Universidad Autónoma de Chihuahua

Facultad de Ingeniería

MONITORED HIVES FOR HONEYBEES: Implementing a Microphone on the Monitoring Device

---

Thesis Protocol.

Student: Leonardo Rafael León Mora.

Thesis director: Dr. Daniel Espinobarro Velázquez.

Suggested advisors:

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

Chihuahua, Chih., June 19, 2023.

1. BACKGROUND

Cada vez que se revisa la colmena de forma manual existe la posiblidad de matar a la reina al extraer uno de los bastidores de la colmena, lo que puede llevar a la perdida total de la misma.

1.1 Dániel Tamás Várkonyi - Dynamic noise filtering for multi-class classification of beehive audio data

In the majority of apiaries, identification of the health condition of a bee colony is done manually by opening and inspecting the hive. Opening the hive, however, introduces certain stress to the colony while changing the micro-climate within the hive. Afterward, bees have to expend considerable energy to re-establish the equilibrium within the beehive. Consequently, **frequent manual inspection of a hive reduces the amount of honey the given bee colony produces**.

Analyzing the colony's sound might reveal certain anomalous events within the hive, like the presence of an intruder or **the preparation of the colony for swarming**. The first technological approaches used for monitoring bees' condition via audio analysis were conducted in the late 20th century using spectral analysis in the range of $0 - 3$ kHz [1].

Applying machine learning (`ML`) to audio analysis for bee queen presence detection and swarming prediction leads to a basic binary classification task, i. e. predicting if the queen is present or not and if the colony is swarming or not. Moreover, **the sound of a bee colony in a "queenless" or swarming state is well distinguishable from its sounds when a queen is present or the colony is not swarming** [2, 3]. **The stress level of the colony exposed to an "anomaly" likely implies a substantially different sound from the one when being in a normal state**.
An important issue, contravening the approaches found in the literature, is that **these problems might not necessarily correspond to binary classification tasks**. For example, identification of various (more than two) types of diseases, intrusion detection by different pests, or estimation of exposure of bees to diverse palettes of chemicals, naturally, calls for multi-label or even multi-class approaches. An assumption here, posing possible difficulties for application of ML techniques, is that **bee sounds corresponding to various classes (labels) might not be so easily distinguishable from each other** as they usually are in the before mentioned binary cases where "anomalous" and "normal" states are well-detectable even for humans.

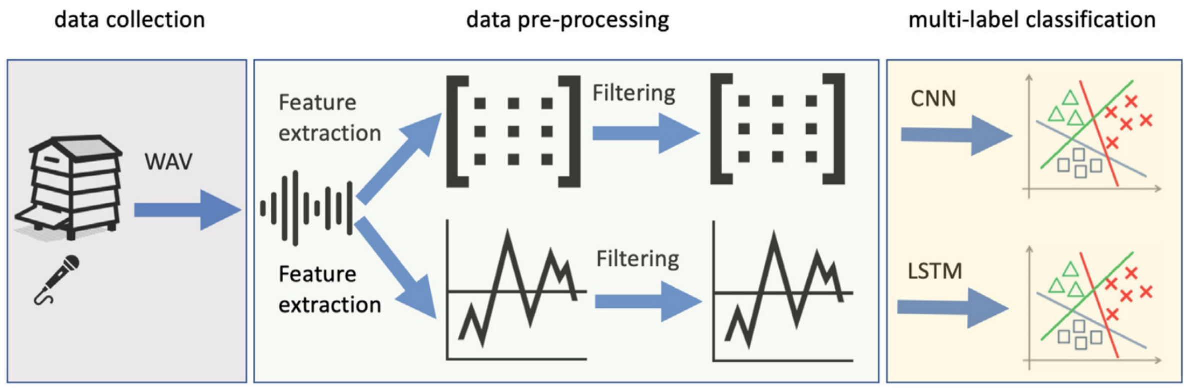

Figure 1.1: A general audio data analytics workflow used in our experiments.

There is a huge gap int he literature considering multi-label or multi-class classification of bee audio data. According to our knowledge. The only work related to multi-label bee sound classification [4], is focused on identifying 12 bumblebee species from their buzzing sounds (from a small-scale dataset using traditional `ML` techniques, such that naïve Bayes decision trees, support vector machines, and random forests).

- **Problem # 1**. The classification accuracy of the audio data analytics work flow, shown in figure 1.1, does not only depend on the discriminative power of the used `ML` techniques, but also, on the used pre-processing steps. One of the basics problems of audio signals is that, even after filtering out irrelevant information, the resulting data represented as time-series have very high dimensionality. Feature extraction (`FE`) methods are utilized to extract and represent the most important features within the audio signal in a compact form (time-series or spectrograms), preserving a significant portion of its original information content. After `FE`, `ML` techniques can be applied on the data (depending on the amount, dimensionality and the task intended to be solved).

- **Problem # 2**. According to our knowledge, there is no reference in the literature to a survey comparing various combinations of audio `FE` (image or time-series), noise filtering (`NF`) and `ML` (image or time-series classification) techniques for bee sound analytics. including thorough hyper-parameter (`HP`) tuning and validation procedures as well as utilizing large datasets.

- **Problem # 3**. According to our knowledge, there is no large-scale publicly available bee sound benchmark data suitable for research on multi-label or multi-class bee sound classification approaches.
  Several `ML` developments, e.g. various deep learning (`DL`) models, have been developed working efficiently on specific problems related to audio processing like, for example, human speech recognition. However, our experiments showed that **these methods are not performing**

well in case of a multi-label bee sound classification problem in which the difference between various classes is inconspicuous.

2. JUSTIFICACIÓN

3. RELACIÓN DE LA PROPUESTA CON LAS MATERIAS DE LA CARRERA

4. OBJETIVOS

5. HIPÓTESIS

6. METODOLOGÍA

7. PLAN DE TRABAJO

8. LUGAR DE DESARROLLO

Bibliography

[1] David G Dietlein. “A method for remote monitoring of activity of honeybee colonies by sound analysis”. In: *Journal of Apicultural Research* 24.3 (1985), pp. 176–183.

[2] M Delia Allen. “The behaviour of honeybees preparing to swarm”. In: *The British Journal of Animal Behaviour* 4.1 (1956), pp. 14–22.

[3] Loren Hord and Errol Shook. “Determining honey bee behaviors from audio analysis”. In: *National Science Foundation Research Experience for Teachers. Appalachian State Univ* (2013).

[4] Anton Gradišek et al. “Predicting species identity of bumblebees through analysis of flight buzzing sounds”. In: *Bioacoustics* 26.1 (2017), pp. 63–76.