



21st International Conference on

Intelligent Systems Design and Applications

On the World Wide Web

December 13-15, 2021

## Honey Bee Queen Presence Detection from Audio Field Recordings using Summarized Spectrogram and Convolutional Neural Networks

Agnieszka Orlowska, **Dominique Fourer**, Jean-Paul Gavini and Dominique Cassou-Ribehart

December 13, 2021



## Purpose of this work



### Goals

- A supervised precision beekeeping method for detecting bee queen from audio field recordings
- **Enhancing the generalization capability of the trained model for unknown audio beehive recordings**
- Evaluation through a beehive-independent cross-validation methodology

⇒ **Proposed approach:** Time-frequency analysis combined with a supervised deep convolutional neural network (CNN)

## Context: the French Starlings Partners Project



des abeilles & nous

Project holder: Dominique Cassou-Ribehart with Jean-Paul Gavini

### Goals:

- Precision beekeeping using AI methods to efficiently monitor beehives health state
- Reducing the stress due to beehive inspection to protect bees
- Collaboration between Starlings Partners and professional beekeepers to collect data for a better understanding of bees behaviors
- Embedded, non-intrusive and low-cost hardware dedicated solution

<https://desabeillesetnous.fr/>

# Content

- 1 Introduction
- 2 Proposed approach
- 3 Numerical Experiments
- 4 Conclusion

## Audio-based Smart Beekeeping

### Principle

Smart monitoring methods for predicting the health state of a beehive.

- Supervised/semi-Supervised machine learning approaches
- Sensors and Embedded hardware systems designed for beehives.
- Multi-modal data collection and processing (eg. audio, weight, hygrometry, temperature, CO<sub>2</sub>, etc.) [Cecchi et al. 2018]

### Motivation

Aiding beekeepers to preserve bio-diversity

A promising and recently emerging field of research since the late 2010s.<sup>1</sup>

---

<sup>1</sup>Cecchi, S., Terenzi, A., Orcioni, S., Riolo, P., Ruschioni, S., & Isidoro, N. (2018, May). A preliminary study of sounds emitted by honey bees in a beehive. In Audio Engineering Society Convention 144. Audio Engineering Society.

## The audio-based approach

Audio signals can convey a lot of relevant informations about the state of a beehive.<sup>2</sup>

	Freq. (Hz)	Signal Pattern	Sender	Possible Sig.
<b>Tooting</b>	300 ~ 500	Pulse sequence	Queen	Prevent hatching of further queens and trigger quacking
<b>Quacking</b>	300 ~ 350	Pulse sequence	Queen	Presence detection, viability of confined queens
<b>Hissing</b>	300 ~ 3600	Single pulse	Colony	Warning signal
<b>Piping</b>	100 ~ 2000	Single pulse	Scout	Triggers colony hissing, prepare for swarming
<b>Recruit</b>	200 ~ 350	Pulse sequence	Forager	Existence and quality of valuable food source

<sup>2</sup>Qandour, A., Ahmad, I., Habibi, D., & Leppard, M. (2014). Remote beehive monitoring using acoustic signals. Tech. Report. CCER, Australia.

## Prediction of the bee queen presence from audio signals

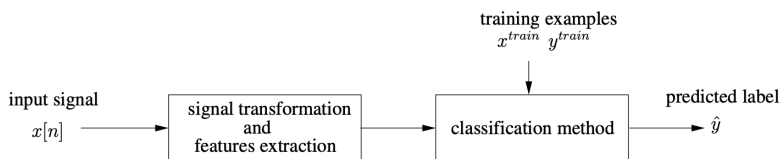


Figure : Illustration of the overall proposed approach.

### Problem Formulation

- Supervised prediction of the bee queen presence ( $\hat{y}$ ) from audio recordings  $x$
- Training of the classification method using annotated data  $x^{train}, y^{train}$  by minimizing the prediction error  $y - \hat{y}$ .

## Materials

- Publicly available dataset introduced in [Nolasco, Benetos 2018]<sup>3</sup> in the Open Source Beehive (OSBH) project and the NU-Hive project<sup>4</sup>.
- Bee signals acquired from six distinct beehives (“no bee” signals are ignored)
- Each audio recording is resampled at rate of  $F_s = 22.05$  kHz
- Each recording is split in one-second-long homogeneous time series (associated to the same annotation label).
- 17,295 distinct individuals where 8,444 ones are labeled as “queen” ( $y = 1$ ) and 8,851 ones are labeled as “no queen” ( $y = 0$ ).

Beehive name	<i>queen</i>	<i>no queen</i>	Total
CF001	0	16	16
CF003	3,700	0	3,700
CJ001	0	802	802
GH001	1,401	0	1,401
Hive1	2,687	1,476	4,163
Hive3	656	6,557	7,213
Total	8,444	8,851	17,295

<sup>3</sup>I. Nolasco and E. Benetos. To bee or not to bee: Investigating machine learning approaches for beehive sound recognition. Proc. DCASE 2018.

<sup>4</sup><https://zenodo.org/record/1321278>.



## Summarized Spectrogram Computation

Given a discrete-time finite-length signal  $x[n]$ , with time index  $n \in \{0, 1, \dots, N - 1\}$ , and an analysis window  $h$ , the discrete STFT of  $x$  can be computed as:

$$F_x^h[n, m] = \sum_{k=-\infty}^{+\infty} x[k]h[n - k]^* e^{-j\frac{2\pi mk}{M}} \quad (1)$$

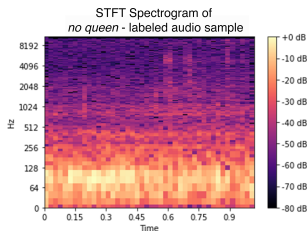
with  $z^*$  the complex conjugate of  $z$ ,  $j^2 = -1$ , and  $m \in \{0, 1, \dots, M - 1\}$  the frequency indices.  $|F_x^h[n, m]|^2$  being the classical Spectrogram.

The Summarized Spectrogram is obtained by down-sampling the spectrogram along the frequency axis as:

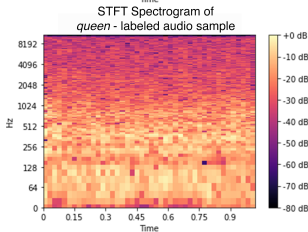
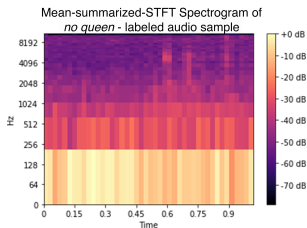
$$\text{SF}_x^h[n, b] = g \left( |F_x^h[n, m_b]|^2 \right)_{\forall m_b \in [b \lfloor \frac{M}{2B} \rfloor, (b+1) \lfloor \frac{M}{2B} \rfloor - 1]} \quad (2)$$

with  $g()$  the summarizing function and  $B$  the desired number of quantization levels.

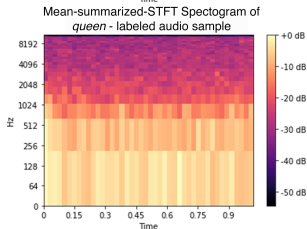
## Summarized Spectrogram Examples (arithmetic mean function)



summarizing  
function



summarizing  
function



## Deep Convolutional Network

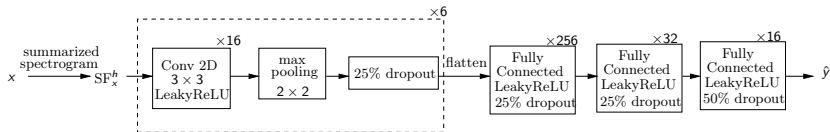


Figure : Diagram of the proposed deep neural network architecture.

### CNN 2d architecture

- 6 convolutional blocks including with a  $3 \times 3$  kernel size with a LeakyReLU activation, followed by a batch normalization, a  $2 \times 2$  max-pooling and a 25% dropout layers.
- 3 fully-connected (FC) layers including 2 dropout layers of respectively 25% and 50%.

## Experimental Setup

### Experiment 1:

We merge the 6 available beehives and then we apply a random split to obtain 70% of the individuals for training and 30% for testing.

### Experiment 2:

We use a 4-fold cross-validation methodology where the beehives are independent. To this end, the folds have been manually created to assign each beehive to a unique fold as detailed in Table 1.

Table : Description of the partitioned dataset investigated in Experiment 2.

Fold	Training set	Testing Set
Fold 1	CJ001 + GH001 + Hive3 + Hive 1	CF001 + CF003
Fold 2	CF001 + CF003 + Hive3 + Hive 1	CJ001 + GH001
Fold 3	CJ001 + GH001 + Hive3 + CF001 + CF003	Hive1
Fold 4	CJ001 + GH001 + Hive1 + CF001 + CF003	Hive3

	Fold 1	Fold 2	Fold 3	Fold 4
<i>queen</i>	3700	1401	2687	656
<i>no queen</i>	16	802	1476	6557
Total	3716	2203	4163	7213

## Hyperparameters Tuning

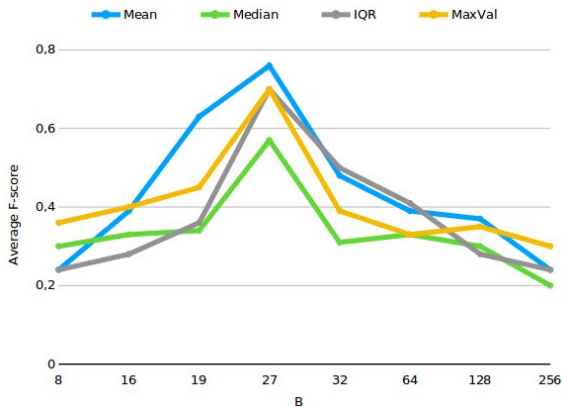


Figure : Average F-measure for different summary function  $g$  and  $B$  value configurations in Experiment 2. The best value is reached for  $B = 27$  using the mean function as  $g$ .

## Comparative Results - Random split

Table : Comparison of the classification results in Experiment 1 (random split).

Method	Features	Label	Precision	Recall	F-score	Accuracy
MFCCS+CNN [Benetos et al. 19]	20×44	Queen	1.00	0.99	0.99	0.99
		No queen	0.99	1.00	0.99	
STFT+CNN	513×44	Queen	1.00	0.93	0.97	0.97
		No queen	0.94	1.00	0.97	
CQT+CNN	513×44	Queen	0.96	0.93	0.95	0.95
		No queen	0.92	1.00	0.95	
mean-CQT+CNN	27×44	Queen	0.98	1.00	0.99	0.99
		No queen	0.99	0.98	0.98	
mean-STFT+CNN	27×44	Queen	<b>0.99</b>	<b>1.00</b>	<b>1.00</b>	<b>1.00</b>
		No queen	<b>1.00</b>	<b>0.99</b>	<b>1.00</b>	
mean-STFT+CNN+DA	27×44	Queen	<b>0.99</b>	<b>1.00</b>	<b>1.00</b>	<b>1.00</b>
		No queen	<b>1.00</b>	<b>0.99</b>	<b>1.00</b>	

- All the methods provide satisfying classification results in a random-split evaluation experiment<sup>5</sup>.

---

<sup>5</sup>Data Augmentation (DA) consists in artificially increasing of 50% the size of the training dataset with examples merged with a white Gaussian noise to obtain a SNR of 30 dB.

## Comparative Results - beehive-independent 4-fold-cross validation

Table : Comparison of the classification results in Experiment 2 (4-fold hive-independent cross-validation).

Method	Features	Label	Precision	Recall	F - score	Accuracy
MFCCs+CNN [Benetos et al. 19]	20x44	Queen No queen	0.36 0.22	0.44 0.16	0.40 0.19	0.31
STFT+CNN	513x44	Queen No queen	0.77 0.33	0.76 0.20	0.66 0.33	0.55
CQT+CNN	513x44	Queen No queen	0.10 0.32	0.07 0.41	0.08 0.36	0.25
mean-CQT+CNN	27x44	Queen No queen	0.25 0.41	0.11 0.65	0.16 0.50	0.38
mean-STFT+CNN	27x44	Queen No queen	0.71 0.81	0.86 0.64	0.78 0.71	0.75
<b>mean-STFT+CNN+DA</b>	27x44	Queen No queen	<b>0.96</b> <b>0.99</b>	<b>0.99</b> <b>0.94</b>	<b>0.96</b> <b>0.96</b>	<b>0.96</b>

- Only the STFT-based methods provide satisfying classification results in a random-split evaluation experiment.
- All the other methods obtain very poor results using of not DA (Accuracy < 0.5)

## Conclusion and future work

### Contributions summary

- A beehive-independent comparative evaluation to investigate the generalization capability of existing bee queen detection methods based on the public Nu-Hive dataset
- A promising approach for both reducing the computational cost and the input size of a CNN-based audio classification method
- Improving the generalization capability of the trained model of the proposed CNN-based neural architecture

### Future work

- A further investigation of the summary spectrogram to explain why this method is efficient
- Optimization of the summary function  $g()$

Code freely available for the sake of reproducible research at  
[https://github.com/agniorlowska/beequeen\\_prediction](https://github.com/agniorlowska/beequeen_prediction)