.shape} to [frames x {n_mfcc}']')
```

Listing 6.2: Prueba para extracción de características.