We are IntechOpen,
the world's leading publisher of
Open Access books
Built by scientists, for scientists

4,800
Open access books available

122,000
International authors and editors

135M
Downloads

Our authors are among the

154
Countries delivered to

TOP 1%
most cited scientists

12.2%
Contributors from top 500 universities

CLARIVATE ANALYTICS
BOOK CITATION INDEX
INDEXED

WEB OF SCIENCE™

Selection of our books indexed in the Book Citation Index
in Web of Science™ Core Collection (BKCI)

Interested in publishing with us?
Contact book.department@intechopen.com

Numbers displayed above are based on latest data collected.
For more information visit www.intechopen.com

Chapter

A Case Study of Wavelets and SVM Application in Coffee Agriculture: Detecting Cicadas Based on their Acoustic and Image Patterns

João Paulo Lemos Escola, Rodrigo Capobianco Guido,
Alexandre Moraes Cardoso,
Douglas Henrique Bottura Maccagnan,
João Marcelo Ribeiro and José Ricardo Ferreira Cardoso

Abstract

One of the main problems in agriculture is crop pest management, which causes financial damage to farmers. This management is traditionally performed with pesticides; however, with a large area of application, it would be more economically viable and more environmentally recommended to know precisely the regions where there is concrete infestation. In coffee farms, cicada makes a distinctive sound when it hatches after years of underground nymph-shaped living. One possibility of contributing to its management would be the development of a device capable of capturing the sound of the adult cicada in order to detect its presence and to quantify crop insects. This device would be spread across the coffee plots to capture sounds within the widest possible area coverage. With monitoring and quantification data, the manager would have more input for decision-making and could adopt the most appropriate management technique based on concrete information on population density separated by crop region. Thus, this chapter presents an algorithm based on wavelets and support vector machines (SVMs), to detect acoustic patterns in plantations, advising on the presence of cicadas.

Keywords: acoustic patterns, support vector machines, wavelets, digital signal processing, cicada

1. Introduction

Since human population has been intensively increasing, the need for food and other products from agricultural fields also grows. To guarantee the production, monoculture is carried out at extensive areas, which intensifies the appearance of pests. Generally, pest control in agriculture is performed with the use of chemical pesticides, often applied even in areas without pest incidence, which raises the cost of production and may cause environmental impacts that affect human health.

Figure 1.
Quesada gigas. *On the left, male emitting acoustic signals. On the right, lateral view of male resting.*

The development of specific hardware and software for pest detection in agriculture can provide support for distinct production forms which reduce negative impacts. In this sense, the capture, recording, and analysis of acoustic signals emitted by insects can be an alternative to optimize the production of certain crops [1].

The cicada (Hemiptera: Cicadidae) is a good example of an insect capable of emitting acoustic signals. In Brazil, the coffee plantation can be attacked by several arthropods and among them, *Quesada gigas* is considered a key pest in the entire state of Minas Gerais and in the northeastern region of the state of São Paulo [2]. Considering that there have been reports on the occurrence of cicadas in coffee plantations from the period between 1900 and 1904, it is important to say that this fact has interfered in the way of conducting the crop: it has practically forced coffee growers to adopt practices for productive system improvement, such as larger spaces between plants, allowing for the mechanization of crop aiming at the use of pesticides to manage this and other pests and diseases that affect productivity.

Cicadas attack crops in the search for sap, their main food. The impacts on the plants occur in the nymphal phase of the cicada when it sucks sap from the host plant root [3]. The *Q. gigas* species, that are the biggest in size the country, can reach 70 mm in length, including wings, and 20 mm in width in the case of males. The females can reach 69 mm in total length and 16.5 mm in width, as shown in **Figure 1**. The size of the insect is probably the reason why it is associated with the impact caused in coffee plantation.

Males usually sound from October to December. In the 1970s, since there was no efficient method for cicadas controlling, many coffee growers had no choice than to eradicate infested crops. Many of them even abandoned their cultivation. However, lately the recommended control has been made through systemic chemical defenses and more recently by the use of a sound trap that attracted *Q. gigas* to a closed spraying system [4].

There are few technological devices used in coffee plantations to monitor and control the cicadas, and maybe, because of this, the most used device is the chemical one. Considering that the methods currently used for mapping and monitoring the populations of cicadas consist basically of the nonautomated counting of the individuals through direct observation, this paper aims to present some applications using digital signal processing and support vector machines (SVMs) as techniques for detecting and monitoring cicadas in crops and forests, reducing time and control costs.

2. Initial considerations

Machine learning is a branch of artificial intelligence that seeks to develop algorithms capable of learning certain behaviors or patterns through examples,

being able to generalize from training. SVMs, for instance, are supervised machine learning methods with superior results compared to other pattern classification procedures, considering binary problems [5].

Thus, one possibility for improving agricultural procedures would be the development of a device capable of capturing the sound of the adult cicada in order to detect its presence, thus monitoring and quantifying crop insects. This device would be spread across the coffee fields to capture sounds within the widest possible area coverage. That allows for the manager to have more input for decision-making, adopting the most appropriate management technique based on concrete information on population density separated by crop region.

Cicadas emit particular frequency components which characterize certain patterns. To detect them, both Fourier and discrete wavelet (DWT) transforms, which convert a time-domain signal to the frequency domain, can be used. Nevertheless, in the latter case, it is still possible to obtain the time support of frequencies [6]. Then, an SVM can be used to refine the results, pointing out the existence of the important patterns from the wavelet-transformed signals. This is just how the proposed approach was implemented. Related works, such as [1, 9–13, 15–18], perform automated data collection for monitoring and corroborate the present work.

3. Application one (AP1): cicada density estimation by audio processing

This solution consists of a system that, from an input *wav* audio file, discriminates between three possibilities, noise, low density, and high density, assisting in monitoring the cicada infestation in the coffee crop. Based on features from the human auditory system, which easily differentiates between these acoustic patterns, the system proves to be efficient. Working similarly to the cochlea in the human ear [7] and based on the DWT packet, an efficient time-frequency mapping [6] is provided.

Based on our assumptions and intending to assess it, we carried out the following preprocessing procedure to convert each acoustic input signal of variable length into a 25-sample long feature vector:

- AP1 PRE-PROCESSING PHASE:

  - BEGINNING.

  - STEP 1: the raw data from the input signal i, recorded as a *wav* file sampled at 44,100 samples/s, 16-bit [7], is extracted and stored as the vector $s_i[\cdot]$, for $(0 \leqslant i \leqslant X - 1)$. Each original *wav* file lasts about 3 s, i.e., $44100 \cdot 3 = 132300$ samples, with variations among them. Since a wavelet-based transformation is used in the next step, we cut the the vectors $s_i[\cdot]$ taking advantage of their central part in such a way that their length became 262,144 samples, which is a power of two;

  - STEP 2: $s_i[\cdot]$ of size 262,144 is converted into its corresponding feature vector, i.e., $f_i[\cdot]$ of size $X = 25$, where $(0 \leqslant i \leqslant X - 1)$. Particularly, $f_i[j]$, i.e., the jth component of $f_i[\cdot]$ where $(0 \leqslant j \leqslant 24)$, corresponds to the normalized energy of the jth Bark scale band of the wavelet-packet transformed input signal $s_i[\cdot]$ at the maximum decomposition level, as in Eq. 1, considering the natural frequency ordering [6], according to **Table 1**. Comments on the wavelet family used are presented ahead.

   ◦ END.

$$\frac{log\,(input\_size)}{log\,(2)} = max\_level \qquad\qquad (1)$$

We apply here the technique described in [8], which is based on paraconsistent logic, to analyze the behavior and suitability from the obtained data, i.e., the feature

Bark[a]	Band range (Hz Hz)	Initial WPT sample	Final WPT sample	Energy range[b]
0	0–100	0	$\lfloor\frac{100}{0.0841}\rfloor = 1189$	0–99.9949
1	100–200	1190	$\lfloor\frac{200}{0.0841}\rfloor = 2378$	100.0790–199.9898
2	200–300	2379	$\lfloor\frac{300}{0.0841}\rfloor = 3567$	200.0739–299.9847
3	300–400	3568	$\lfloor\frac{400}{0.0841}\rfloor = 4756$	300.0688–399.9796
4	400–510	4757	$\lfloor\frac{510}{0.0841}\rfloor = 6064$	400.0637–509.9824
5	510–630	6065	$\lfloor\frac{630}{0.0841}\rfloor = 7491$	510.0605–629.9931
6	630–770	7492	$\lfloor\frac{770}{0.0841}\rfloor = 9155$	630.0772–769.9355
7	770–920	9156	$\lfloor\frac{920}{0.0841}\rfloor = 10939$	770.0196–919.9699
8	920–1080	10,940	$\lfloor\frac{1080}{0.0841}\rfloor = 12841$	920.054–1079.9281
9	1080–1270	12,842	$\lfloor\frac{1270}{0.0841}\rfloor = 15101$	1080.0122–1269.9941
10	1270–1480	15,102	$\lfloor\frac{1480}{0.0841}\rfloor = 17598$	1270.0782–1479.9918
11	1480–1720	17,599	$\lfloor\frac{1720}{0.0841}\rfloor = 20451$	1480.0759–1719.9291
12	1720–2000	20,452	$\lfloor\frac{2000}{0.0841}\rfloor = 23781$	1720.0132–1999.9821
13	2000–2320	23,782	$\lfloor\frac{2320}{0.0841}\rfloor = 27586$	2000.0662–2319.9826
14	2320–2700	27,587	$\lfloor\frac{2700}{0.0841}\rfloor = 32104$	2320.0667–2699.9464
15	2700–3150	32,105	$\lfloor\frac{3150}{0.0841}\rfloor = 37455$	2700.0305–3149.9655
16	3150–3700	37,456	$\lfloor\frac{3700}{0.0841}\rfloor = 43995$	3150.0496–3699.9795
17	3700–4400	43,996	$\lfloor\frac{4400}{0.0841}\rfloor = 52318$	3700.0636–4399.9438
18	4400–5300	52,319	$\lfloor\frac{5300}{0.0841}\rfloor = 63020$	4400.0279–5299.9820
19	5300–6400	63,021	$\lfloor\frac{6400}{0.0841}\rfloor = 76099$	5300.0661–6399.9259
20	6400–7700	76,100	$\lfloor\frac{7700}{0.0841}\rfloor = 91557$	6400.0100–7699.9437
21	7700–9500	91,558	$\lfloor\frac{9500}{0.0841}\rfloor = 112960$	7700.0278–9499.9360
22	9500–12,000	112,961	$\lfloor\frac{12000}{0.0841}\rfloor = 142687$	9500.0201–11999.9767
23	12,000–15,500	142,688	$\lfloor\frac{15500}{0.0841}\rfloor = 184304$	12000.0608–15499.9664
24	15,500–22,050	184,305	$\lfloor\frac{22050}{0.0841}\rfloor = 262143$	15500.0505–22050.0000

[a]*Bark band.*

[b]*Range in f_i[Bark band] Hz.*

At that level, which is the maximum, and considering the original signal sampling rate of 44,100 samples/s, each sample of the transformed signal has a resolution of $\frac{44,100}{2^{18+1}} = 0.0841$ Hz. The energy of each one of the 25 sets is calculated separately and, then, normalized based on its division by the total energy. Thus, $f_i[j]$ is the jth normalized energy, $(0 \leqslant j \leqslant 24)$, for a certain input signal $s_i[\cdot]$. The symbol $\lfloor\cdot\rfloor$ in the fourth column of the table represents a rounding floor operation. It is required because a sample is obviously always an integer number. Due to that rounding, the frequency range we obtained from the WPT tree is only an approximation to the Bark scale; however, it does not make any difference in practice.

Table 1.
The 25 sets of samples, from the 18-th level WPT of size 262,144, is used to mimic the bark scale.

vector. Thus, the data vectors were represented in the paraconsistent plane as point P, according to **Figure 2**. Ideally, the closer the P is from corner $(G_1, G_2) = (0, 1)$, the better our feature vectors separate between the classes, disregarding any specific classifier. The next step was to choose the best wavelet family, that is, the family that puts P closer to that corner.
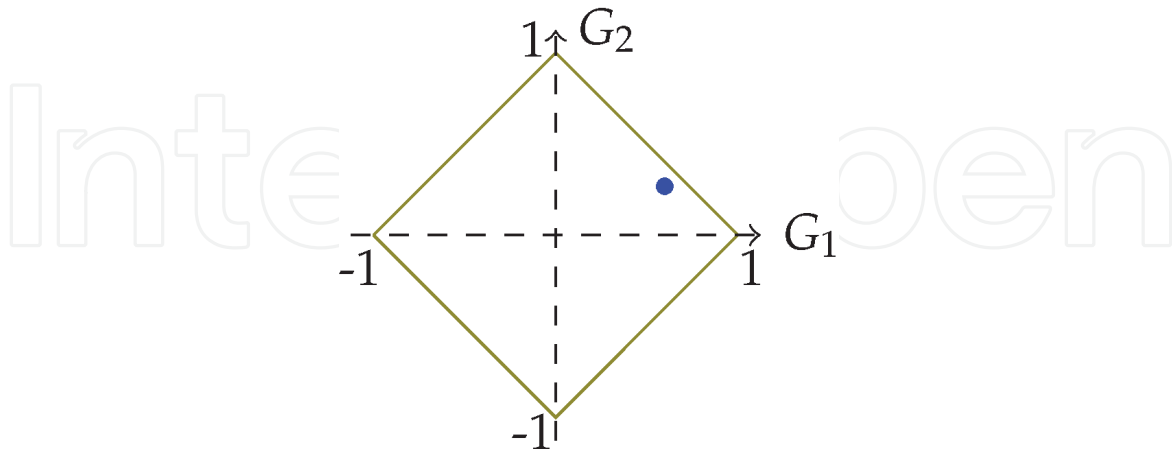

Figure 2.
The paraconsistent plane where the axes G_1 and G_2 represent the degrees of certainty and contradiction, respectively. $P = (G_1, G_2) = (\alpha - \beta, \alpha + \beta - 1)$, drawn in blue just to exemplify, is an important element for our analysis: The closer it is to the corner $(1,0)$, the weaker the classifier associated with the features vector can be. The values of α and β are derived from intra-class and inter-class analyses, respectively, as detailed in [8].
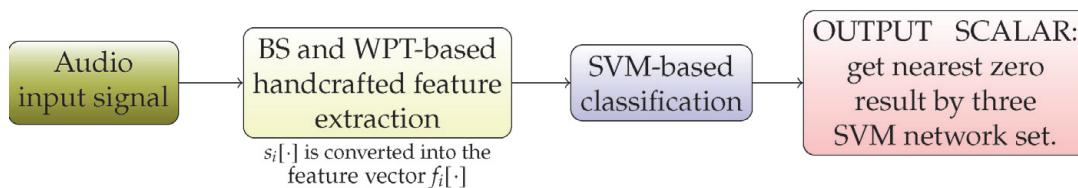

Figure 3.
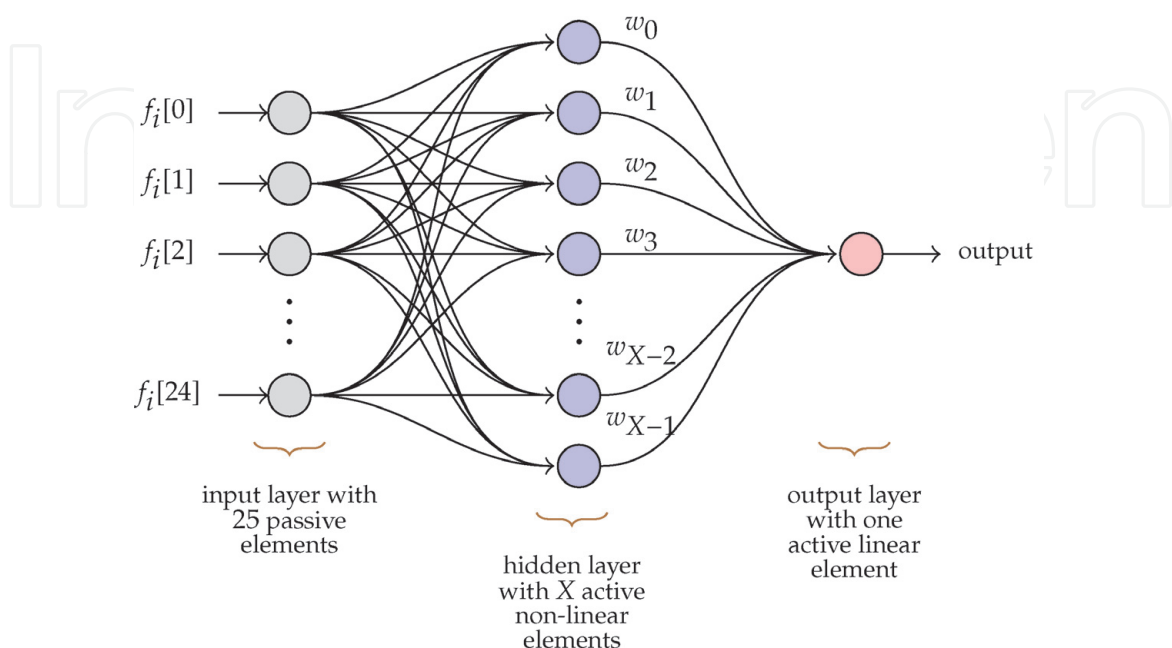*The experimental setup for the proposed application one.*

Figure 4.
The SVM structure used in application one approach. The weights determined during the supervised part of the training are $\{w_0, w_1, ..., w_{X-1}\}$. The output element linearly combines the outputs of the hidden layer with the weights.

Proceeding, we estimate that an SVM could be a proper classifier to interpret the feature vectors we selected during the preprocessing step, because of its excellence in terms of binary classifications [5]. Accordingly, **Figure 3** shows the complete setup for the proposed approach. It is divided into two phases, i.e., training and testing, with four steps each. As discussed ahead, a total of X and Y vectors were isolated to carry out each phase, respectively, where $X + Y$ corresponds to the number of acoustic files in the database.

The SVM has been implemented, as described in [5], in such a way that it receives the input vectors defined in the preprocessing step. In **Figure 3**, we illustrate the proposed setup, which is divided into two phases, training and testing, with four steps each. As we will discuss later, a total of X and Y vectors have been isolated to perform each phase, respectively, where $X + Y$ corresponds to the number of acoustic files in the database.

The detailed procedures are as follows:

- AP1 TRAINING PHASE:

    ○ BEGINNING

    ○ All the X vectors $f_i[\cdot]$ were used to train an SVM with 25 input passive elements, X hidden active non-linear elements and one output active linear element, as in **Figure 4**. X elements were used in the hidden layer to allow for a simple and effective training scheme, as explained in [5]. The kth element in the hidden layer uses a function of the form

    $$e^{-\text{Euclidian\_Distance}\left(k-\text{th\_training\_vector},\text{input\_vector\_under\_analysis}\right)},$$

    ○ implying that the kth element outputs 1 for the kth training vector and a value in the range (0–1) for the others, where $(0 \leqslant k \leqslant X - 1)$. This corresponds to a non-supervised procedure. There is no weight between the input and the hidden layers, however, there are X between the hidden and the output layer. To find them, a linear system of X equations in X unknowns is established and solved, implying in a supervised task. In that system, the closest resultant value from the SVM set corresponds the answer.

    ○ END.

Once the training procedures are over, the system is ready for testing, as follows.

- AP1 TESTING PHASE:

    ○ BEGINNING.

    ○ Each one of the Y testing vectors of size 25 are passed through the trained SVM and the corresponding output is verified: SVM with result closest to zero will be elected;

    ○ END.

4. Application two (AP2): cicada density estimation by image processing

Visually, cicadas are quite noticeable in the farming environment, so a management hypothesis would be the inclusion of a camera to permit visual detection of pests, adjusting the data capture interval for sending to a web server.

The tests were performed by using images captured in the coffee crop and divided into three classes that represent the pest insidious density: high, medium, and none. The submitted images have been converted to grayscale and have a size of 320×240 pixels, as in **Figure 5**.

Similar to AP1, an SVM is used to classify the preprocessed input energies; however, in this case we use the N normalized energies of each wavelet sub-band, instead of 25 Bark bands, according to the decomposition level defined in each test instance. In **Figure 6**, we present a flowchart that illustrates the process of capturing and processing the images.

In **Figure 7**, we have the SVM illustration with $n = 4^{level}$ entries, corresponding to the number of energies of the selected level.

Figure 5.
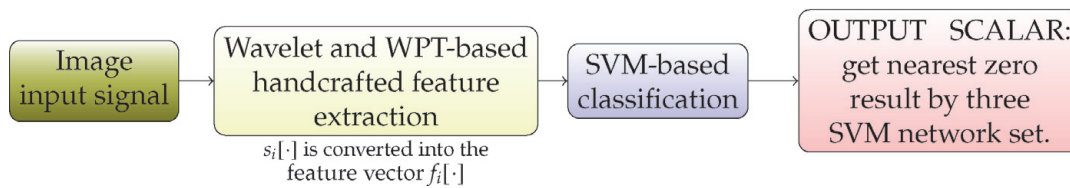*Examples of images used: high, low, and zero density, respectively.*

Figure 6.
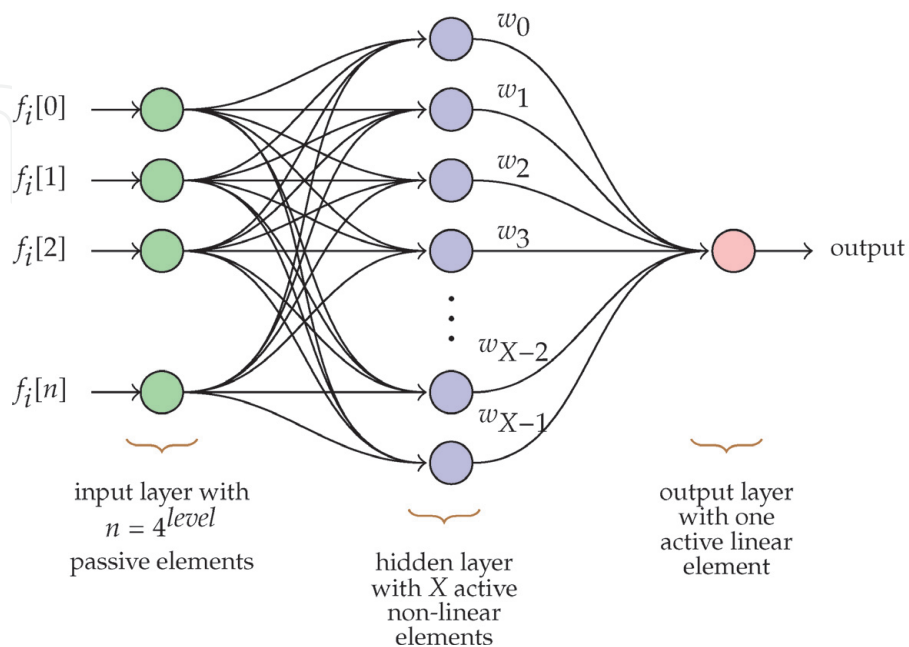*The experimental setup for the proposed application two.*

Figure 7.
The SVM structure used in application two. Similar to AP1, the weights determined during the supervised part of the training are $\{w_0, w_1, \ldots, w_{X-1}\}$. The output element linearly combines the outputs of the hidden layer with the weights.

5. Tests and results

The implementation of the algorithms proposed here were performed using Java programming language.

Wavelet	Train files/class	Test files	Percentage
DAUB6	2	99	94.95/100
DAUB8	3	96	96.88/100
DAUB4	4	93	96.77/100
DAUB4	5	90	96.67/100

Table 2.
Test results from AP1.
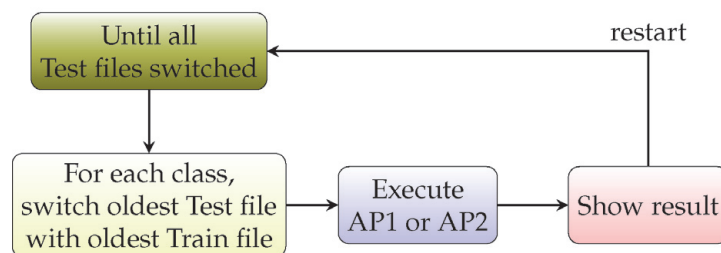
Figure 8.
Cross-validation algorithm.

Wavelet	Level	Train files/class	Test files	Percentage
COIF6	1	2	99	68.69/100
COIF6	1	3	96	81.25/100
DAUB68	1	4	93	81.72/100
DAUB32	1	5	90	83.33/100
DAUB12	2	2	99	86.87/100
DAUB52	2	3	96	87.50/100
DAUB16	2	4	93	87.10/100
DAUB38	2	5	90	87.78/100
DAUB18	3	2	99	86.87/100
DAUB68	3	3	96	89.58/100
DAUB24	3	4	93	90.32/100
DAUB16	3	5	90	90.00/100
COIF24	4	2	99	90.91/100
DAUB24	4	3	96	92.71/100
COIF6	4	4	93	92.47/100
DAUB22	4	5	90	92.22/100

Table 3.
Test results from AP2.

5.1 AP1

Thirty-five files from each class were collected and used. The tests were performed with those files not used for training, where two to five training files were experimented. For each set of training files, the DWT maximum level and mean decomposition level were tested with each of the 46 wavelet filters presented in **Table 2**. The best result, i.e., 96.88% accuracy, was obtained with Haar filter, showing excellent results and confirming our hypothesis of viability of using this system in coffee crop.

The cross-validation procedure was performed to present the best result in **Table 2**. The algorithm developed for cross-validation is illustrated in **Figure 8**.

5.2 AP2

To develop AP2, an application that uses digital image processing to estimate cicada density in a coffee crop was adopted, where 35 high-density class files, 35 low-density class files, and 35 files considered by an expert as zero density were used. In **Tables 3** and **4**, we present the results of the tests performed in the laboratory, which demonstrate the viability of future implantation in that system.

6. Conclusions

Both systems are being implemented in hardware for real-time coffee crop deployment using ESP8266 devices and their derivatives, integrating with cloud server for storing and organizing data to aid farmer decision-making. In future work, we must present practical results of their implementations.

Both AP1 and AP2 are systems that can be used as an additional form of coffee crop pest control and management. However, one possibility to be studied by the present research group is the combined use of the modalities in a single system in order to obtain even more improved results.

An important feature that was used for the laboratory experiments was the Java [14] Serialization class, which allowed for the wavelet transform in both AP1 and AP2 to be performed only once, storing its result on disk, making it possible to recover its value in repetitive cross-validation testing, substantially reducing equipment processing time.

	High density	Low density	Null density	Total
High density	30	3	0	33
Low density	1	31	1	33
Null density	0	0	33	33
Total	31	34	34	99

Table 4.
Best confusion matrix from AP2.

Author details

João Paulo Lemos Escola[1*†], Rodrigo Capobianco Guido[2†],
Alexandre Moraes Cardoso[1†], Douglas Henrique Bottura Maccagnan[3†],
João Marcelo Ribeiro[4†] and José Ricardo Ferreira Cardoso[1†]

1 Instituto Federal de São Paulo, Barretos, SP, Brazil

2 Instituto de Biociencias, Letras e Ciências Exatas, Univ. Estadual Paulista
(UNESP) (São Paulo State University), SP, Brazil

3 Universidade Estadual de Goiás, Novo Horizonte II, Iporá, GO, Brazil

4 Instituto Federal do Sul de Minas Gerais, Muzambinho, MG, Brazil

*Address all correspondence to: jpescola@ifsp.edu.br

† These authors contributed equally.

IntechOpen

References

[1] Mankin RW et al. Perspective and promise: A century of insect acoustic detection and monitoring. American Entomologist. 2011;**57**(1):30-44

[2] Martinelli NM, Zucchi RA. Cicadas (Hemiptera: Cicadidae: Tibicinidae) associated with coffee: Distribution, hosts and key to species (in Portuguese). Anais da Sociedade Entomolgica do Brasil. 1997:133-143

[3] De Souza JC. Coffee Cicada in Minas Gerais: Historical, Reconnaissance, Biology, Damage and Control (in Portuguese). Belo Horizonte: EPAMIG; 2007

[4] Maccagnan DHB. Cicada (Hemiptera: Cicadidae): Emergence, Acoustic Behavior and Sound Trap Development (in Portuguese) [Tese de Doutorado. PhD thesis]. Faculdade de Filosofia, Ciências e Letras da Universidade de São Paulo; 2008

[5] Haykin S. Neural Networks and Learning Machines. HAYKIN, Simon. 3/E. India: Pearson Education; 2010

[6] Guido RC. Effectively interpreting discrete wavelet transformed signals. IEEE Signal Processing Magazine. 2017; **34**(3):89-100

[7] Bossi M, Goldberg E. Introduction to Digital Audio Coding and Standards. Springer Science & Business Media; 2012

[8] Guido RC. Paraconsistent feature engineering. IEEE Signal Processing Magazine. 2019;**36**(1):154-158

[9] Dawson DK, Efford MG. Bird population density estimated from acoustic signals. Journal of Applied Ecology. 2009;**46**(6): 1201-1209

[10] Eliopoulos PA, Potamitis I, Kontodimas DC. Estimation of population density of stored grain pests via bioacoustic detection. Crop Protection. 2016;**85**:71-78

[11] Eliopoulos PA et al. Detection of adult beetles inside the stored wheat mass based on their acoustic emissions. Journal of Economic Entomology. 2015; **108**(6):2808-2814

[12] Marques TA et al. Estimating animal population density using passive acoustics. Biological Reviews. 2013; **88**(2):287-309

[13] Gardiner T, Hill J. A comparison of three sampling techniques used to estimate the population density and assemblage diversity of Orthoptera. Journal of Orthoptera Research. 2006: 45-51

[14] Haverlock K. Object serialization, Java, and C++. Dr. Dobb's Journal: Software Tools for the Professional Programmer. 1998;**23**(8):32-35

[15] Langer F et al. Geometrical stem detection from image data for precision agriculture. 2018. arXiv preprint arXiv: 1812.05415

[16] Burgos-Artizzu XP et al. Real-time image processing for crop/weed discrimination in maize fields. Computers and Electronics in Agriculture. 2011;**75**(2):337-346

[17] Li Y et al. In-field cotton detection via region-based semantic image segmentation. Computers and Electronics in Agriculture. 2016;**127**: 475-486

[18] Burgos-Artizzu XP et al. Improving weed pressure assessment using digital images from an experience-based reasoning approach. Computers and Electronics in Agriculture. 2009;**65**(2): 176-185