

> Bioc@pt

Capteurs automatiques de biodiversité en forêt

Métaprogramme BIOSEFAIR

Bilan de projet : 2021 - 2024

Juillet 2024

Le projet Bioc@pt explore de nouveaux moyens d'acquérir des données de biodiversité, plus rapides et moins coûteuses que la détermination morphologique classique en laboratoire, en automatisant les relevés sur le terrain (par des capteurs) et la détermination des espèces (par des techniques d'intelligence artificielle appliquées à la reconnaissance photographique), dans une triple optique de biovigilance, de suivis de biodiversité et d'étude des patrons spatiotemporels de biodiversité. Cette démarche exploratoire a concerné deux groupes taxonomiques forestiers importants et relativement mal connus (insectes et bryophytes) et pourra être adaptée à d'autres groupes.

Un test de faisabilité de reconnaissance visuelle autonome par apprentissage profond a été conduit sur des photos de spores de bryophytes, échantillonnées en forêt par aspiration active de particules aéroportées avec un capteur Cyclone Sampler automatique. Nous avons développé un algorithme de type CNN apte à discriminer des objets de type « spore » et constitué une banque photo des diaspores de 54 espèces de Bryophytes, incluant les espèces recensées par inventaire naturaliste sur les sites d'échantillonnage par aspiration. Sur les données d'apprentissage, le taux de reconnaissance global est insuffisant, entre 63% et 80%, et la sensibilité et la précision moyenne sont faibles pour l'ensemble des espèces, avec des confusions fréquentes entre certaines espèces. Malgré la faible quantité de matériel biologique observée, les 2 algorithmes d'apprentissage (entraînés respectivement sur des spores de

petite et de grande taille) ont été appliqués aux photographies d'échantillons du cyclone sampler, après détournage automatique des objets présents en imagerie, contenant aussi des objets tronqués ou autres que des spores. Cette procédure a fourni des résultats difficiles à exploiter. Entraîner un modèle IA de détection objet EfficientDet, qui détecte les objets ciblés sans traitement préalable de l'image et qui prend en compte la taille relative des spores pour la classification, pourrait s'avérer plus efficace et son test est en cours.

Dans Bioc@pt, nous avons également développé un prototype de capteur d'insectes à piège photographique automatique et testé son fonctionnement en forêt. Pour ce projet pilote, nous avons choisi un groupe d'espèces de coléoptères xylophages (buprestes *Agrilus* sp.), qui faisait déjà l'objet d'investigations écologiques dans notre équipe. En collaboration avec la société Cap2020 pour l'ingénierie mécatronique du capteur (boîtier, capteur optique rétro-éclairage, communication, sauvegarde, autonomie énergétique), nous avons transformé un piège Lindgren vert, sélectif à l'égard des buprestes, en un prototype de capteur photographique sélectif, non destructif, automatique et connecté. La mise à jour à distance du script pilotant la fréquence régulière des prises de vue et la télétransmission des clichés sont opérationnels. La géométrie d'une nouvelle chambre de prise de vue, avec collerette anti-retour, optimisée par EFNO par impression 3D après observation du comportement des insectes piégés, a été intégrée à la version 2 du boîtier de capteur Cap2020 et testée avec succès en forêt à l'été 2024. Le panneau solaire placé au-dessus du piège suspendu en canopée pour recharger des batteries n'a pas livré une autonomie énergétique suffisante et reste à améliorer.

En parallèle de cette calibration d'un capteur de terrain automatisé, nous avons procédé à un test d'identification par reconnaissance visuelle autonome des photos de coléoptères *Agrilus*. Si l'objectif à moyen terme est bien un piège photographique automatique avec algorithme embarqué de traitement des photos, nous avons ici travaillé sur un outil de tri automatique de laboratoire, habilité à assurer la reconnaissance automatique (classification par apprentissage profond) de coléoptères photographiés au laboratoire, de façon à automatiser le dépouillement d'échantillons provenant de pièges classiques. Nous n'avons pas exploré l'automatisation de la manipulation des spécimens par un robot trieur, mais testé des algorithmes de reconnaissance des spécimens photographiés en série avec un microscope numérique de laboratoire. Un réseau de neurone personnalisé a été créé puis entraîné pour discriminer les classes (espèces) d'une photothèque labellisée d'apprentissage, aujourd'hui constituée de plusieurs milliers de clichés du taxon cible, i.e. les 7 principales espèces régionales du genre *Agrilus*, et d'une 8e espèce non régionale d'*Agrilus*, *Agrilus planipennis*, une espèce exotique envahissante aux portes de l'Europe. Dans un second temps, nous avons ajouté un deuxième groupe d'espèces test, constituant un groupe externe aux buprestes *Agrilus* et d'intérêt pour notre laboratoire, en intégrant 8 espèces de scolytes également rencontrées dans les chênaies dépérissantes. Le diagramme précision/rappel et la matrice de confusion ci-dessous confirment les bonnes performances du réseau de neurones à discriminer les taxa. Avec la photothèque d'apprentissage augmentée, le taux de reconnaissance global, meilleur que dans le test sur les spores de bryophytes, atteint 90%. Il n'y pas de confusion entre les deux groupes (scolytes et buprestes), davantage de confusion entre espèces au sein du groupe des buprestes qu'entre scolytes, et plusieurs espèces sont reconnues à 100%. Un autre résultat intéressant est la démonstration de la transférabilité du système : l'algorithme entraîné avec la photothèque de laboratoire s'est montré capable de discriminer les espèces d'*Agrilus* sur des photos provenant du capteur de terrain, pourtant moins qualitatif au niveau optique que le microscope numérique de labo. La finalisation du capteur à buprestes Bioc@pt a été inscrite dans deux projets EFNO en cours d'examen, et ce capteur est voué à servir de preuve de concept. L'adaptation de sa brique électronique (capteur photo, sauvegarde/ télétransmission, alimentation, éclairage) à d'autres pièges est envisagée.

Résultats détaillés

Bilan des études bryologiques et mycologiques

Echantillonnage de spores de bryophytes en forêt par capteur automatique (Cyclone Sampler)

Récolte de 19 échantillons par aspiration active de particules aéropartées, en forêt, sur les sites de Tours Grandmont (Indre-et-Loire) et Bourges Larçay (Cher). L'aspirateur utilisé (Cyclone Sampler de Burkard Manufacturing Co Ltd) récolte des échantillons à sec, en tubes Eppendorf.

Un protocole de traitement (hydratation, centrifugation, dépôt sur lame) a été mis au point pour extraire le matériel récolté à sec par le cyclone sampler et réaliser la préparation aqueuse entre lame et lamelle pour observation et photographie au microscope numérique. Un protocole de photographie sur microscope numérique a été mis au point.

Pour chaque échantillon, les photographies ont été prises à grossissement :

x500 (photos plus nettes, moins de problèmes de profondeur de champ en fonction de l'épaisseur des objets)

x1000 (photographies moins nettes en raison du « voile » créé par la lamelle à ce grossissement, mais potentiellement plus de détails, par exemple granulométrie ou ornementation de surface, sur les objets photographiés).

Un certain nombre de difficultés techniques ont compliqué l'étude :

Très peu de matériel biologique (spore) récolté en raison sans doute de la capture à air à sec, le flux d'air induit pouvant aussi faire ressortir les spores. La capture d'objets de taille et épaisseur différentes a compliqué la mise au point des photos sous lamelle.

Une capture en milieu aqueux, puis filtrage serait à tester.

Photo au microscope : beaucoup de temps nécessaire à la prise de photos, avec de nombreux clichés sans objet d'intérêt, besoin de réhydrater en cours de prise de vue (avec risque de "fuite" de matériel)

Test d'identification des photos de spores de Bryophytes par reconnaissance visuelle autonome

Constitution d'une base d'apprentissage, par préparations entre lame et lamelle pour des spores de 54 espèces de Bryophytes et 3 espèces de Champignons, ainsi que pour des propagules (6 espèces de Bryophytes) et élatères (1 espèce de Bryophyte). Les préparations ont été réalisées à partir d'individus fertiles fraîchement récoltés ou issus de collection (herbiers).

Développement d'un algorithme pour détourner les spores dans les images et les analyser par critères de forme, taille et couleur ; élaboration des matrices de confusion sur les échantillons de validation et de test.

Calculé sur la matrice entière, le taux de reconnaissance global (sur le nombre total de cas, combien ont été attribués à la bonne espèce ?) ne dépasse pas 80% sur la matrice de validation et descend à 63% sur l'échantillon de test dans le cas du grossissement x 1000. La sensibilité est faible (sur toutes les occurrences de l'espèce, 62 à 68% seulement sont correctement prédites). La précision moyenne est faible aussi : sur tous les individus attribués à une espèce par l'algorithme, 23% sont attribués à tort à cette espèce)

Dans une optique d'étude de la dispersion d'espèces rares, ou de détection d'espèces pathogènes ou envahissantes, ces résultats ne seraient pas satisfaisants.

Tableau 1. Analyse des matrices de confusion sur la base d'apprentissage (matrices d'entraînement (= training) sur 80% de la base d'apprentissage et matrice de validation sur les 20 % restants). Accuracy = taux de reconnaissance globale. Erreur globale = 1 - accuracy ; Spécificité = Nombre de cas où le nom de l'espèce n'est pas attribué (à raison) à un individu qui n'est pas de cette espèce (TN) sur le nombre réel d'individus qui ne sont pas de cette espèce ; taux de faux positifs = Nombre de cas où le nom d'espèce est attribué (à tort) à un individu qui n'est pas de cette espèce sur le nombre réel d'individus qui ne sont pas de cette espèce ; taux de faux négatifs = Nombre de cas où le nom d'espèce n'est pas attribué (à tort) à un individu de cette espèce sur le nombre réel d'individus qui sont de cette espèce ; Sensibilité (=recall) = ratio entre les prévisions positives correctes et le nombre total d'occurrences réelles de l'espèce ; précision = ratio entre prévisions positives correctes et nombre total de prévisions positives. Score F1 = moyenne harmonique de précision et sensibilité.

%	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 des 2 algorithmes d'apprentissage (entraînés respectivement sur des spores de petite et de grande taille) **aux photographies d'échantillons du cyclone sampler** : la procédure utilisée a fourni des résultats difficiles à exploiter. Une amélioration de la procédure est à travailler.

Bilan des études entomologiques

Développement mécatronique d'un prototype de capteur entomologique photographique

Un capteur entomologique « intelligent », sélectif, non destructif et connecté et à l'intelligence artificielle pour la détection et l'identification visuelle automatiques d'insectes cibles a été développé. Il a été testé sur un groupe d'espèces de coléoptères xylophages pionniers (buprestes *Agilus* sp.), qui comporte des espèces cosmopolites invasives, d'autres patrimoniales ou impliquées dans les dépérissements forestiers.

La réalisation du capteur a bénéficié des résultats des projets régionaux (IR CANOPEE (2019-2022) et Entomoc@pt) en collaboration avec la société Cap2020.



La photographie d'insectes capturés par le piège Lindgren (non destructif) repose sur les modules suivants : boitage (robustesse, légèreté, encombrement), capteur optique (position de la caméra, détection de présence, déclenchement de l'acquisition d'image), éclairage, communication (modem et télétransmission réseau GSM/bas débit), sauvegarde (enregistrement des images sur carte mémoire numérique), autonomie énergétique (batterie rechargeable). La fréquence de prise de vue est programmée, systématique et régulière et non déclenchée.

Mise en œuvre d'un piège Lindgren vert automatique et calibration de la détection

Deux prototypes du Lindgren vert automatique ont été placés en juin 2022 en forêt domaniale de Vierzon (18), dans des parcelles connues pour leurs abondances en coléoptères *Agilus*. Ces capteurs ont fonctionné tous les

jours du 17 juin au 28 juillet, au moins 6 heures d'allumage entre 8h et 20h (les Agrilus sont diurnes), i.e. 12 acquisitions de 30 minutes, avec 6 photos, espacées de 5 minutes et sous 3 éclairages différents, par acquisition. Soit par semaine 84 acquisitions = 504 images x 3 éclairages = 7,7 Go de mémoire.

Lors de la phase terrain 2022, le piège n'a pas permis la capture d'insectes. Une adaptation a été conçue et testée en 2024 afin d'obliger l'insecte entré dans le piège à passer devant le capteur.

Bilan des tests d'identification des photos de coléoptères par reconnaissance visuelle autonome

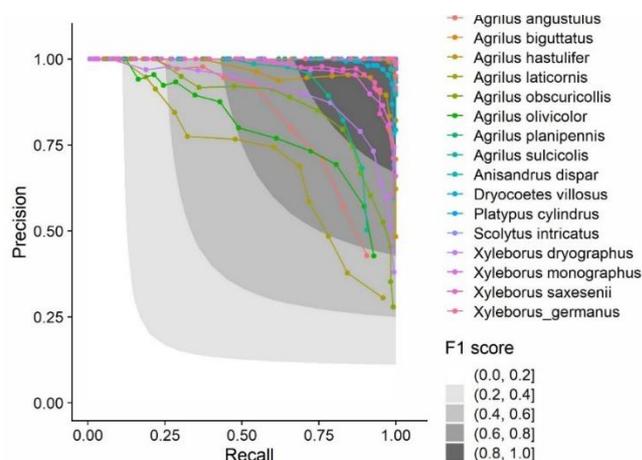
Pour l'automatisation de l'identification photographique des coléoptères, 3 workflows ont été comparés : (i) piège photographique automatique avec algorithme de traitement embarqué, (ii) piège photographique automatique avec post-traitement sur serveur de calcul distant, (iii) piège classique avec traitement automatique du lot de spécimens au laboratoire : identification photographique par système de type Entomoscope ou identification génétique par métabarcoding.

Le projet Bioc@pt s'est consacré à la calibration de l'algorithme sur photographies déjà isolées. La base d'images d'apprentissage est constituée de plusieurs milliers de clichés du taxon cible, i.e. les 7 principales espèces régionales du genre *Agrilus*, de quelques milliers de clichés d'autres coléoptères prises dans les mêmes conditions (pour jouer le rôle de groupe externe dans la classification), et d'une 8e espèce non régionale d'*Agrilus*, *Agrilus planipennis*, une espèce exotique envahissante originaire d'Asie, qui cause des dégâts importants aux USA et qui est aux portes de l'Europe.

Un réseau de neurone a été créé et entraîné pour (i) distinguer les *Agrilus* des autres coléoptères et (ii) discriminer les 8 classes d'*Agrilus*.

Les résultats de performance du modèle sur 8 classes de photos sont encourageants : 100% des agriles sont reconnus comme des agriles, 98% des non-agriles sont reconnus comme des non-agriles.

Dans un second temps, nous avons ajouté un deuxième groupe d'espèces test, constituant un groupe externe aux buprestes *Agrilus* en intégrant 8 espèces de scolytes rencontrées dans les chênaies déperissantes. Le taux de reconnaissance global atteint 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;

De multiples tests sont encore en cours pour analyser les performances des réseaux de neurones avec la photothèque d'apprentissage augmentée. Un des résultats intéressants est la démonstration de la transférabilité du système : l'algorithme entraîné avec la photothèque de laboratoire (images d'*Agrilus* prises au microscope électronique) s'est montré capable de discriminer les espèces d'*Agrilus* sur des photos provenant du capteur de terrain CapTrap Vision, pourtant moins qualitatif au niveau optique. Nous avons montré que la dégradation artificielle de la résolution des photos et le détournement des imagerie segmentées affecte peu l'efficacité de classification.

Valorisation

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>