TreeVibes: Modern tools for global monitoring of trees against borers

Iraklis Rigakis,[1] Ilyas Potamitis,[2] Nicolaos-Alexandros Tatlas,[1b] Stelios Potirakis,[1c] and Stavros Ntalampiras[3]

[1] *Department of Electrical and Electronics Engineering, University of West Attica, Athens, 12244, Greece.*
[2] *Department of Music Technology & Acoustics, Hellenic Mediterranean University, Rethymno, 74100, Greece*
[3] *Department of Computer Science, University of Milan, Milan, 20133, Italy*

Key message
Remote, automatic monitoring of probed trees against borers at global scales is currently technologically feasible. Vibroacoustic recorders transmit short clips on cloud services where deep learning networks screen that data.

Abstract. Is there a wood-feeding insect inside a tree or wooden structure? We investigate several ways on how deep learning approaches can massively scan recordings of vibrations stemming from probed trees to infer their infestation state with wood-boring insects that feed and move inside wood. The recordings come from remotely controlled devices that sample the internal soundscape of trees on a 24/7 basis and wirelessly transmit brief recordings of the registered vibrations to a cloud server. We discuss the different sources of vibrations that can be picked up from trees in urban environments and how deep learning methods can focus on those originating from borers. Our goal is to match the problem of the accelerated—due to global trade and climate change—establishment of invasive xylophagus insects by increasing the capacity of inspection agencies. We aim at introducing permanent, cost-effective, automatic monitoring of trees based on deep learning techniques, in commodity entry point as well as in wild, urban and cultivated areas in order to effect large-scale, sustainable pest-risk analysis and management of wood boring insects such as those from the *Cerambycidae* family (longhorn beetles).

1

I.   INTRODUCTION

The establishment of pest populations outside their native ranges is facilitated by climatic change and global trade [1]. Larvae of non-native invasive species breed in wooden pallets and are accidentally transported through ports, handling facilities or truck roads in places that are not biologically adapted to regulate their multiplication (i.e. low plant resistance and absence of natural enemies) [2].

In Europe alone, wood boring species of the coleopteran family *Cerambycidae* (longhorn beetles) native at various parts of Asia, are now considered established or establishing in Spain, Italy, Austria, Germany, Finland, France, Montenegro, Switzerland and Turkey. Longhorn beetles are attacking at least 140 different tree species including citrus and stone fruits(peach, nectarines, plums, cherries and apricots) as well as forest woodland (infested trees become unsuitable for pulp and wood exploitation). Stone fruit production and production of roundwood for industrial uses is estimated to several billions, and, therefore, the cost of damages to these products is significant as well as the cost of eradication efforts and control of these invading insects [1]. Cerambycid female borers lay their eggs under the bark or in physical cavities or wounds/cracks of their host trees. Newly hatched borers chew directly into the heartwood. Chewing under the bark in living wood severely damages water- and sap-conducting tissues. Adults emerge from infested trees in spring or summer after penetrating the bark, therefore causing an additional problem: the exiting tunnels become the entry points for several plant pathogens. The repeated tunneling from many borers, over many generations, gradually weakens the health of the tree, causes structural instability (wind breakage), drop of fruits and leads to the decline and eventual death of susceptible trees.

There is a wide bibliography [3] on optical [4-5], laser vibrometry [6], piezoelectric sensors, and accelerometers [7-10], used to detect locomotion and feeding sound of larvae or adult pests inside the tree trunk. A mildly intrusive and widely applied method for inspection of commodities is based on inserting a piezoelectric probe in the tree trunk to listen for potential internal audio activity due to feeding and locomotion (i.e. passive acoustic detection) [11-18]. The feeding activity is audible [12-17] in two of the biological stages of the insect a) when the mature larvae tunnel into the sapwood or heartwood to form a pupal chamber (February-April), and b) after their exoskeleton is fully hardened and the adults dig emergence tunnels through the bark to exit the trunk (late spring to summer). The larva activity has lower audio imprint than adult activity whereas during the egg and pupation stages in the pupal chamber the pest is inaudible. In brief, the benefit of a piezoelectric probe is that it is portable and practical, has lower cost than competitive methods (e.g. vibrometry), is much more sensitive than microphones, does not require mains supply, calls for minor training and there are commercial products available for practitioners.

Currently, all detection methods are manually applied. A trained technician must examine and decide *in situ* on the state of the infestation. The current manual approach has several shortcomings:

a) Field visits and frequent manual inspection of trees and plants is costly, cumbersome and impractical to be scaled to large numbers of trees.

b) The listener has limited time to inspect a single tree and the larvae could be present but inactive during the inspection time for several reasons: for example, the pest may happen to be in an inaudible biological stage (e.g. egg or pupa) but will evolve in the short run, or the trunk may have low infestation load and the pest may not be chewing during the inspection's time-slot.

We have shown in [18] that a piezoelectric device can record and transmit the vibrations picked by a probe inserted in a tree. The emphasis of this work is not on hardware implementation but on the nature of this particular vibrational signal and its classification in the presence of other vibrational interferences commonly existing in the field. We introduce fast, automatic screening of vibrational records based on deep-learning models looking at the spectrogram of the internal vibrations that can

extend for weeks to months before reaching a decision on the infestation state of the tree (see Fig.1 for a depiction of the main idea). We have used a commercial version of a piezoelectric device and recorded several thousand transmitted recordings extending over a period of six months mostly in urban environments. We train various deep learning approaches each one having its own merits. The database (the first of its kind for this kind of problem) is available for download at http://www.kaggle.com/potamitis/treevibes along with the associated deep learning code. Our vision is ambitious: remote, automatic surveillance of trees against borers at global scales based on deep nets.

In this work, we demonstrate that this vision is technologically feasible; it creates services currently inexistent and is hampered only by the current cost of materials that can only drop in the future. The structure of this work is as follows: We first examine the signal of wood-boring insects based on the example of *Xylotrechus chinensis* (Coleoptera: *Cerambycidae* - Tiger longicorn beetle), an Asian woodborer, causing high mortality of *Morus* trees (mulberries) in Greece. In the context of this work, audio is based on the vibrations caused by the pests cracking the tree fibres. It is possible that elsewhere, different types of wood and different borers produce sounds with different spectral content but extended literature of the field (see [12-24] and the references therein) show that acoustic emission cannot be avoided. We then examine the vibrational soundscape of trees in urban spaces and forests and analyse the practical benefit of automatic remote surveillance of trees against borers. Subsequently, we describe deep learning techniques as applied to the spectrogram of vibrations originating from piezoelectric probes inserted in tree trunks. Finally, we conclude on future prospects especially on how our approach can be connected to the internet of things (IoT) reaching global scales.
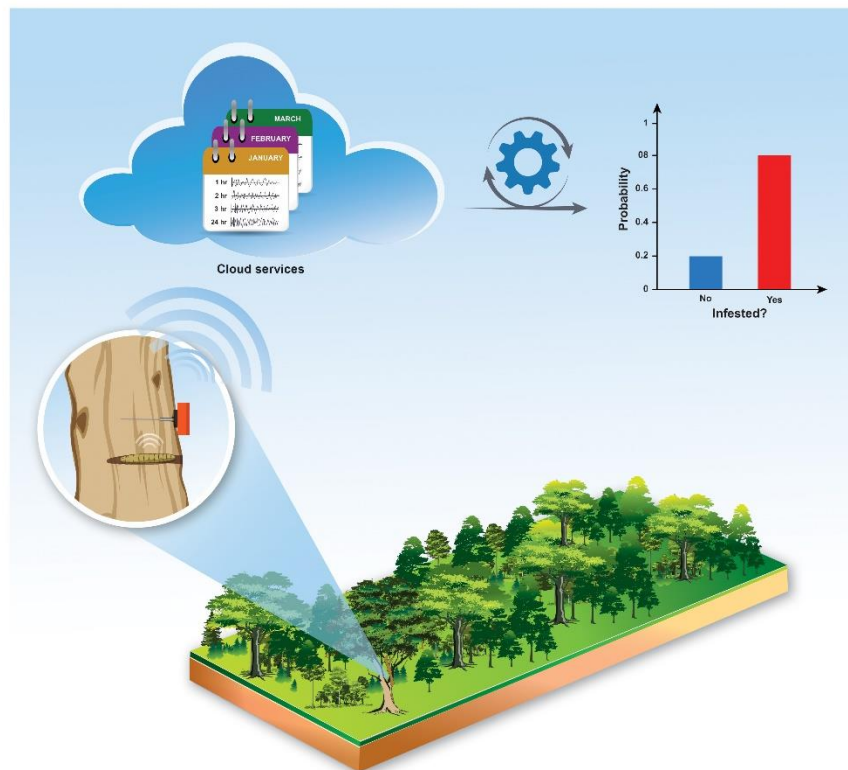

Figure 1. The concept of screening massively transmitted snippets of vibrations stemming from inside the tree due to the feeding or moving sounds of larvae.

II.   MATERIALS AND METHODS

In this section we start with basic principles of vibrational recorders and the nature of the signal recorded under different environmental conditions.

The device

The core of the sensor we used for listening to vibrations caused by borers is the piezoelectric crystal. This is an electromechanical system (the crystal and an embedded amplifier) that reacts to compression by converting it to a fluctuation of an electrical charge. Therefore, it is closer to the concept of a seismometer than that of a microphone. In the context of our application, compression is inflicted by any vibration inside the wood, while the electrical fluctuation can easily be converted to an audio signal that can be stored, compressed and transmitted. A metal waveguide (see Fig. 2-left) is a metal bar, functioning as a sound coupler between the wood and the sensor probe.

The circuit is constantly in sleep mode, wakes up on a predefined time schedule (e.g. 20 sec every hour) and takes a recording before going to sleep again. The recording duration and the density of the sampling is configurable through the reporting server. This means that there is a bidirectional wireless communication between the deployed devices and the reporting cloud server. The recordings are stored in the SD card, and the time-stamp is passed to the filename. All audio recordings are compressed using the open source opus compressor prior to sending them over the communication channel (see Fig. 2-right for a field application). The bit rate is 24KBPS at a sampling frequency of 8 kHz. The device uses a global SIM card, therefore any tree can be tracked from anywhere in the world. There is no need to recharge the device as it has an embedded solar panel that provides enough power for its low-power electronics. Therefore, it can stay on a tree for an indefinite time-period, sampling and transmitting the internal vibrations of the tree. The location of the device appears on the world map of the server as the device carries a global positioning system decoder (GPS). All data are communicated through the mobile network. Further details of a proof of concept of this approach can be found in [18].

Figure 2. Left: The device is a recorder of vibrations that are picked up by the metal probe. Right: The device attached to a mulberry listening for the cerambycid pest *Xylotrechus chinensis*.

The signal

Back in 2017, in the island of Crete, Greece, and in Spain, it was observed [25-26] that several mulberries appeared to bear exit tunnels (see Fig. 3-left) that had not been observed before. Suspicious trunks were sliced and the larvae found were subjected to polymerase chain reaction

(PCR) analysis that showed that they belonged to the invasive cerambycid *X. chinensis* (see Fig. 3-right). We took several trunks like in Fig. 3-left to the laboratory, where we made several recordings using a multichannel recorder (Tascam DR-680MKII) (Fig. 4).

Figure 3. Left: Exit tunnels of *X. chinensis* in a mulberry in Crete, Greece. Right: Adult *X. chinensis*.

Figure 4. Recording in the lab with a multi-channel recorder.

In Fig. 5 (top) one can see a typical example of these recordings. Generally, the internal soundscape of a healthy tree —excluding externally induced vibrations — is silent at the level of audio sounds we seek. If it is infested one expects to hear a train of pulses like in Fig. 5-middle and the rate of insect bursts can be used to estimate the likelihood that the tree is infested [27]. The train consists of a number of bursts, each one corresponding to a crack of fibers as the borers feed and move (see Fig. 5-bottom for a single burst). We can confidently attribute these impulses to *X. chinensis* because

the recordings have been taken in the controlled environment of the lab, the adults have emerged some months after the recording and the trunk has been subsequently sliced and examined for other possible insects. Looking at Fig. 5, one may suppose that the detection of borers is an easy task: an envelope follower or a simple thresholding could reveal the impulses. This is not the case as field recordings can be more complex than laboratory recordings.
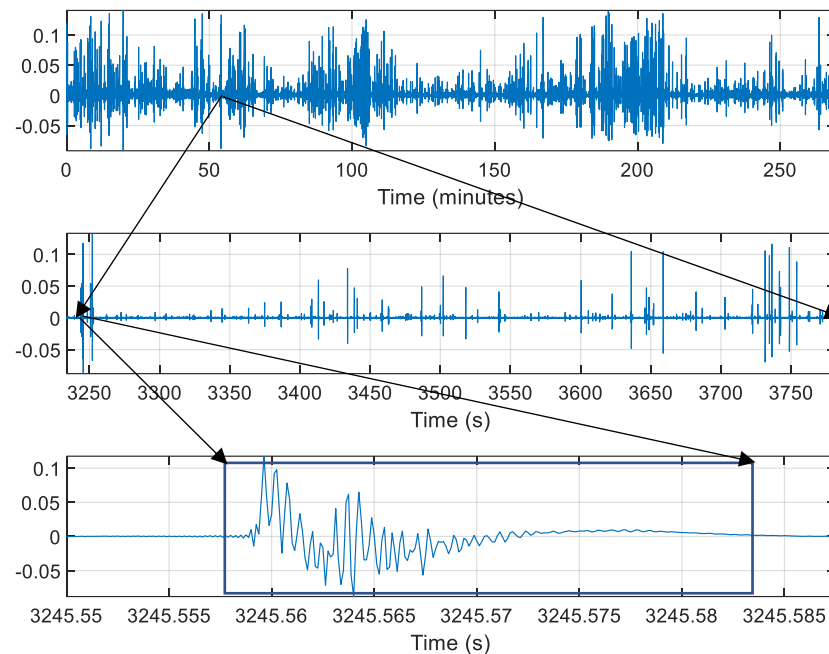

Figure 5. Looking a recording of an infested mulberry with *X. chinensis* at different time-scales. Top: A 4h 30 min continuous recording. Middle: A pulse train between 53m 50sec-63m 20sec. Bottom: A single burst corresponding to a crack of fiber.

A soundscape is a combination of sounds that arises from an environment. It refers to both the natural acoustic environment (animal vocalizations, weather sounds, rain) and sounds created by humans (traffic sounds, corns, footsteps, vocalizations). One may expect that the internal soundscape of a tree in the field is quiet and dull. However, it is not. With the term 'internal', we mean everything that a recording element located inside the tree would register. In the context of this work, we are interested only in sounds of borers but these must be discerned against any other possible forms of vibration.

a) As mentioned in the introduction, depending on the biological cycle of the pest, it can be noisy or cryptic. Therefore, snippets taken from trees in urban spaces can be rich in vibrations originating from traffic, footsteps, vocalizations of dogs, birds and humans, shaking of the branches and leaves due to the wind and uncountable other unpredictable audio sources. Some of these vibrations propagate in the wood and reach the metal probe of the device. Therefore, recordings can be very noisy sometimes to the point that external noise dominates over the impulsive sound of the borer.

b) One does not know if there are borers in the tree and even if there are, one cannot know their number and location inside the tree. Some of these impulses are feeble because they originate from a location distant to the probe. Depending on the kind of the wood, the probe can detect feeding sounds within a sphere of 1.5-2m radius.

6

In Fig. 6 we gather characteristic examples of biophony, anthropophony and geophony taken from the transmitted field recordings of the TreeVibes database. In each sub-figure the top figure corresponds to the time-domain signal and the bottom to its spectrogram (i.e. the Short-Time Fourier Transform which is a representation of the change over time of the frequency composition of the signal). The sampling rate is 8 kHz and we use a hamming window of 512 samples with 50% overlap. Fig. 6a is taken from a young pine tree (not a host of *X. chinensis*) in a forest with no signs of wounds or degradation. The recording is quiet with some distant bird chirps mainly seen in the spectrogram near 4 kHz. This is a typical recording of a healthy tree with a quiet background (usually at nights). Fig. 6b is taken from a mulberry with severe visual signs of infestation seen also as vertical strips in the spectrogram, indicative of impulsive audio events. The recording was taken in summer; therefore, it is most probably an adult digging his/her way out. The tree is located near a busy street. At 4-5 sec there are human vocalizations whereas from 12-18sec a passing-by car that vibrates the tree. The impulses of adults digging their tunnel out in summer are much stronger than the sound of larvae in the beginning of the year. Yet, both sounds are clearly audible. Fig. 6c is an infested tree but the bird vocalizations are very strong. Fig 6d is a healthy apricot tree. The recording is taken under heavy wind and rain. All impulses are due to weather conditions and shaking of branches and leaves that result into vibrations. Healthy trees in calm weather may register occasional impulses (but not trains of impulses) that are due to tree metabolism related to humidity levels and dilations. Borers create a characteristic repeated pattern in the form of a pulse train and not isolated events.
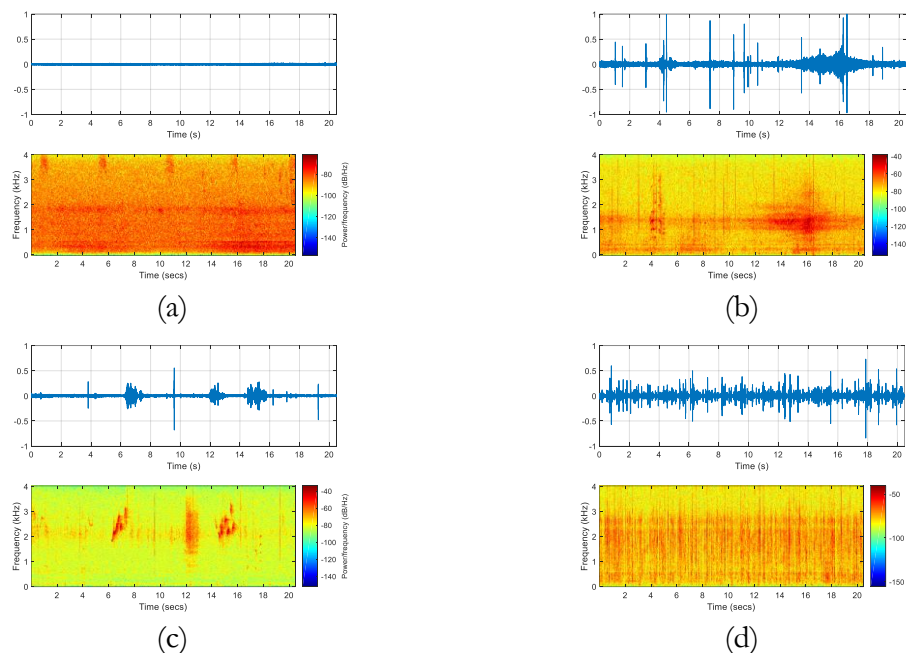

Figure 6. Transmitted vibrational recordings from different trees under different circumstances. a) Healthy young pine in a calm day. No train of impulses, b) infested tree in alley near heavy traffic. Impulse trains in the presence of human vocalizations and traffic, c) An infested tree in the presence of strong birds' vocalizations, d) healthy, young, apricot tree. Recording taken during a heavy shower. All impulses are due to rain, and shaking branches/leaves due to strong wind.

In Fig. 7, we compare the spectral profiles of long duration recordings taken from an infested and a non-infested trunk. A different borer in a different tree could create acoustic emissions with a different spectral profile; nevertheless, it would not be flat like the non-infested one. The power spectral density (PSD) one-sided estimate, of each recording sampled at 8 kHz is found using

Welch's overlapped segment averaging estimator. To elaborate further, the signal is divided into sections of length of 512 samples. The modified periodograms are computed using a Hamming window of the same length as the window. The overlapping in windowing equals to the 50% of the window length.
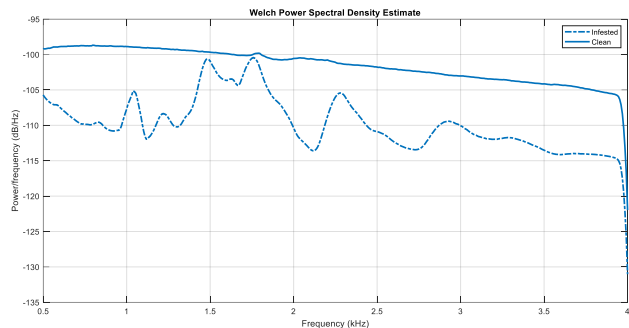

Figure 7. Spectral profile of an infested vs a non-infested trunk through the power spectral density estimate (Welch method). The duration of both recordings is longer than 1h 30 min. The non-infested trunk is flat indicating practically, a white noise profile. The averaged spectrum of the infested trunk contains the spectral profile of multiple impulses. The variations in the relative magnitude of signal energy at different frequencies indicate that it is an infested trunk.

Figure 8. A case of highly infested tree as indicated by the recordings but without any visible signs of infestation (picture taken March 2020, at 35º19'54.3"N 25º07'54.9"E).

The new services are gathered in Table 1.

Automatic interception of infested trees and timber cargos at commodity entry-points
Integration of information from larger time spans (weeks to months)
Replacement of all decisions based on late visual assessment of trees with early vibrational detection
Provision of permanent time-stamped evidence (recordings) interpretable even by non-specialists

| Decentralization of the problem of decision making by bringing the knowledge of specialists to remote areas |
| Delimitation of infestation areas that can reduce use of pesticides and infestation rate by removing infested trees at early stage |

Table 1. A list of benefits using automatic screening of trees' vibrational records.

III.   RESULTS

The practical value of knowledge

Phytosanitary interception at commodity entry points (e.g. airports, harbors, stations, lorry parks, cargo depots and quarantine facilities) is the first line of defense against invasive insects [1-3]. Wooden pallets, wood products, ornamental trees, plants but also cargos of fruits and other agricultural products are typically examined before importation using visual inspection and various technological means [10, 13]. Effective interception of potential pests including but not limited to quarantine species already intercepted in the past, is crucial [3]. Though not impossible, it is increasingly difficult to achieve eradication of establishing or established invasive species after initial arrival. Interception is currently based on visual inspection and *manual* application of several technologies.

This work introduces the novel service of automatic screening of wood-related imports. In short, devices are attached to the trees in storage facilities, the vibrational soundscape of the trees is sampled for the whole quarantine period and clearance is provided automatically after deep learning models have finished screening the vibrational record of the shipment, otherwise a rapid cargo rejection is inflicted. Because it does not involve human attendance (one can attach the device and leave), it can be applied to a larger scale than it is currently done. In addition, since it integrates a longer time span of observations than the human service currently applied, it is anticipated that it will be more accurate.

Another service that currently does not exist is based on transmitting the systematic registration of vibrations to cloud services. The audio data serve as a permanent record of evidence and the process of cross-examination by trained bioacousticians is decentralized in the sense that the trees under investigation, the stored audio records and the human specialists need not be in the same place— pretty much as the way telemedicine is applied.

Due to current manual limitations, only 2%/year of incoming shipments is inspected in US [3]. Therefore, more often than not, invading species are not intercepted at commodity entry-points and —as an example family — *Cerambycidae* beetles are establishing in new locations. Post-border surveillance and containment is easier if the first establishment of the invasive species is detected and localized as early as possible. Forests and parks nearby commodities' entry points are most at risk. If the invasive species attack trees of urban ornamental greenery in public spaces, like in the case of *X. chinensis* for mulberries and *Rynchoforus ferrugineous* (*curculionidae*) for palms in Crete, the trees are left untreated until they die without consideration of their aesthetic value [2]. Even in such a case, the automatic screening of vibrational records from trees offers new services and introduces a possible revision of the currently applied protocol. Regarding urban spaces, workers in ornamental greenery assess visually whether the trees already have exit tunnels, discoloration/damage of leaves, signs of rotten tissue and any other visual symptoms of health decline and cut down only the ones that are heavily infested or dead. However, this is too late: visual symptoms appear 1-2 years after the first infestation as regards *cerambycidae*/*curculionidae*, which means that by the time their traces are visible, the borers have completed several generations inside the tree and have escaped to infest new ones. What we suggest is to remove the trees with positive acoustic records and not to base

9

inspection and assessment on visual records. Even if no other treatment is applied, this procedure is expected to delay the degradation of urban greenery relying on the specific tree species.

Let us give a lucid example on the dilemmas phytosanitory personnel face on a daily basis and how these can be answered with automatic screening of vibrational records. Should we cut down the mulberry in Fig. 8 (one of the 3500 mulberries in the city of Heraklion in Crete) knowing that the city is infested with *X. chinensis* and *morus* tree is the primary host? The decision to cut down trees is of grave importance both in terms of financial cost (i.e. removal and secure destruction costs) and in terms of ecological impact. Pest specialists would definitely refuse the cutting, as the tree has no symptoms of degradation and looks perfectly healthy. Yet, an examination on the upper-part of its trunk shows long and vivid pulse trains of vibrations. Again, recordings can serve as evidence and the pulse rate can assess the infestation status (heavy or low). Removing the tree will locally degrade the greenery but the alternative is to remove it dead, two-three years later, while escaping adults and their descendants will have infested a large number of healthy trees thus accelerating the degradation on a regional level. On the contrary, removing it a year and a half prior to the visual symptoms will significantly prolong ornamental greenery even if no other treatments are applied.

A different protocol may apply in the cases of trees of economic importance like orchards of stone fruits as heavy infestations lead to fruit drop. In such cases, the usual procedure is the removal and the immediate destruction of all infested trees, as well as those present within a variable radius of the infestation. The decision, however, to characterize a tree as infested is again based on visual signs. As mentioned above, this approach has poor effects because when visual signs are prominent enough to characterize a tree as infested, many generations of adult pests have already escaped. Therefore, removal of trees based on visual assessment of symptoms is not sufficient to stop the invasion to new areas, and to limit the damage where pests are already established. When borers are established, pest control may involve aerial and ground bait pesticide sprays, but their efficiency depends on knowing the time and location of insect infestations as early as possible. The advantage of probing the trees is that they can reveal the problem as early as first generation larvae and automatically tag their location (the transmitting device carries a GPS).

The database

It is quite straightforward to acquire recordings in acoustically challenging conditions (i.e. due to background interference) from trees that are not infested by borers. By simply inserting the probe in trees known to be healthy, one may easily get most of the typical sources of background vibrational interference (traffic, vocalizations, wind etc). It is more complicated, however, to get recordings from infested trees, as the ultimate way to verify infestation is to cut down the tree — which is generally illegal in public spaces — and slice it until one finds the larvae.

We gathered the recordings in two ways: a) By attaching the device on trees that had serious visual signs of attack and manually verifying the existence of pulse trains from the audio and visual inspection of spectrograms. b) From mulberries that had been cut down with permission by authorities (heavily infested trunks or dead trees).

The database is composed of 33 folders with audio recordings taken from 35 different trees and a corresponding annotation csv file. The folders contain emitted recordings over a period of six months. The recordings are in wav format but are actually decompressed after being received in an ogg format. The sampling frequency is 8 KHz. The first 27 folders are used for training and validation and the last 6 for testing.

The data set of the target insect is composed of 4165 field and 53676 laboratory recordings mostly at 20 sec. Training Folders: Infested (train pulses from borers) 1-6,11-23, #recs 731. Clean: 7-10, 24-25, 35, #recs 1754. Total training data #recs 2485. Test Folders: #26-#34.

Deep-learning as applied to spectrograms of vibro-acoustic signals
Deep learning (DL) architectures have a modular layer composition where the layers close to the input learn to extract low-level features and subsequent layers rely on the previous layer(s) to synthesize patterns of higher abstraction (e.g. starting from edges and textures and ending in objects) [28-30]. As it is impractical to listen manually to hundreds of thousands of clips transmitted to a cloud server from a large number of trees, there is need for an automatic process that screens these recordings. Deep learning techniques can provide fast classification (as rule of thumb 5 ms/recording in a single GPU), as they can discern between train pulses originating from borers and events from other external sources of vibrations (as human listeners can). We achieve that by transforming the audio recording to an image through the spectrogram (i.e. the Short-time Fourier Transform is a 2D representation like an image) and feeding the images to a DL model. In the case of the spectrogram the 'object' is a spectral blob that corresponds to a vibration source. It is important in our case to not only detect trains of pulses originating from borers but also learn to discern between impulsive events belonging to different sources, which vibrate the tree, although they are located outside it. The operational model calculates the probability of infestation of a tree based on a long history of recordings that can span weeks.

Verification Experiments
We performed 10-fold validation cross-validation on field data to estimate how different convolutional neural networks (CNNs) models are expected to perform in general when used to make predictions on data not used during training. The procedure had a single parameter k=10 referring to the number of groups that a given data corpus was to be split into. Each group, in turn, was held out as a test data set and the remaining groups made the training data set. We fitted a model on the training set and evaluated it on the test set. The accuracy over each fold was measured and the mean score over 10-folds along with the standard deviation is reported in Table 2. We applied a type of data augmentation with rolling of recordings at a random point to randomize the point in time the impulses appeared. In this work, our aim is not to fine-tune the hyper-parameters of the classifiers through grid-search. The images used to feed all CNNs are the spectrogram of the recordings using an FFT size of 256 and 50% overlap, resulting to a 129x1251 matrix.

Table 2. Mean accuracy/std.dev using a 10-fold cross validation scheme	
CLASSIFIERS (ranked by their parameters size)[†]	
ResNet50 (98 MB)	93.68/1.58
Xception (88 MB)	**94.16/0.99**
DenseNet121 (33 MB)	93.56/1.60
EfficientNetB0 (29 MB)	93.80/1.72
MobileNet (16 MB)	**93.84/1.16**
[†]Adam optimizer(learn_rate=0.001, decay=1e-6) see Appendix.	

Table 2. 10-fold cross-validation results for various deep learning models

We compared a set of state of the art deep learning models to find the best-performing model that is most generalizable, has the least loss, and is the most suitable to be embedded for the task to be performed. In Table 2 we give emphasis on models with small memory imprint (EfficientNetB0, MobileNet) with a view to embedding them in the probes instead of running them on the server level. It can be seen that, among the five models, the EfficientNetB0 and the MobileNet compare

11

favorably to the larger models, while the best scoring Xception had the best convergence and training performance.

To further elaborate on the verification accuracy we use precision, recall and F_1 score metrics on a random 20% holdout data for the best performing model. Precision (P) is defined as the number of true positives (Tp) over the number of true positives plus the number of false positives (Fp). Recall (R) is defined as the number of true positives (Tp) over the number of true positives plus the number of false negatives (Fp). These quantities are also related to the (F_1) score, which is defined as the harmonic mean of precision and recall.

$$P = \frac{Tp}{Tp + Fp}, \qquad P = \frac{Tp}{Tp + Fn}, \qquad F1 = 2\frac{P*R}{P+R}$$

High precision relates to a low false positive rate (false alarm), and high recall relates to a low false negative rate (miss). High scores for both show that the classifier is returning accurate results (high precision), as well as returning a majority of all positive results (high recall). We did not try to fine-tune classifiers through grid-search and voting schemes of different models as optimization of classifiers is not the focus of this work.

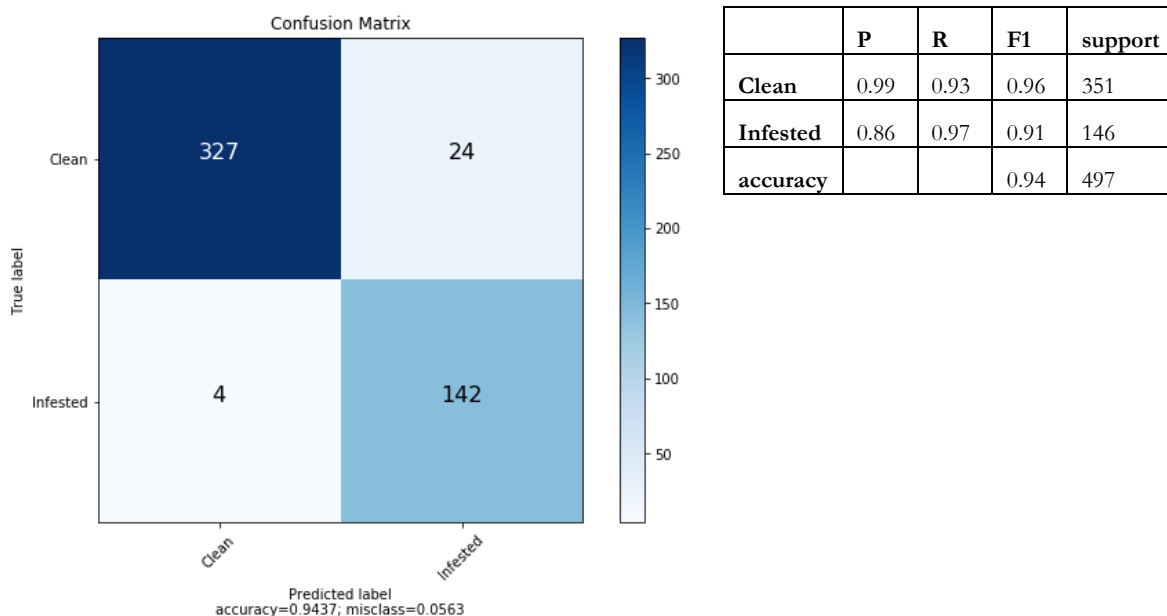|	P	R	F1	support
Clean	0.99	0.93	0.96	351
Infested	0.86	0.97	0.91	146
accuracy			0.94	497

Fig. 9. Automatic verification of borers' infestation in trees in the field data based on a 20% holdout set. (Left) Confusion matrix of a using an Xception model. (Right) Precision (P), Recall (R) and *F1* scores.

Finally, in Fig. 10 we demonstrate how automatic assessment on the infestation status of a tree takes place once the CNN is operational: the probed tree provides a folder of snippets spanning a time interval and this folder is directly fed to the trained CNN with spectrograms of vibrations being the input and probability of infestation the output. Probabilities are averaged for all snippets and normalized to unity by diving with the number of snippets.
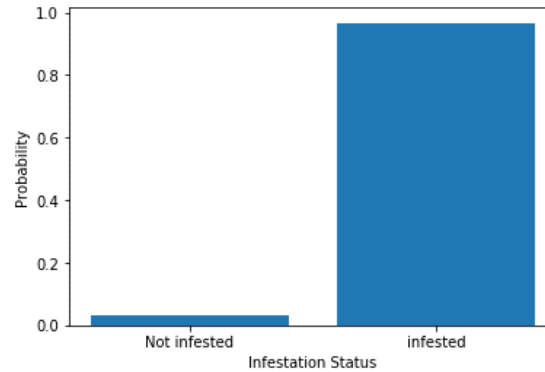
Figure 10. Infestation state of a tree after examining the folder no. 33 of the provided database containing 753 recordings spanning several days. The probability is derived by averaging the probabilities of all cases and normalizing to one.

IV.   CONCLUDING REMARKS AND FURTHER STEPS

To our point of view vibrational sensors attached to trees that have a bidirectional wireless communication with cloud-servers, hold much promise for detecting invasive wood-feeding insects in various novel applications. Borers can be detected during their larva and adult stage when they move and feed. The algorithms can automatically integrate data from long time spans (daily, weekly, monthly) to infer the infestation state of a tree. The number of nodes applied will increase as the cost of electronics decreases and technology improves when 5G wireless communication is widely adopted. Automatic screening of vibrational data can be carried out at the server allowing the efficient and rapid processing of thousands of recordings allowing novel services to emerge as automatic creation of infestation maps and predictive modelling of invasion and spread. In the era of global trade and climate change, modern tools to monitor remotely trees for borers before they colonize and establish in new habitats can lead to novel services and modernize inspection agencies. This in turn can have a significant reduction on the economic damage caused by pests and spraying costs related to treatments and increase productivity with a lower impact on the environment and human health.

APPENDIX

The full database and associated csv file and Python codes to reproduce all results of this work can be downloaded from https://www.kaggle.com/potamitis/treevibes
IMEI a unique identifier for each device, GPS_LAT, GPS_LONG stand for latitude and longitude coordinates. VISUAL stand for visual signs of degradation, AUDIO stands for the result of human listener, CONFIRMATION stands for the trunk-slicing process until larvae have been found and COMMENTS relate to some observation of the location.

FOLDER	IMEI	GPS_LAT	GPS_LONG	VISUAL	AUDIO	CONFIRMATION	COMMENT

1	867584031542538	35.3317795	25.1290474	B	1	YES	busy city center, pavement, drill bit sensor
2	867584033227260	35.3318176	25.1290665	B	1	YES	busy city center, pavement, drill bit sensor
3	867584031542538	35.3316269	25.1290836	A	1	NO	busy city center, pavement, drill bit sensor
4	867584033227260	35.3314171	25.1288147	A	1	NO	busy city center, pavement, drill bit sensor
5	867584031542538	35.3195114	25.1315899	A	1	YES	near street, not city center, drill bit sensor
6	867584033227260	35.319397	25.1314068	A	1	YES	near street, not city center, drill bit sensor
7	867584033228409	35.5079269	24.0136147	X	0	YES	orange tree, rain-wind

	VISUAL INSPECTION OF TREES
A	exit tunnels, heavy infestation
B	exit tunnels, low infestation
X	no exit tunnels, no signs of infestation

Let us see the structure of a typical filename such as F_20200222172127_4_T14.0 and what it means:

F_20200222172127_4_T14.0
2020 stands for year
02 stands for the month
22 indicates the day
17 indicates the hour
21 indicates the minute
27 indicates the second
_4_ is a file-in-folder index
T14.0 stands for 14.0°C

Author contribution statement

IR developed that electronics and partly analyzed the data. IP gathered the database and wrote the manuscript. NT designed the study and helped to analyze the data. SP proofread and revised the manuscript and supervised technical analysis. SN provided suggestions on the experimental design and analyses of deep learning techniques. All authors read and approved the manuscript.

Conflict of interest

Iraklis Rigakis is the founder of INSECTRONICS (www.insectronics.net) that fabricates the vibroacoustic recorders used in this study. The authors declare that they have no other conflict of interest.

REFERENCES

1   Aukema JE, Leung B, Kovacs K, Chivers C, Britton KO, Englin J, et al. (2011) Economic Impacts of Non-Native Forest Insects in the Continental United States. PLoS ONE 6(9): e24587. https://doi.org/10.1371/journal.pone.0024587
2   Cocquempot C, Lindelöw Å (2010) Longhorn beetles (Coleoptera, Cerambycidae). Chapter 8.1. In: Roques A et al. (Eds) Alien terrestrial arthropods of Europe. BioRisk 4(1): 193–218. doi: 10.3897/biorisk.4.56
3   Poland, Therese M.; Rassati, Davide. 2019. Improved biosecurity surveillance of non-native forest insects: a review of current methods. Journal of Pest Science. 92(1): 37-49. https://doi.org/10.1007/s10340-018-1004-y.

4   Perles, A.; Mercado, R.; Capella, J.V.; Serrano, J.J. Ultra-Low Power Optical Sensor for Xylophagous Insect Detection in Wood. Sensors 2016, 16, 1977.

5   Ashry, I., Mao, Y., Al-Fehaid, Y., Al-Shawaf, A., Al-Bagshi, M., Al-Brahim, S., Ng, T.K. & Ooi, B.S. Early detection of red palm weevil using distributed optical sensor. Scientific Reports 10, 3155 (2020).

6   Zorović, M., Čokl, A. Laser vibrometry as a diagnostic tool for detecting wood-boring beetle larvae. J Pest Sci 88, 107–112 (2015). https://doi.org/10.1007/s10340-014-0567-5

7   Francis Juanes, Visual and acoustic sensors for early detection of biological invasions: Current uses and future potential, Journal for Nature Conservation, Volume 42, Pages 7-11, 2018.

8   Mankin RW, Hagstrum DW, Smith MT, Roda AL, Kairo MTK. Perspective and Promise: A century of insect acoustic detection and monitoring. American Entomologist. 2011; 57(1):30-44.

9   Liu H, Lee SH, Chahle JS. A review of recent sensing technologies to detect invertebrates on crops. Precision. Agriculture. 2017; 18:635-666.

10   K Sindhura Bhairavi, Badal Bhattacharyya, Nang Sena Manpoong, Partha Pratim Gyanudoy Das, Elangbam Bidyarani Devi and Sudhansu Bhagawati, Recent advances in exploration of acoustic pest management: A review, Journal of Entomology and Zoology Studies 2020; 8(3): 2056-2061

11   Koubaa, A.; Aldawood, A.; Saeed, B.; Hadid, A.; Ahmed, M.; Saad, A.; Alkhouja, H.; Ammar, A.; Alkanhal, M. Smart Palm: An IoT Framework for Red Palm Weevil Early Detection. Agronomy 2020, 10, 987.

12   Alexander Sutin, Alexander Yakubovskiy,Hady R. Salloum, Timothy J. Flynn, Nikolay Sedunov, and Hannah Nadel, Towards an Automated Acoustic Detection Algorithm for Wood-Boring Beetle Larvae (Coleoptera: Cerambycidae and Buprestidae), Journal of Economic Entomology, 112(3), 2019, 1327–1336

13   Sutin, Alexander & Flynn, Timothy & Salloum, Hady & Sedunov, Nikolay & Sinelnikov, Yegor & Hull-Sanders, Helen. (2017). Vibro-acoustic methods of insect detection in agricultural shipments and wood packing materials. 1-6. 10.1109/THS.2017.7943503.

14   Mankin RW, Osbrink WL, Oi FM, Anderson JB. Acoustic detection of termite infestations in urban trees. Journal of Economic Entomology. 2002; 95(5):981-988.

15   Mankin RW, Weaver DK, Grieshop M, Larson B, Morrill W. Acoustic system for insect detection in plant stems: comparisons of Cephus cinctus in wheat and Metamasius callizona in bromeliads. Journal of Agricultural and Urban Entomology. 2004; 21(4):239-248.

16   Mankin RW, Samson PR, Chandler KJ. Acoustic detection of Melolonthine larvae in Australian sugarcane. Journal of Economic Entomology. 2009; 102(4):1523- 1535.

17   Mankin RW, Moore A. Acoustic detection of Oryctes rhinoceros (Coleoptera: Scarabaeidae: Dynastinae) and Nasutitermes luzonicus (Isoptera: Termitidae) in palm trees in urban Guam. Journal of Economic Entomology. 2010; 103:1135-1143

18   Potamitis, I.; Rigakis, I.; Tatlas, N.-A.; Potirakis, S. In-Vivo Vibroacoustic Surveillance of Trees in the Context of the IoT. Sensors, 2019, 19, 1366.

19   Pinhas J, Soroker V, Hetzroni A, Mizrach A, Teicher M, Goldberger J. Automatic acoustic detection of the red palm weevil. Computer and Electronics in Agriculture. 63(2):131-139, 2008.

20   Y. Fujii, M. Noguchi, Using acoustic emission monitoring to detect termite activity in wood, Forest Prod. J., 40(1) (1990), pp. 34-36.

21   Potamitis, T. Ganchev, D. Kontodimas, On Automatic Bioacoustic Detection of Pests: The Cases of Rhynchophorus ferrugineus and Sitophilus oryzae, Journal of Economic Entomology, Volume 102, Issue 4, 1 August 2009, Pages 1681–1690, https://doi.org/10.1603/ 029.102.0436

22   Siriwardena KAP, Fernando LCP, Nanayakkara N, Perera KFG, Kumara ADNT, Nanayakkara T. Portable acoustic device for detection of coconut palms infested by Rynchophorus ferrugineus (Coleoptera: Curculionidae). Crop Protection. 2010; 29(1):25-29.

23 Lampson BD, Han YJ, Khalilian A, Greene J, Mankin RW, Foreman EG. Automatic detection and identification of brown stink bug, Euschistus servus and southern green stink bug, Nezara viridula, using intraspecific substrateborne vibrational signals. Computer and Electron., in Agriculture. 2013; 91:154-159.

24 Mankin RW, Al-Ayedh HY, Aldryhim Y, Rohde B. Acoustic Detection of Rhynchophorus ferrugineus (Coleoptera: Dryophthoridae) and Oryctes elegans (Coleoptera: Scarabaeidae) in Phoenix dactylifera (Arecales: Arecacae) Trees and Offshoots in Saudi Arabian Orchards. Journal of Economic Entomology. 2016; 109(2):622-628

25 Leivadara E, Leivadaras I, Vontas I, Trichas A, Simoglou K, Roditakis E, Avtzis DN (2018) First record of Xylotrechus chinensis (Coleoptera, Cerambycidae) in Greece and in the EPPO region. Bulletin OEPP/EPPO Bulletin 48(2), 277–280. https://doi.org/10.1111/epp.12468

26 Sarto i Monteys, V.; Torras i Tutusaus, G. A New Alien Invasive Longhorn Beetle, Xylotrechus chinensis (Cerambycidae), Is Infesting Mulberries in Catalonia (Spain). Insects 2018, 9, 52.

27 R. W. Mankin, A. Mizrach, A. Hetzroni, S. Levsky, Y. Nakache, and V. Soroker, Temporal and spectral features of Sounds of wood-boring Beetle Larvae: Identifiable patterns of Activity enable improved discrimination from background noise, Florida Entomol.2008, 91 (2), 241-248.

28 LeCun, Y., Bengio, Y. & Hinton, G. Deep learning. Nature, 521,436–444 (2015).

29 Olga Russakovsky, Jia Deng, Hao Su, Jonathan Krause, Sanjeev Satheesh, Sean Ma, Zhiheng Huang, Andrej Karpathy, Aditya Khosla, Michael Bernstein, Alexander C. Berg and Li Fei-Fei. ImageNet Large Scale Visual Recognition Challenge.Intern. Journal of Computer Vision, 2015

30 François Chollet, Deep Learning with Python, Second Edition, 2021, Manning Publications.