Bioc@pt

Automatic biodiversity sensors in forest

BIOSEFAIR Metaprogram

Project summary: 2021 - 2024

July 2024

The Bioc@pt project explores new ways of acquiring biodiversity data that are faster and less costly than traditional morphological identification in the laboratory, by automating field surveys (using sensors) and species identification (using artificial intelligence techniques applied to photographic recognition), with a threefold focus on biomonitoring, biodiversity monitoring, and the study of spatiotemporal patterns of biodiversity. This exploratory approach involved two important and relatively poorly understood forest taxonomic groups (insects and bryophytes) and can be adapted to other groups.

A feasibility test of autonomous visual recognition using deep learning was conducted on photos of bryophyte spores, sampled in the forest by actively aspirating airborne particles with an automatic Cyclone Sampler sensor. We developed a CNN-type algorithm capable of discriminating between "spore"-type objects and built a photo bank of diaspores from 54 bryophyte species, including those recorded by naturalist inventory at the aspiration sampling sites.

Based on the training data, the overall recognition rate is insufficient, ranging between 63% and 80%, and the average sensitivity and accuracy are low for all species, with frequent confusion between certain species. Despite the small amount of biological material observed, the two learning algorithms (trained on small and large spores, respectively) were applied to photographs of cyclone sampler samples, after automatic cropping of objects present in thumbnails, which also contained truncated objects or objects other than spores. This procedure produced results that were difficult to use. Training an EfficientDet AI object detection model, which detects targeted objects without prior image processing and takes into account the relative size of spores for classification, could prove more effective and is currently being tested.

In Bioc@pt, we also developed a prototype insect sensor, i.e. an automatic camera trap, and tested its operation in the forest. For this pilot project, we chose a group of wood-boring beetle species (Agrilus buprestids), which were already the subject of ecological investigations by our team.

In collaboration with Cap2020 for the mechatronic engineering of the sensor (housing, backlit optical sensor, communication, backup, energy autonomy), we have transformed a green Lindgren trap, selective for buprestids, into a prototype selective, non-destructive, automatic, and connected photographic sensor. Remote updating of the script controlling the regular frequency of shots and remote transmission of images are now operational. The geometry of a new camera chamber, with an anti-return collar, optimized by EFNO using 3D printing after observing the behavior of trapped insects, has been integrated into version 2 of the Cap2020 sensor box and successfully tested in the forest in the summer of 2024. The solar panel placed above the trap suspended in the canopy to recharge the batteries did not provide sufficient energy autonomy and still needs to be improved.

In parallel with this calibration of an automated field sensor, we conducted a test of autonomous visual recognition identification of photos of Agrilus beetles.

While the medium-term goal is indeed an automatic camera trap with an embedded photo processing algorithm, we have worked here on an automatic laboratory sorting tool capable of automatically recognizing (classification by deep learning) beetles photographed in the laboratory, in order to automate the analysis of samples from conventional traps. We have not explored the automation of specimen handling by a sorting robot, but have tested algorithms for recognizing specimens photographed in series with a digital laboratory microscope. A custom neural network was created and trained to discriminate between classes (species) in a labeled learning photo library, which now consists of several thousand images of the target taxa, i.e., the seven main regional species of the genus Agrilus, and an eighth non-regional species of Agrilus, Agrilus planipennis, an invasive alien species at the gates of Europe. In a second step, we added a second group of test species, constituting a group external to Agrilus buprestids and of interest to our laboratory, by integrating eight species of bark beetles also found in declining oak forests.

The precision/recall diagram and confusion matrix below confirm the neural network's good performance in discriminating between taxa. With the expanded training image library, the overall recognition rate, which is better than in the bryophyte spore test, reaches 90%. There is no confusion between the two groups (bark beetles and buprestids), more confusion between species within the buprestid group than between bark beetles, and several species are recognized with 100% accuracy. Another interesting result is the demonstration of the system's transferability: the algorithm trained with the laboratory photo library proved capable of discriminating between Agrilus species in photos taken with the field sensor, even though the latter is of lower optical quality than the digital laboratory microscope. The finalization of the Bioc@pt buprestes sensor has been included in two EFNO projects currently under review, and this sensor is intended to serve as proof of concept. The adaptation of its electronic component (photo sensor, backup/remote transmission, power supply, lighting) to other traps is being considered.

Detailed results

Summary of bryological and mycological studies

Sampling of bryophyte spores in forests using an automatic sensor (Cyclone Sampler)

Collection of 19 samples by active aspiration of airborne particles in forests at the Tours Grandmont (Indre-et-Loir) and Bourges Larçay (Cher) sites. The aspirator used (Cyclone Sampler from Burkard Manufacturing Co Ltd) collects dry samples in Eppendorf tubes.

A processing protocol (hydration, centrifugation, deposition on a slide) was developed to extract the material collected dry by the cyclone sampler and prepare an aqueous solution between the slide and cover slip for observation and photography using a digital microscope. A digital microscope photography protocol was developed.

For each sample, photographs were taken at the following magnifications:

x500 (sharper photos, fewer depth of field problems depending on the thickness of the objects)

x1000 (less sharp photographs due to the "veil" created by the coverslip at this magnification, but potentially more detail, e.g., grain size or surface ornamentation, on the photographed objects).

A number of technical difficulties complicated the study:

Very little biological material (spores) was collected, probably due to dry air capture, as the induced air flow can also cause spores to be released. The capture of objects of different sizes and thicknesses complicated the development of photos under a microscope slide.

Capture in an aqueous medium, followed by filtering, should be tested.

Microscope photography: a lot of time was needed to take photos, with many shots of no interest, and the need to rehydrate during shooting (with the risk of material "leaking").

Testing the identification of bryophyte spore photos using autonomous visual recognition

Creation of a learning database, using slide and cover slip preparations for spores from 54 species of bryophytes and 3 species of fungi, as well as for propagules (6 species of bryophytes) and elaters (1 species of bryophyte). The preparations were made from freshly harvested fertile individuals or from collections (herbariums).
Development of an algorithm to cut out the spores in the images and analyze them by shape, size, and color criteria; development of confusion matrices on validation and test samples.
Calculated on the entire matrix, the overall recognition rate (out of the total number of cases, how many were assigned to the correct species?) does not exceed 80% on the validation matrix and drops to 63% on the test sample in the case of 1000x magnification. Sensitivity is low (only 62 to 68% of all occurrences of the species are correctly predicted). Average accuracy is also low: of all individuals assigned to a species by the algorithm, 23% are incorrectly assigned to that species.
From the perspective of studying the dispersion of rare species or detecting pathogenic or invasive species, these results would not be satisfactory.

Table 1. Analysis of confusion matrices based on the training set (training matrices on 80% of the training set and validation matrix on the remaining 20%). Accuracy = overall recognition rate. Erreur globale = Overall error = 1 – accuracy; Specificity = Number of cases where the species name is not (correctly) assigned to an individual that is not of that species (TN) out of the actual number of individuals that are not of that species; taux de faux positifs = False positive rate = Number of cases where the species name is (incorrectly) assigned to an individual that is not of that species out of the actual number of individuals that are not of that species; taux de faux négatifs = False negative rate = Number of cases where the species name is not (incorrectly) assigned to an individual of that species out of the actual number of individuals that are of that species; Sensibilité = Sensitivity (=recall) = ratio between correct positive predictions and the total number of actual occurrences of the species; precision = ratio between correct positive predictions and the total number of positive predictions. F1 score = harmonic mean of precision and sensitivity.

	X 500		x 1000	
%	Training	Validation	Training	Validation
Accuracy	**72**	**68**	**80**	**63**
Erreur globale	**28**	**32**	**20**	**37**
Spécificité moy	99 +/- 2	99 +/- 2	99,5 +/- 0,6	99 +/- 0,99
Taux moy de Faux Positifs	**1,1 +/- 2**	**1,2 +/- 2,2**	**0,55 +/- 0,6**	**0,99 +/- 0,99**
Taux moy de Faux Négatifs	35 +/- 28	38 +/- 29	32 +/- 23	38 +/- 27
Sensibilité (Recall) moy	**64 +/- 28**	**62 +/- 29**	**68 +/- 23**	**62 +/- 27**
Précision moy	**77 +/- 18**	**73 +/- 26**	**72 +/- 20**	**68 +/- 19**
Score F1 moy	**66 +/- 25**	**69 +/- 27**	**68 +/- 21**	**60 +/- 21**
Balanced Accuracy moy	82 +/- 14	80 +/- 14	84 +/- 12	80 +/- 13

Application of the two learning algorithms (trained on small and large spores, respectively) to photographs of cyclone sampler samples: the procedure used produced results that were difficult to exploit. Improvements to the procedure are needed.

Summary of entomological studies

Mechatronic development of a prototype photographic entomological sensor

A "smart," selective, non-destructive, connected entomological sensor equipped with artificial intelligence for the automatic detection and visual identification of target insects has been developed. It has been tested on a group of pioneer wood-boring beetle species (Agrilus sp.), which includes invasive cosmopolitan species, heritage species, and species involved in forest decline.

The sensor was developed using the results of regional projects (IR CANOPEE (2019-2022) and Entomoc@pt) in collaboration with the company Cap2020.

Photography of insects captured by the Lindgren trap (non-destructive) is based on the following modules: housing (robustness, lightness, size), optical sensor (camera position, presence detection, image acquisition trigger), lighting, communication (modem and GSM/low-speed network remote transmission), backup (image recording on digital memory card), energy autonomy (rechargeable battery). The shooting frequency is programmed, systematic, and regular, and not triggered.

Implementation of an automatic green Lindgren trap and calibration of detection

Two prototypes of the automatic green Lindgren trap were placed in June 2022 in the Vierzon state forest (18), in plots known for their abundance of Agrilus beetles. These sensors operated every day from June 17 to July 28, for at least 6 hours between 8 a.m. and 8 p.m. (Agrilus beetles are diurnal), i.e., 12 acquisitions of 30 minutes, with 6 photos, spaced 5 minutes apart and under 3 different lighting conditions per acquisition.

This equates to 84 acquisitions per week = 504 images x 3 lighting conditions = 7.7 GB of memory.

During the 2022 field phase, the trap did not capture any insects. An adaptation was designed and tested in 2024 to force insects entering the trap to pass in front of the sensor.

Results of tests to identify beetle photos using autonomous visual recognition

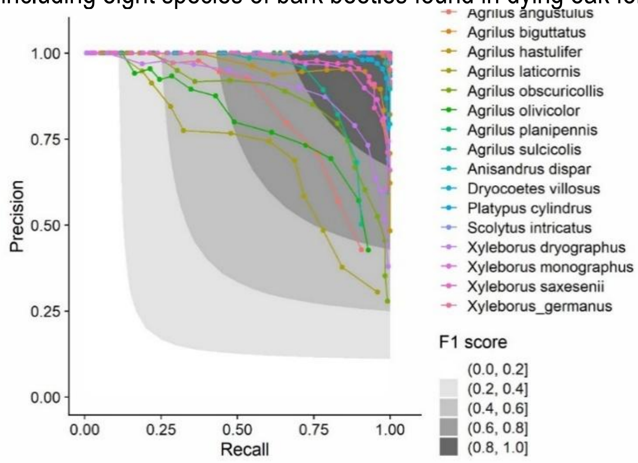
For the automation of photographic identification of beetles, three workflows were compared: (i) automatic camera trap with embedded processing algorithm, (ii) automatic camera trap with post-processing on a remote computing server, (iii) conventional trap with automatic processing of the batch of specimens in the laboratory: photographic identification using an Entomoscope-type system or genetic identification by metabarcoding.

The Bioc@pt project focused on calibrating the algorithm using photographs that had already been isolated. The training image database consists of several thousand photographs of the target taxa, i.e. the seven main regional species of the Agrilus genus, several thousand photographs of other beetles taken under the same conditions (to serve as an external group in the classification), and an eighth non-regional species of Agrilus, Agrilus planipennis, an invasive alien species native to Asia, which causes significant damage in the US and is on the verge of reaching Europe.

A neural network was created and trained to (i) distinguish Agrilus from other beetles and (ii) discriminate between the eight classes of Agrilus.

The model's performance results on eight classes of photos are encouraging: 100% of Agrilus beetles are recognized as Agrilus beetles, and 98% of non-Agrilus beetles are recognized as non-Agrilus beetles.

In a second step, we added a second group of test species, constituting a group external to Agrilus buprestes, by including eight species of bark beetles found in dying oak forests. The overall recognition rate reached 90%.



	A.lati	X.mono	X.germ	X.dryo	D.vill	A.obsc	A.bigu	A.disp	A.sulc	S.intr	P.cyli	A.hast	A.plan	A.angu	X.saxe	A.oliv
A.lati	62	0	0	0	0	15	0	0	3	0	0	4	0	3	0	6
X.mono	0	165	0	26	2	0	0	0	0	0	0	0	0	0	1	0
X.germ	0	0	210	0	0	0	0	0	0	0	0	0	0	0	1	0
X.dryo	0	4	0	138	3	0	0	0	0	0	0	0	0	0	7	0
D.vill	0	0	0	3	150	0	0	0	0	0	0	0	0	0	1	0
A.obsc	3	0	0	0	0	107	0	0	0	0	0	0	0	1	0	3
A.bigu	0	0	0	0	0	0	124	0	1	0	0	0	0	0	0	0
A.disp	0	0	0	1	1	0	0	63	0	0	0	0	0	0	0	0
A.sulc	10	0	0	0	0	6	2	0	82	0	0	12	0	3	0	2
S.intr	0	0	0	0	0	0	0	0	0	225	0	0	0	0	0	0
P.cyli	0	0	0	0	0	0	0	0	0	0	182	0	0	0	0	0
A.hast	0	0	0	0	0	4	0	0	0	0	0	92	0	0	0	0
A.plan	0	0	0	0	0	0	0	0	0	0	0	0	49	0	0	0
A.angu	11	1	0	0	1	27	0	0	2	0	0	1	0	71	0	2
X.saxe	0	0	0	4	1	0	0	0	0	0	0	0	0	0	105	0
A.oliv	3	0	0	0	0	23	0	0	2	0	0	1	0	3	0	62

Buprestes : Agrilus laticornis=A.lati; Agrilus obscuricollis=A.obsc; Agrilus biguttatus=A.bigu; Agrilus sulcicolis=A.sulc; Agrilus hastulifer=A.hast; Agrilus planipennis=A.plan; Agrilus angustulus=A.angu; Agrilus olivicolor=A.oliv; **Scolytes** : Xyleborus saxesenii=X.saxe; Xyleborus monographus=X.mono; Xyleborus_germanus=X.germ; Xyleborus dryographus=X.dryo; Dryocoetes villosus=D.vill; Anisandrus dispar=A.disp; Scolytus intricatus=S.intr; Platypus cylindrus=P.cyli;

Multiple tests are still underway to analyze the performance of neural networks with the augmented learning photo library. One interesting result is the demonstration of the system's transferability: the algorithm trained with the laboratory photo library (images of Agrilus taken with an electron microscope) proved capable of discriminating between Agrilus species in photos taken with the CapTrap Vision field sensor, despite the latter's lower optical quality. We have shown that artificially degrading the resolution of photos and stimulating the cropping of segmented thumbnails has little effect on classification efficiency.

Outputs

Le Borgne, H. & Bouget, C. 2023. Suivis de biodiversité par la reconnaissance automatique des espèces sur photographies : perspectives et défis. Naturae (6) : 75-96, https://doi.org/10.5852/naturae2023a6

Le Borgne, H. & Bouget, C. 2023. Suivis acoustiques de biodiversité : perspectives et défis en milieu continental terrestre. Naturae (8) : 129-150, https://doi.org/10.5852/naturae2023a8

Le Borgne, H. & Bouget, C. 2024. La reconnaissance des espèces basée sur l'ADN : applications, perspectives et défis en milieu continental terrestre. Naturae 2024 (3): 31-67. https://doi.org/10.5852/naturae2024a3