

MARS 2023

RAPPORT DE RECHERCHE

ANALYSE AUTOMATISÉE DE LA DIVERSITÉ ACOUSTIQUE, DE LA DÉTECTION D'ESPÈCES AUX INDICES BIOACOUSTIQUES

Remerciements

Ce rapport de recherche bénéficie de l'environnement et certains des algorithmes développés dans le cadre de la Chaire nationale en IA, Bioacoustique ADSIL, co-financée par l'ANR, l'AID et la DGA,

PI Glotin, 2021-2025, ANR-20-CHIA-0014-01.



LIS CNRS UTLN

Stéphane Chavin, Pierre Mahé, Thomas Hermet, Nicolas Deloustal, Hervé Glotin
LIS DYNi université de Toulon, France
contact : stephanechvn@gmail.com

In coll.with : Jean-François Jetté, Marc-André Villard, Felix Couture-Maheu, Joëlle Spooner, et Anouck Simard MFFP, Québec

Table des matières

1.	Introduction	2-3
2.	Matériel et Méthode	3-7
	2.1 Jeu de donnée et annotations	3
	2.2 Réseau de neurones à convolution de type YOLO	5
	2.3 Indices bioacoustiques.....	6
3.	Résultats	7-12
	3.1 Détection d'espèces	7-10
	3.2 Suivi de la présence d'espèces en fonction du temps et le la localisation	11
	3.3 Suivi des indices acoustiques	12
4.	Discussion	14
	4.1 Fusion des oiseaux et des orthoptères dans un même réseau	14
	4.2. Les résultats de l'indice bioacoustiques ACI semblent corrélés aux détections d'oiseaux	14
5.	Conclusion	14
6.	Outils de dépôt de fichier en ligne et traitement automatisé	15-16

1. Introduction

La bioacoustique est un domaine scientifique interdisciplinaire qui étudie les sons produits par les organismes vivants et tente de les associer à un comportement ou un état de santé de ces derniers. Cette étude implique la biologie, la physique, la psychologie, l'ingénierie acoustique et d'autres domaines connexes. La bioacoustique se concentre sur les sons produits par les animaux, leur utilisation pour la communication, la navigation, la reconnaissance et la localisation des proies, ainsi que l'effet des sons sur le comportement et la santé des animaux. Les sons produits par les animaux peuvent varier en fréquence, en amplitude et en durée, et peuvent être produits de différentes manières, tel que par des vocalises, des percussions ou des mouvements du corps.

Parmi les animaux terrestres capables de produire des sons, on retrouve dans un premier temps les oiseaux, qui émettent des vocalises dans des fréquences proches de 8 kHz et qui, par la succession de ces dernières, forment un chant audible (*Metcalf, O. C., Barlow, J, 2021*). Ensuite, on retrouve dans un second temps les Chiroptères qui émettent des vocalises dans la bande de fréquence des ultrasons. Ces vocalises sont utilisées pour la communication mais également pour la localisation dans l'espace (*Aguilar Silva, C., Machado 2022*). Dans cette liste on retrouve également les Anoures, qui émettent des vocalises tandis que pour finir les Orthoptères émettent eux des vibrations, pour la plupart dans les ultrasons, par le frottement de leurs ailes par exemple (*Metcalf, O. C., Barlow, J, 2021 ; Yves Bas, Raphael Colombo 2021 ; Schmidt, 2013*).

Sur le terrain, des microphones, des enregistreurs numériques et des logiciels d'analyse sont utilisés pour collecter et analyser les données acoustiques, tentant ainsi d'apporter des conclusions selon ce qui est étudié. La bioacoustique a de nombreuses applications importantes dans la conservation de la biodiversité et notamment dans la gestion de la faune. Concernant la conservation de la biodiversité, la bioacoustique peut être utilisée pour surveiller les populations d'animaux rares et menacés, ce qui peut aider à protéger ces populations. Dans la gestion de la faune, la bioacoustique peut être utilisée pour comprendre les habitudes de migration et de reproduction des animaux, ce qui peut aider à protéger les populations animales.

Dans un environnement donné, de multiples espèces de groupes différents émettent des sons et forment ainsi ce que l'on peut appeler la biophonie. En ajoutant également le bruit naturel de l'environnement, le vent, la pluie, les ruisseaux etc, appelé géophonie, ainsi que le bruit d'origine anthropique, l'anthropophonie, on forme ce qui se nomme le paysage acoustique (*Sueur et al. 2014*).

L'étude d'un paysage acoustique relève donc, du fait du nombre de composants impliqués, d'un défi de séparation des composants. Différentes méthodologies ont déjà été apportées pour répondre à cette

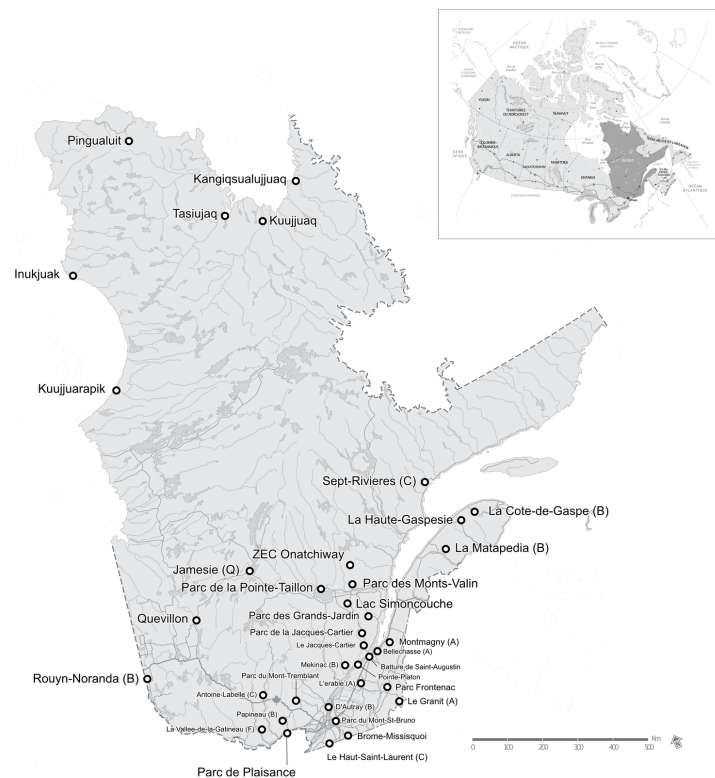
problématique dont le calcul d'indices permettant d'attribuer une valeur à un enregistrement qui, en comparaison avec d'autres enregistrements, indique si l'activité est supérieure ou non. Bien que cette méthode soit importante, il a également été tenté une approche par utilisation de réseaux de neurones capables de reconnaître une vocalise et de l'attribuer à une espèce (Sevilla, Glotin H. 2017). Étant de plus en plus utilisé et notamment dans le son, ces réseaux de neurones ont cependant montré des failles dans l'étude des paysages acoustiques et notamment en raison de la superposition des vocalises de différentes espèces dans un même enregistrement.

Ainsi, la méthode proposée dans ce rapport est une méthode de calcul d'indices acoustiques complétée par un apprentissage automatique par réseau de neurones. En effet, le réseau de neurones à convolution de type YOLO s'est montré performant concernant la réduction des anomalies causées par les chœurs et a donc été adapté à l'étude des sons et à la reconnaissance de vocalises d'oiseaux ainsi que les sons d'Orthoptères.

2. Matériels et Méthode

2.1. Jeu de données et annotations

Pour la réalisation d'un algorithme capable de détecter et classifier des vocalises il est important d'avoir une base de données annotées par des



source : ressources-naturelles.canada.ca

Figure 1. Cartographie d'une partie des stations d'enregistrements sur le territoire du Québec

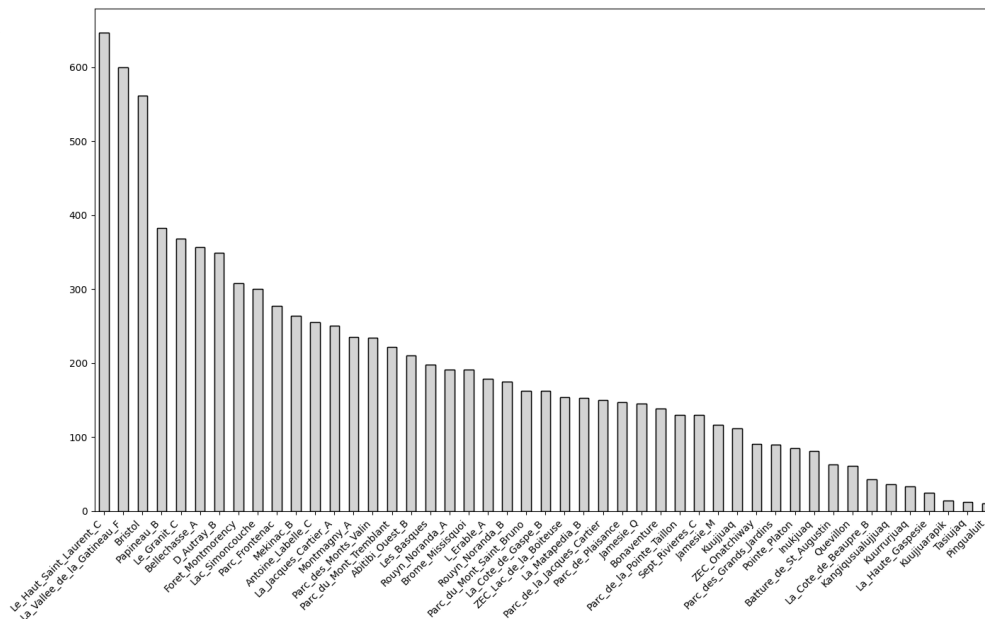
Une partie des stations ont été disposées sur cette carte afin de mettre en évidence l'étendu des zones d'écoutes.

ornithologues en mesure d'apporter des centaines d'enregistrements annotés de la manière la plus précise possible. Dans le cas de ce travail, deux bases de données, issues du travail de plusieurs spécialistes ornithologues, composées de 11.366 annotations d'oiseaux (*BdQc_Database_Bird_Frog_Mammal_2022-10-03*) et 1.426 annotations d'insectes (*Data_Insectes_BdQc_JFJ_2022-10-29*) ont été compilées. Ces annotations avec un temps de début et fin de la vocalise sont précises à la seconde et réunissent plusieurs centaines d'espèces au total (Figure 2a.). La durée des annotations est la plupart du temps comprise entre 1 et 5 secondes (Figure 2c.). La banque de données annotées est également accompagnée par une banque d'enregistrements complète correspondant aux enregistrements journaliers sur plus de 60 stations différentes et à des heures différentes (Figure 2b.)

a.

Annee	Cellule	Site	Date_Heure	Obs	Task	TaskOn	Duree	Statut	Vol	?	Code	On	Off	On (sec)	Off (sec)	Class	#	Type	Amp	Pur	note specifique au signal sonore
2016	Lac Simoncouche	132_116_F01	20160518_230000	vg	pc	0	3	done			pscr	0	300	0	180	frog	cl3	song	bad	good	
2016	Lac Simoncouche	132_116_F01	20160518_230000	vg	pc	0	3	done			lisi	0	300	0	180	abio					
2016	Lac Simoncouche	132_116_F01	20160520_050000	jf	pc	0	3	done	20		oven	0	2	0	2	bird	1	song	mid	mid	
2016	Lac Simoncouche	132_116_F01	20160520_050000	jf	pc	0	3	done	20		lisi	0	10	0	10	abio	1	abio	mid	mid	
2016	Lac Simoncouche	132_116_F01	20160520_050000	jf	pc	0	3	done	20		swth	1	3	1	3	bird	1	call	mid	mid	qui-brrrrr call
2016	Lac Simoncouche	132_116_F01	20160520_050000	jf	pc	0	3	done	20		heth	6	7	6	7	bird	1	call	mid	mid	flight call
2016	Lac Simoncouche	132_116_F01	20160520_050000	jf	pc	0	3	done	20		wiwr	11	17	11	17	bird	1	song	bad	bad	
2016	Lac Simoncouche	132_116_F01	20160520_050000	jf	pc	0	3	done	20		btbw	120	122	80	82	bird	1	song	mid	mid	
2016	Lac Simoncouche	132_116_F01	20160520_050000	jf	pc	0	3	done	20		wbsp	129	133	89	93	bird	1	song	bad	mid	
2016	Lac Simoncouche	132_116_F01	20160520_050000	jf	pc	0	3	done	20		oven	142	144	102	104	bird	1	song	mid	good	
2016	Lac Simoncouche	132_116_F01	20160521_010000	csp	pc	0	3	done		jf	pscr	0	20	0	20	frog	cl3	song	bad	good	

b.



c.

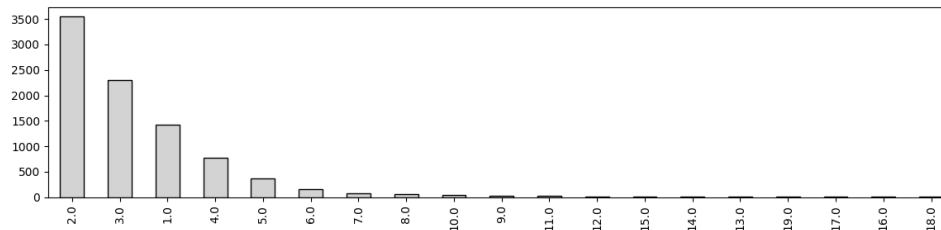


Figure 2. Base de donnée annoté sur les enregistrements du Québec

Exemple du formalisme de l'annotation des fichiers audio (a), détail du nombre d'enregistrements par stations (b) et représentation de la durée des annotations (c).

YOLO (You Only Look Once) est un modèle de détection d'objet (*Redmon, J. et al 2016*), c'est-à-dire qu'il a été entraîné à reconnaître des formes, des textures et d'autres caractéristiques visuelles. Ses utilisations sont principalement dans la reconnaissance sur des images provenant de caméras. Cependant, en considérant qu'une vocalise possède également une forme sur la représentation du son en spectrogramme, c'est-à-dire en appliquant une transformée de Fourier au signal, on peut donc apprendre au réseau à reconnaître la forme de cette dernière avec le modèle YOLO (Chavin S. et al 2023). La version de YOLO choisie pour répondre à la problématique est YOLO v5 (*Jocher, G. 2020*). Cette dernière se décline en plusieurs niveaux de performance. Celle utilisée dans ce travail est YOLOV5L, c'est à-dire la version Large, avec un plus grand nombre de paramètres et donc une plus grande complexité. Il a été préféré de travailler avec cette version afin d'obtenir de meilleurs résultats malgré le fait qu'elle soit plus lente et plus lourde. En effet la différence entre la version Large et Small peut atteindre près de 15% de mean Average Precision (mAP), soit une importante différence de classification et de précision.

Parmi toutes les annotations présentes dans la base annotée, seules celles provenant d'enregistrements avec un bon rapport signal à bruit et une bonne amplitude ont été conservées, ce choix s'est fait en sélectionnant seulement les annotations avec la précision 'good' concernant l'amplitude et la qualité des enregistrements. Le but étant de reconnaître la forme de la vocalise, il est important que cette dernière soit de qualité pour l'entraînement du réseau de neurones. L'extraction des spectrogrammes a été réalisée avec une taille de fenêtre de 1.024 secondes et une hop_length de 512. La durée d'un spectrogramme est de 8 secondes pour les oiseaux et 5 secondes pour les orthoptères, avec un overlap de 2 secondes. Toutes les espèces n'ont pas été conservées pour l'apprentissage. En effet ce sont principalement celles avec une vocalise reconnaissable mais surtout les espèces qui ont été les plus annotées par les spécialistes. Ainsi, un set d'entraînement et un set de validation ont été mis en place avec les valeurs présentées dans le Tableau I. Ce sont donc 21 espèces d'oiseaux qui ont été sélectionnées, avec entre 58 et 493 annotations pour chaque espèce.

Concernant les Orthoptères, une liste de 8 espèces a été sélectionnée parmi les 58 espèces présentes dans les annotations. Au total, sur les 8 espèces, ce sont plus de 300 annotations avec certaines espèces à seulement une dizaine d'annotations. L'entraînement tente de prendre en compte, lorsque cela est possible, ~ 70% du jeu de données annoté tandis que la validation ~ 30%. Il a été vérifié que la répartition soit telle qu'il y a plus ou moins 2 fois plus d'annotations dans l'entraînement que dans la validation, principalement pour les oiseaux. Pour l'obtention des bounding box, c'est-à-dire une annotation en temps/fréquence, nécessaire à l'entraînement de YOLO, il a été décidé d'attribuer une valeur seuil en fréquence pour les Orthoptères, entre 5kHz et 11kHz) tandis que pour les oiseaux, ayant des vocalises plus complexes l'ensemble des annotations ont été redimensionnées à la main avec l'outil d'annotation LabelMe.

Cette action a pour effet d'améliorer la détection de différentes vocalises malgré la superposition de ces dernières. Il est important que l'ensemble des vocalises présentes sur un spectrogramme soient annotées. Dans le cas contraire, il sera compliqué pour le réseau de faire la distinction entre le bruit et les vocalises d'intérêts.

Tableau 1. Répartition des annotations pour l'entraînement et la validation de YOLOv5

Code	Nom français	Nom latin	Entraînement	Validation	Total
alfl	Moucherolle des aulnes	<i>Empidonax alnorum</i>	164	77	241
amre	Paruline flamboyante	<i>Setophaga ruticilla</i>	69	28	97
blja	Geai bleu	<i>Cyanocitta cristata</i>	102	17	119
btbw	Paruline bleue	<i>Setophaga caerulescens</i>	50	20	70
btnw	Paruline à gorge noire	<i>Setophaga virens</i>	72	42	114
coye	Paruline masquée	<i>Geothlypis trichas</i>	237	106	343
gcki	Roitelet à couronne dorée	<i>Regulus satrapa</i>	78	33	111
heth	Grive solitaire	<i>Catharus guttatus</i>	180	80	260
lisp	Bruant de Lincoln	<i>Melospiza lincolni</i>	40	18	58
mawa	Paruline à tête cendrée	<i>Setophaga magnolia</i>	73	31	104
nawa	Paruline à joues grises	<i>Oreothlypis ruficapilla</i>	166	70	236
oven	Paruline couronnée	<i>Seiurus aurocapilla</i>	158	59	217
rcki	Roitelet à couronne rubis	<i>Regulus calendula</i>	60	29	89
revi	Viréo aux yeux rouges	<i>Vireo olivaceus</i>	116	38	154
rwbl	Carouge à épaulettes	<i>Agelaius phoeniceus</i>	118	55	173
sosp	Bruant chanteur	<i>Melospiza melodia</i>	44	14	58
swsp	Bruant des marais	<i>Melospiza georgiana</i>	211	95	306
swth	Grive à dos olive	<i>Catharus ustulatus</i>	75	44	119
veer	Grive fauve	<i>Catharus fuscescens</i>	69	18	87
wtsp	Bruant à gorge blanche	<i>Zonotrichia albicollis</i>	352	141	493
yrwa	Paruline à croupion jaune	<i>Setophaga coronata</i>	61	33	94
Nombre d'annotation :			2495	1048	3543

Code	Nom français	Nom latin	Entraînement	Validation	Total
alal	Némobie des prés	<i>Allonemobius allardi</i>	41	38	79
alfa	Némobie striée	<i>Allonemobius fasciatus</i>	39	33	72
pscu	Criquet des champs	<i>Pseudochortippus curtispennis</i>	9	6	15
sccu	Scuddérie des bois	<i>Scudderia curvicauda</i>	8	7	15
scpi	Scuddérie à ailes larges	<i>Scudderia pistillata</i>	27	16	43
stgc	Éclaireur élégant	<i>Stéthompyhma gracile</i>	48	38	86
unco	Unknown Conocephalinae	-	11	8	19
untk	Unknown Ticker	-	24	15	39
Nombre d'annotation :			207	161	368

2.3. Indices Bioacoustiques

Il existe plusieurs types d'indices pour l'étude bioacoustique, et selon ce que l'on cherche à étudier. D'abord les indices qui utilisent l'amplitude ou l'intensité de l'énergie sonore des signaux. Ensuite, certains indices vont analyser le niveau de complexité du temps, de la fréquence et/ou de l'amplitude du signal. Enfin, d'autres indices prennent en compte le paysage sonore.

Les indices sélectionnés pour répondre à la problématique de ce projet sont l'ACI (Acoustic Complexity Index), l'ADI (Acoustic Diversity Index), l'AEI (Acoustic Entropy Index), BI (Bioacoustics Index), H (Shannon Entropie) et NDSI (Normalized Difference Soundscape Index) et ont été appliqués sur l'ensemble des enregistrements. Alors que l'indice NDSI va séparer les bandes de fréquences, par l'application de masques, en biophonie, géophonie et anthropophonie pour calculer la part de biophonie et donc indiquer si cette dernière est haute ou non, l'indice ACI lui va mesurer la complexité de l'enregistrement en mesurant les variations en temps et en fréquences. Ainsi, plus il y a de variations, c'est-à-dire dans le cas des oiseaux qu'il y a davantage d'espèces différentes qui chantent ensemble, alors plus l'indice ACI sera haut.

3. Résultats

3.1. Détection d'espèces

Parmi les 21 espèces d'oiseaux, on obtient entre 28% et 90% de moyenne des précisions (Average Précision) avec une moyenne des moyennes des précisions de 67% (mean Average Precision).

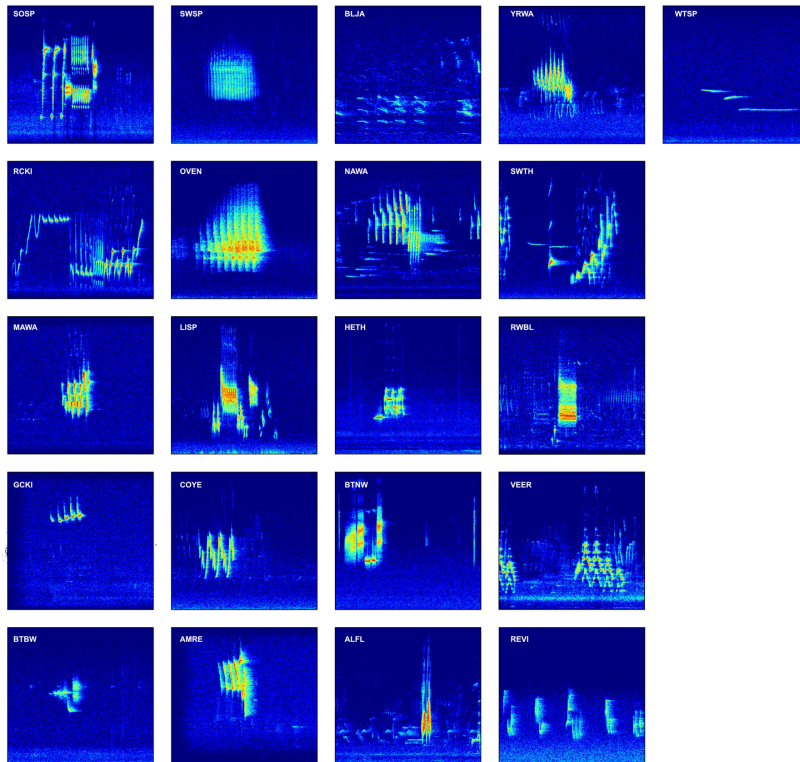


Figure 3. Liste des différentes vocalises détectées

Exemple des 21 types de vocalises d'oiseaux détectées par le réseau de neurones YOLO. Les spectrogrammes ont été réalisés Short Time Fourier Transform (STFT) avec une taille de fenêtre de 1.024 secondes et une hop_length de 512.

Ces chiffres indiquent que 67% des annotations ont bien été prédites par le réseau (Figure 4), cependant, avec une précision de 76%, on peut donc dire que le modèle ne prédit que rarement de fausses détections. En effet il semble que ce soit plutôt une sous-estimation des détections plutôt qu'un problème de confusion malgré une confusion de maximum 11% entre certaines espèces. En regardant la forme des annotations (Figure 3), on peut se rendre compte que certaines vocalises sont très ressemblantes, notamment heth avec swth ainsi que mawa avec amre. Ces ressemblances peuvent expliquer la confusion des vocalises mais il est constaté qu'il est possible de filtrer ces confusions. En effet le réseau prédit un score de confiance pour chaque détection et ce score est faible lorsque la confusion est probable. Ainsi, en ne prenant que les détections assurées par le réseau, il est possible d'obtenir des scores importants. Il a notamment été testé d'entraîner un modèle sur des détections avec un score de confiance supérieur à 80% et cela a montré un gain de performance de 15% pour le réseau. Dans le cas d'une simple détection d'une espèce dans un milieu, cela est intéressant mais le risque de manquer certaines vocalises est important. Au contraire, si l'objectif est d'essayer de compter le nombre de vocalises dans un enregistrement, alors il est préférable d'utiliser le modèle avec un plus gros risque de confusion mais beaucoup plus de détections.

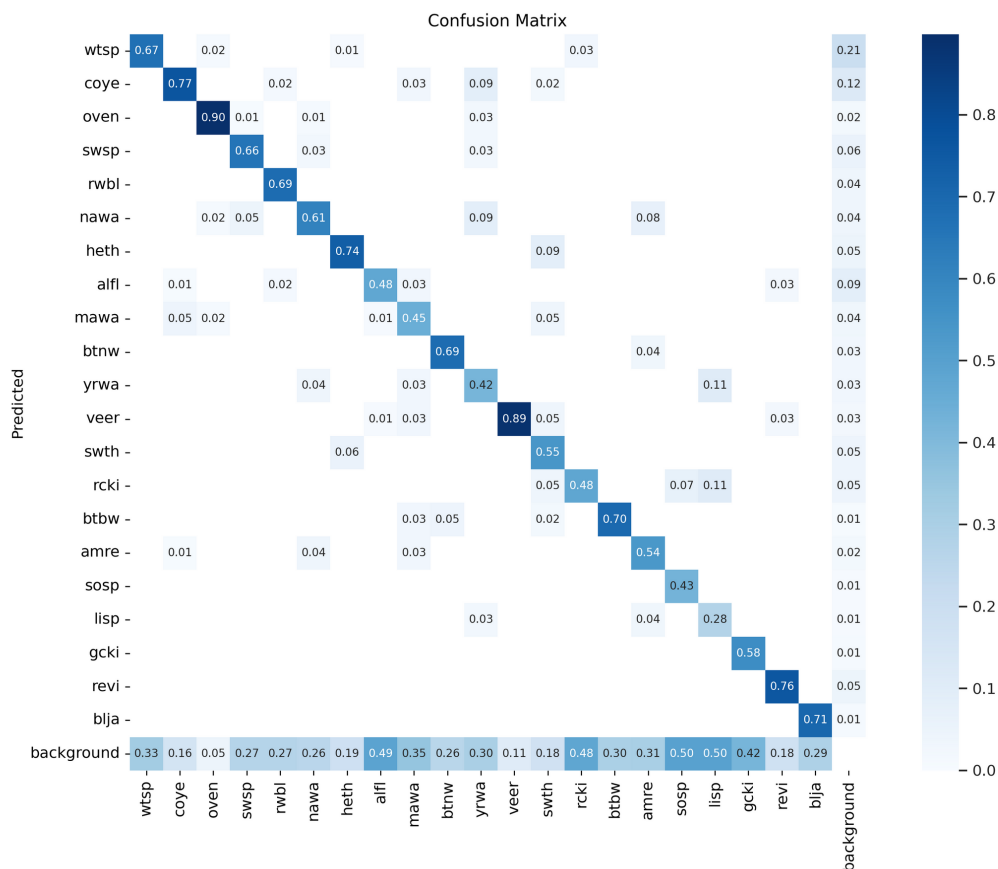


Figure 4. Matrice de confusion de la validation du réseau après un entraînement sur les vocalises d'oiseaux

Sur cette figure sont représentées les valeurs de prédiction du réseau en fonction des vrais labels. La colonne tout en bas et tout à droite représente le bruit des enregistrements.

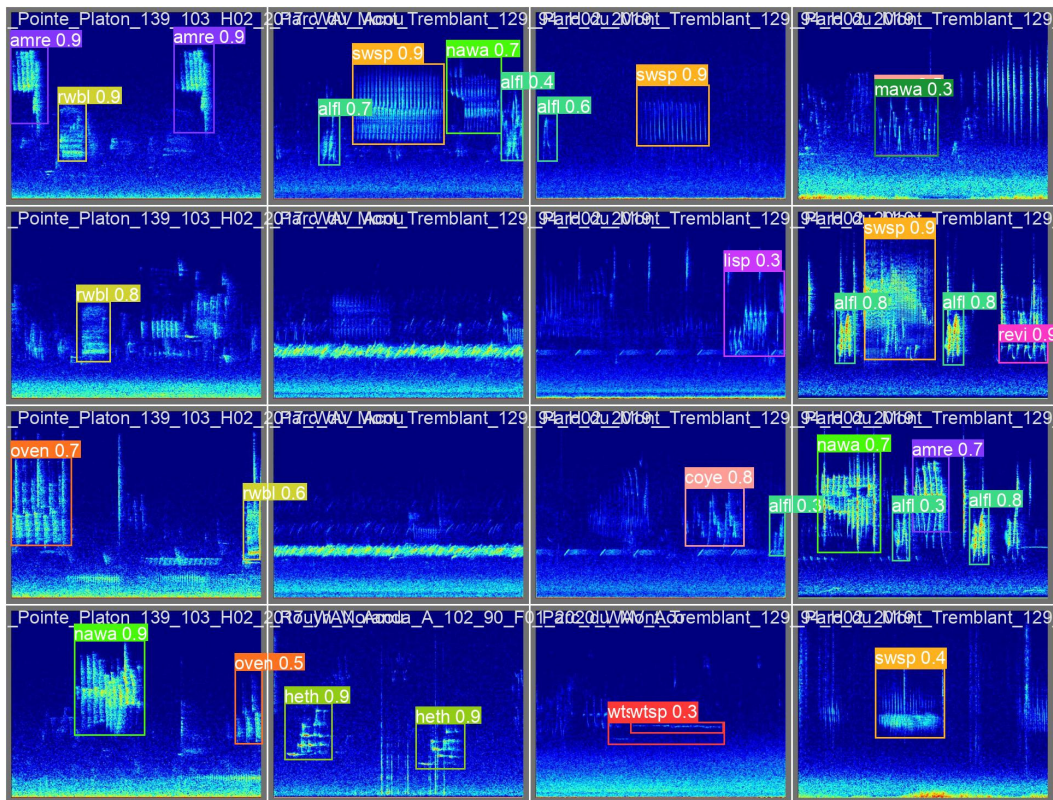


Figure 5. Prédications d'oiseaux par réseau sur des spectrogrammes

Exemple des prédictions avec placement des bounding box sur des spectrogrammes non annotés. La valeur suivant le code de l'espèce en couleur représente la confiance du réseau sur sa prédiction.

De plus, comme le montre la Figure 5, les détections sont bonnes malgré la superposition de certaines espèces. En effet cela représentait un problème très complexe et cette nouvelle méthode permet d'améliorer les détections dans ces conditions. Il est important de noter que les bounding box tracées par le modèle sont bonnes et prennent bien en compte l'intégralité de la vocalise. C'est notamment cela qui permet au modèle de permettre une superposition des détections.

Dans le cas des orthoptères, les scores de classifications sont compris entre 79 et 100% et la confusion est inférieure à 14% comme le montre la figure 7. Les espèces qui sont bien classifiées sont principalement scpi et stgc, suivies par untk. La confusion entre sccu et stgc peut s'expliquer par le fait que ces deux espèces ont un motif proche mais avec un rythme différent, tandis que la confusion entre sccu et untk s'explique car les deux espèces ont un motif différent mais avec un rythme proche. Comme dans le cas des oiseaux, la forte prédiction dans le background peut signifier que de nouvelles détections non labélisées ont été faites. Ce modèle permet donc de détecter des insectes de manière assez précise. Il a cependant été observé que les espèces untk et unco sont des espèces inconnues de ticker et de *Conocephalinae*. La difficulté avec les insectes est en grande partie la variation en fréquence des sons selon la météo.

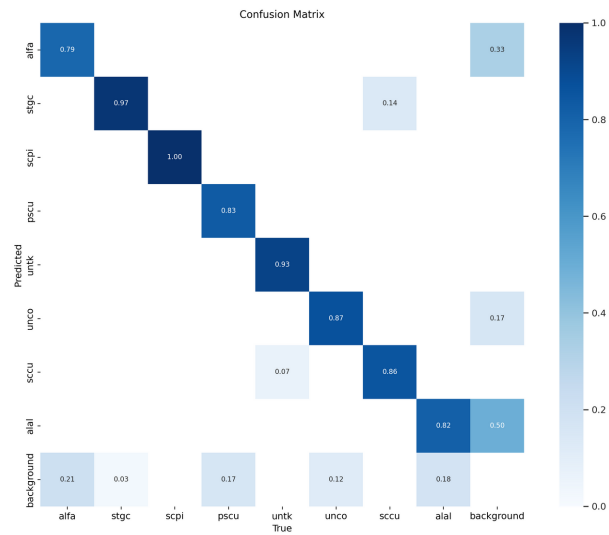


Figure 7. Matrice de confusion de la validation du réseau après un entraînement sur les sons d'orthoptères

Sur cette figure sont représentées les valeurs de prédiction du réseau en fonction des vrais labels. La colonne tout en bas et tout à droite représente le bruit des enregistrements.

En effet il a été vu avec les spécialistes qu'un son peut varier en fonction de la température, passant notamment de 4.5kHz à 8kHz avec un écart de 14°C. Il serait donc intéressant de connaître la température au moment de l'enregistrement afin de mieux traiter les détections.

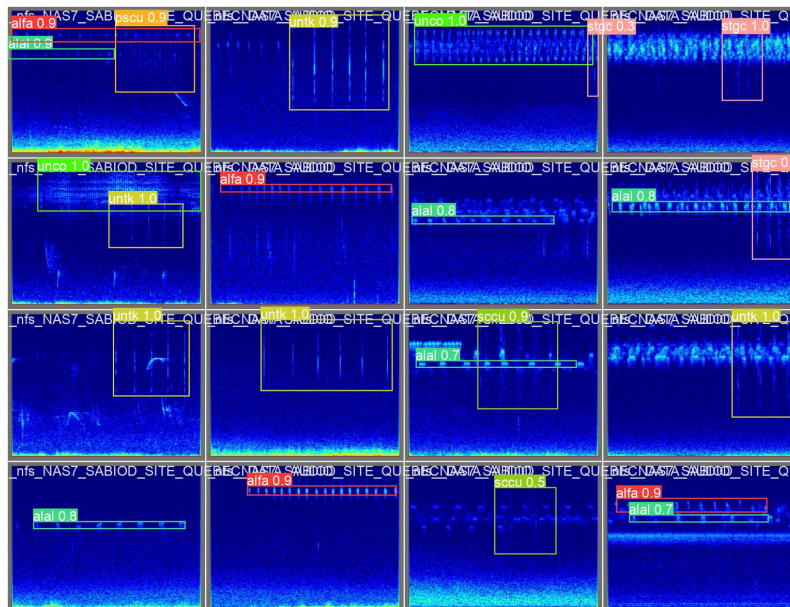


Figure 8. Prédications d'orthoptères par le réseau sur des spectrogrammes

Exemple des prédictions avec placement des bounding box sur des spectrogrammes non annotés. La valeur suivant le code de l'espèce en couleur représente la confiance du réseau sur sa prédiction.

3.2.

Suivi de la présence d'espèces en fonction du temps et le la localisation

En passant le modèle sur les 8 premières minutes de tous les enregistrements de 38 stations, on peut donc tracer la courbe de détections en fonction du temps et donc tenter de mettre en évidence la présence de certaines espèces sur différentes stations et mettre en relation la localité et cette présence. En effet la figure 9 montre l'exemple de 4 stations parmi les 38 sélectionnées. Dans la station humide du site Jamnesie comme celui d'Autray, l'espèce la plus détectée est wtsp.

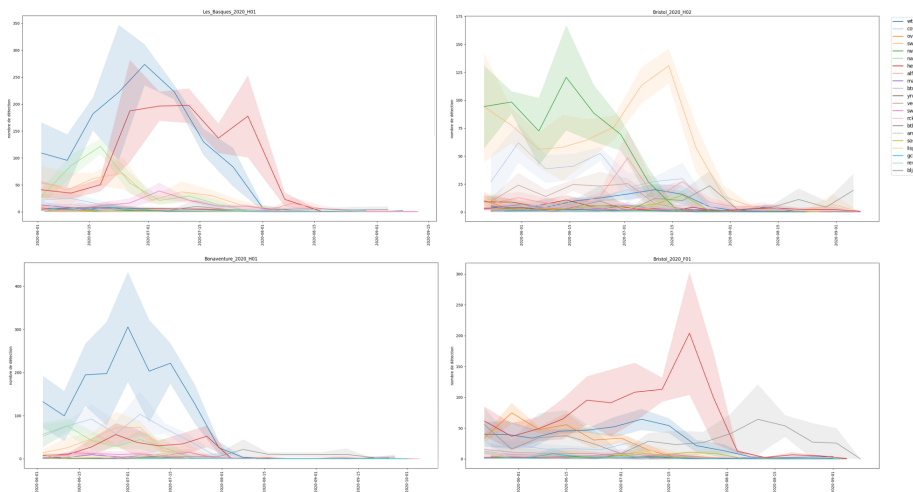


Figure 9. Évolution des détections d'oiseaux en fonction du temps sur 4 stations différentes.

Comparaison du nombre de détection en fonction du temps sur la station humide Les Basques, Bonaventure ainsi que la station forestière et humide de Bristol 2020. Calcul réalisé en prenant la moyenne de la semaine avec mesure de l'écart-type.

C'est-à-partir du milieu du mois d'août que les détections diminuent, quelle que soit la station. En effet cela est corrélé avec le fait que les oiseaux sont surtout présents au printemps. Cependant, on observe que dans les stations humides et forestières de Bristol, les détections de blja sont en hausse à la période où les autres espèces semblent être absentes. Ceci est également observé sur la seconde station humide d'Abitibi Ouest B.

En ce qui concerne la détection des orthoptères, le même travail a permis de réaliser des graphiques similaires mais comme le montre la figure 10, qui présente à gauche les détections d'oiseaux sur deux stations et à droite les détections d'orthoptères sur deux stations, il semblerait que c'est à partir d'août que les détections sont les plus nombreuses. Ce résultat peut alors être mis en corrélation avec le résultat précédent et donc supposer un lien entre la présence d'oiseaux et celle des orthoptères. De plus, lorsque l'on regarde pour la même station, Abitibi Ouest B, ce pic d'orthoptères est bien visible au début du mois d'août, là où les détections d'oiseaux sont très basses. Il est cependant possible que cela soit corrélé avec la température qui, en effet, atteint son maximum à cette période et qui a été au dessus des normales de saison durant le mois d'août 2020.

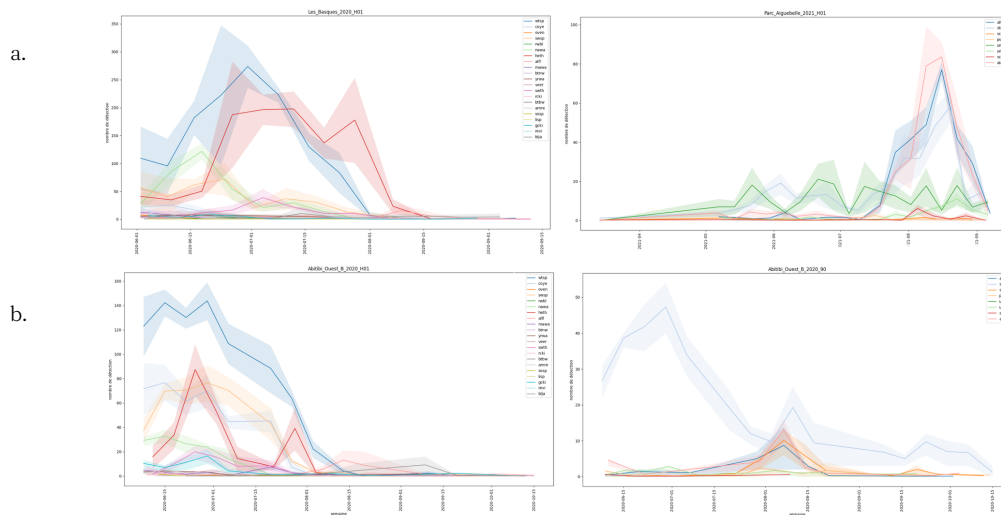


Figure 10. Évolution des détections d'oiseaux et d'orthoptères en fonction du temps sur 3 stations différentes.

Comparaison du nombre de détection en fonction d'oiseaux, à gauche, et d'orthoptères, à droite entre 2 stations différentes (a) et deux stations identiques (b).

3.3. Suivi des indices acoustiques

Concernant les indices, il est souvent pris comme référence l'ACI, pour mettre en évidence la diversité des espèces d'oiseaux dans un enregistrement. En calculant les différents indices sur l'intégralité des enregistrements de chaque station, on obtient donc la figure 11 qui montre l'évolution de l'ACI en fonction du temps et sur différentes stations. On observe que la valeur d'ACI des jours du mois d'août sont plus faibles qu'en juillet par exemple, et cela pour toutes les stations. Ces faibles valeurs sont donc en corrélation avec les faibles détections de YOLO pour cette même période.

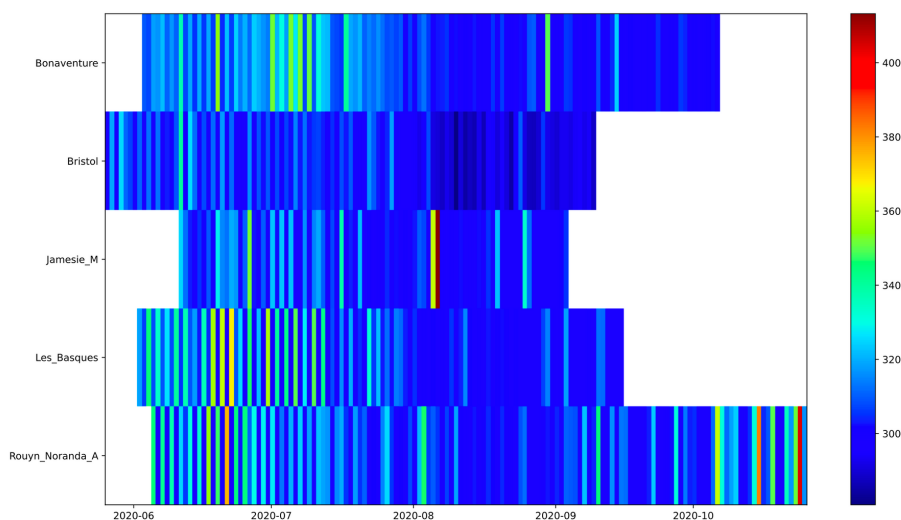


Figure 11. Valeurs d'ACI sur 5 stations différentes en fonction du temps
Évolution de la valeur médiane d'ACI par journée, matin, soir et nuit confondus, sur 5 stations. La valeur d'ACI est comprise entre 280 et 410.

En prenant comme date le 1er juillet dans la station de Bonaventure, on s'aperçoit qu'à cette date la valeur d'ACI semble être la plus haute de toute la session d'enregistrement. La date du 1er juillet est également celle avec le plus de détections de cette session d'après la figure 9, avec plus de 600 détections et au moins 6 espèces différentes. En analysant cette journée et en affichant les différentes mesures d'ACI selon les périodes de la journée, (Figure 12) on observe que les valeurs d'ACI sont dans un premier temps plus haute le matin que le reste de la journée, et qu'elles se situent entre 350 et 390 contre 309 à 317 le 20 août de la même année. Si l'on analyse les valeurs sur une année, on peut voir que les valeurs sont entre 300 et 400 mais que le premier quartile est entre 300 et 310, le second quartile entre 310 et 350 et le troisième quartile entre 350 et 400. Ainsi, cela signifie que la journée du 20 août est une journée qui se situe dans le premier quartile, donc une journée à faible diversité et confirmé par un faible nombre de détection, tandis que le 1er juillet est une journée à forte diversité confirmé encore par un grand nombre de détection.

Ainsi, il semblerait que les deux méthodes soient intéressantes pour répondre à la question et ses dernières se complètent, avec les indices pour mettre en évidence à la fois une forte présence d'oiseaux et également une diversité, et le réseau de neurones à convolution de type YOLO pour la classification et la détection des espèces présentes.

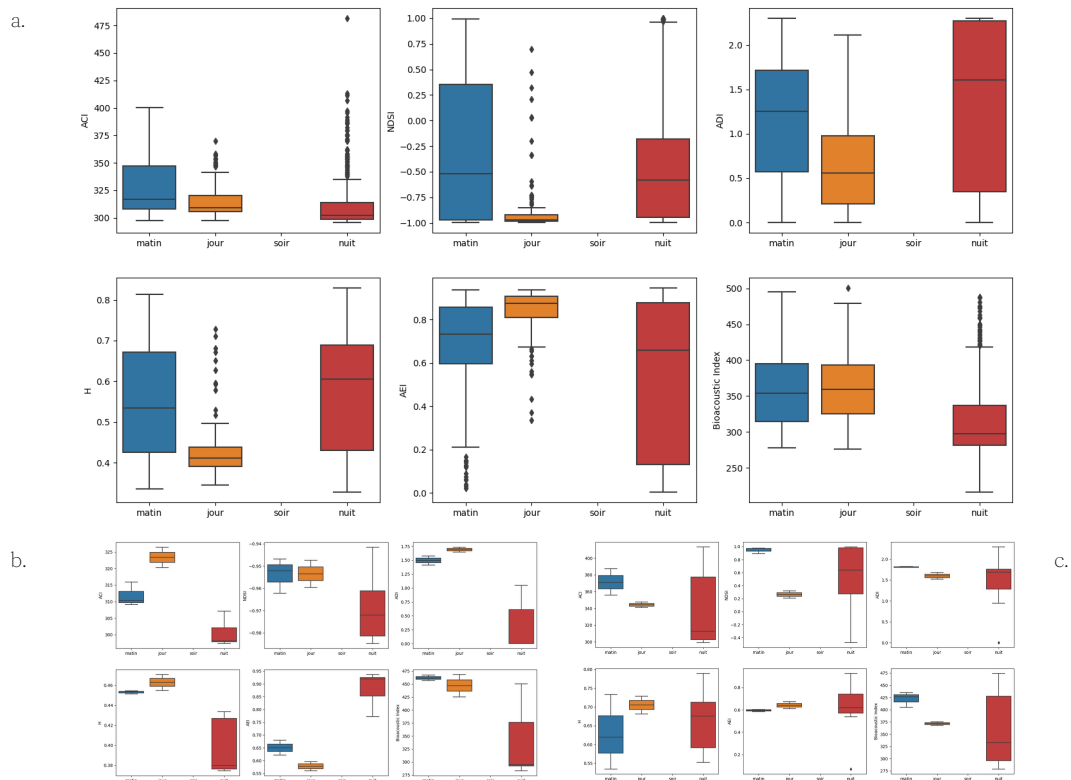


Figure 12. Indices bioacoustiques de la station Bonaventure

Valeur d'indices sur l'intégralité de l'année 2020 (a), pour la journée du 20 août (b) et pour la journée du 1er juillet 2020 (c). Les six indices sont représentés et les périodes de la journée également. Le matin correspond à une horaire d'enregistrement entre 6h et 9h, le jour entre 9h et 16h et la nuit de 19h à 6h.

4. Discussion

4.1. Fusion des oiseaux et des orthoptères dans un même réseau

L'objectif est d'avoir un détecteur qui permet de d'affirmer la présence d'oiseaux mais également d'orthoptères dans un enregistrement afin notamment de suivre des migrations. Dans ce rapport, il a été réalisé un réseau propre aux orthoptères et un propre aux oiseaux. Il a été cependant essayé de mélanger les deux ordres dans un même réseau mais il semblerait que ceci induit une baisse des scores de classification de chaque espèce au sein des ordres. La forte diversité des vocalises, que ce soit pour les oiseaux ou pour les orthoptères est probablement un facteur important de cette perte d'efficacité. Cependant, dans plusieurs datasets dans lesquels un mélange d'ordres a été effectué, de meilleurs résultats ont été obtenus, ce qui semble suggérer que nous devons apporter des modifications au dataset et essayer d'autres types de configuration des annotations.

4.2. Le résultat de l'indice bioacoustiques ACI semblent corrélés aux détections d'oiseaux

Comme on a pu le montrer dans la partie résultats, les pics d'activités mesurés par les indices bioacoustiques et notamment l'ACI, sont également détectés avec les détections d'espèces par le réseau YOLO. En ce qui concerne les orthoptères, il n'est pas encore très claire d'attribuer un indice à leur présence. En effet il est encore compliqué de détecter les orthoptères en raison de la modulation importante de leur sons induite par les variations de température. Une amélioration du réseau sera donc nécessaire et une modulation de ce dernier pourrait être envisagée si les enregistrements contenaient une information précise de la température.

5. Conclusion

Afin d'apporter une conclusion à ce travail, un outil d'analyse d'enregistrements automatique a été mis au point. Cet outil est capable de détecter 21 espèces d'oiseaux ainsi que 8 espèces d'orthoptères puis donner les valeurs de 6 indices bioacoustiques différents que sont l'ACI, l'ADI, l'AEI, la NDSI, l'entropie de Shannon (H) et BI. La détection des vocalises est faite par l'utilisation d'un réseau de neurones à convolution de type YOLO et les scores concernant les oiseaux sont de 67% de mAP avec une précision de 76%. Concernant les insectes, les scores sont compris entre 79 et 100% et la confusion est inférieure à 14%. Le point important, qui représente une évolution majeure dans le suivi des espèces est la réduction des biais causés par les chœurs. En effet, en utilisant la méthode de YOLO, il est possible de détecter plusieurs espèces qui chantent en même temps et éviter les fausses détections.

6. Outils de dépôt de fichier en ligne et traitement automatisé

Dans le but de simplifier et d'accélérer l'analyse des enregistrements sur le terrain, il a été mis en place un système automatisé sur les ordinateurs du laboratoire. En effet, en utilisant un File Transfert Protocol, noté ftp, il est possible de mettre en place un outil permettant de déposer des fichiers sur un espace de stockage et récupérer les résultats traités par les GPU de l'Université de Toulon. Les résultats de ce processus sont un score des indices décrit dans la partie méthode, ainsi qu'un fichier contenant les détections réalisées par le modèle YOLO. En combinant ces deux types de résultats, il sera ainsi possible de connaître à la fois le niveau d'activité d'un fichier mais également les espèces présentes sur cet enregistrement et le nombre de détection par minute d'enregistrement de chacune.

Pour déposer un fichier/dossier/.zip la procédure est la suivante :

```
sftp oiseaux@saphir2.lis-lab.fr
cd MFFP
cd new
put folder/file/.zip
```

```
mdp : vucde,5Grav
```

Les résultats sont ensuite déposés dans un dossier au nom du fichier déposé sur <http://sabiod.lis-lab.fr/pub/Quebec2/>

Le temps de calcul est estimé à environ 1 minute pour un fichier de 15 minutes et le script tourne toutes les 30 secondes, ce qui permet de prendre en compte tous les nouveaux ajouts rapidement.

Dans la première version de l'outil en ligne, seuls les oiseaux sont pris en compte dans les détections. Sur la page <http://sabiod.lis-lab.fr/pub/Quebec2/>, une page est créée au nom du fichier déposé et contient 3 fichiers :

```
Dans le fichier SMU02932_20210529_063300, il y a : 10 especes
Ces especes sont : ['alfl', 'amre', 'blja', 'coye', 'gcki', 'mawa', 'nawa', 'swsp', 'wtsp', 'yrwa'] et ont ete
compte [2, 3, 4, 2, 1, 4, 33, 21, 81, 6] fois sur les 8 premieres minutes
ce qui represente [0.5, 0.75, 1.0, 0.5, 0.25, 1.0, 8.25, 5.25, 20.25, 1.5] detection par minutes

Les indices sont :
      filename          ACI      NDSI      ADI      H      AEI  Bioacoustic Index
0 SMU02932_20210529_063300.wav 402.856158 0.799093 1.666066 0.611438 0.618593      273.970325
Cet enregistrement represente une forte activiïte biologique
```

Figure 13. Exemple de résultat d'un dépôt d'enregistrement

Cette figure présente la mise en forme des résultats dans le fichier 'Compilation.txt' obtenu après le dépôt d'un fichier audio.

Le premier, 'Compilation.txt' comprend l'ensemble des résultats qui découlent de l'analyse (Figure 1). Le second fichier, 'YOLO_detection__X_Y.csv', X étant le jour de l'analyse et Y le mois. Ce fichier csv correspond au dataframe des détections effectuées par YOLO. Les colonnes sont les suivantes : file; idx; espece; x; y; w; h; conf; annot; midl. La colonne 'file' correspond au nom du fichier traité, 'idx' correspond au numéro de l'image traité par YOLO (la première image, idx 0, est associée au spectrogramme entre 0 et 8 seconde, la deuxième image, idx 1, est elle associée au spectrogramme entre 8 et 16 secondes etc.). La colonne 'espece' comprend le numéro de l'espèce tandis que le nom de cette dernière est présent dans la colonne 'annot'. Enfin, le temps en seconde de la détection est présent dans la colonne nommée 'midl' et correspond au point central de la vocalise. Les colonnes 'x', 'y', 'w', 'h' correspondent aux coordonnées de la détection en ratio de pixels (annotations au format YOLO v5). Le dernier fichier 'results.csv', correspond au tableau des différents indices pour le fichier déposé.

Références

- Aguiar Silva, Claysson de et al. (2022). "Listening in the dark : acoustics indices reveal bat species diversity in a tropical savannah". In : Bioacoustics, p. 1-16.
- Chavin S., Deloustal N, Glotin H. (2023). "Suivi acoustique et identification temps-fréquence par transfert deep learning sur YOLO : gestion des chœurs".
- Fairbrass, Alison J et al. (2017). "Biases of acoustic indices measuring biodiversity in urban areas". In : Ecological Indicators 83, p. 169-177.
- Jocher, G. (2020). YOLOv5 by Ultralytics (Version 7.0) [Computer software]. <https://doi.org/10.5281/zenodo.3908559>
- Metcalfe, Oliver C et al. (2021). "Acoustic indices perform better when applied at ecologically meaningful time and frequency scales". In : Methods in Ecology and Evolution 12.3, p. 421-431.
- Pijanowski, Bryan C et al. (2011). "What is soundscape ecology? An introduction and overview of an emerging new science". In : Landscape ecology 26.9, p. 1213-1232.
- Redmon, J., Divvala, S., Girshick, R., & Farhadi, A. (2016). *"You only look once: Unified, real-time object detection"*. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 779-788).
- Renaud Pierangelo, (2022). Rapport de stage Master 2 Analyse du paysage sonore & indices bioacoustiques après un incendie dans le massif des Maures (Var, France)
- Sánchez-Giraldo, Camilo et al. (2020). "Ecoacoustics in the rain : understanding acoustic indices under the most common geophonic source in tropical rainforests". In : Remote Sensing in Ecology and Conservation 6.3, p. 248-261.
- Sueur, Jérôme et al. (2014). "Acoustic indices for biodiversity assessment and landscape investigation". In : Acta Acustica united with Acustica 100.4, p. 772-781.
- Sevilla, H. Glotin, (2017). *"Audio bird classification with inception-v4 extended with time and time-frequency attention mechanisms"*, WINNER of LIFE CLEF Challenge, Working Notes of CLEF 1866, ceur-ws.org/Vol-1866/paper_177.pdf, 2017.