

ANALYSE DES PAYSAGES SONORES NATURELS PAR INTELLIGENCE ARTIFICIELLE EN LIEN AVEC LES CHANGEMENTS CLIMATIQUES

Jérôme SUEUR (MNHN) / OLIVIER ADAM (INSTITUT D'ALEMBERT) / Nicolas OBIN (IRCAM)

biodiversité, acoustique, intelligence artificielle (IA), apprentissage profond, changements environnementaux

. Contexte scientifique et sociétal

Depuis une décennie, la nécessité grandit de suivre au mieux et de manière non invasive la biodiversité en lien avec les changements globaux, notamment climatiques dont il faut estimer les effets sur les espèces animales et leurs environnements. De nombreuses méthodologies ont été développées, notamment les pièges photographiques (camera trapping), l'imagerie satellitaire (remote sensing), ou encore l'ADN environnemental (meta-barcoding). L'écoacoustique est une discipline récente qui offre des possibilités complémentaires de suivi de la biodiversité en continu sur des échelles temporelles et spatiales importantes. Les enregistreurs audios autonomes permettent d'ores et déjà d'acquérir de manière non-invasive des jeux de données massifs dans des milieux écologiques stratégiques sensibles aux changements climatiques, comme les forêts tropicales, les milieux d'altitude ou les fonds sous-marins. Ces jeux de données contiennent des informations écologiques cruciales sur les changements de la biodiversité qu'il convient d'extraire avec des méthodes objectives et automatisées.

Les avancées menées en intelligence artificielle (IA) appliquée au traitement du signal audio ont permis la réalisation de systèmes d'écoute artificielle de plus en plus performants dans des tâches complexes, comme la détection, la reconnaissance, et la séparation d'événements sonores dans des domaines aussi variés que la parole, la musique, les environnements sonores urbains, et en robotique. Ces avancées montrent depuis quelques années un développement important dans les domaines de l'écologie, comme le prouve la multiplication des compétitions internationales de détection en bio-acoustique (par exemple CLEF, bird audio detection challenge, DCLDE). En particulier, les réseaux de neurones profonds convolutionnels (CNN) rendent aujourd'hui possible la détection d'événements sonores au sein d'environnements réels avec des taux de performance extrêmement élevés. Ils permettent en outre d'extraire les sources sonores détectées par filtrage à partir des masques associés à chacune des sources sonores appris par le réseau. Plus récemment, des avancées importantes ont été réalisées en apprentissage de réseaux de neurones profonds avec des potentialités extrêmement importantes pour la bioacoustique et l'écoacoustique : d'une part, l'apprentissage faiblement supervisé permet d'opérer un passage à l'échelle de l'analyse sur de grandes quantités de données sans nécessiter un surplus de supervision humaine (tri, segmentation, annotation) ; d'autre part, l'apprentissage de « représentation démêlées » (« disentangled representations ») offre la possibilité de décomposer par apprentissage faiblement supervisé les facteurs de variabilités au sein d'un environnement sonore.

L'extension de ces avancées récentes en IA à l'analyse de paysages sonores ouvre de nouvelles perspectives à l'écoacoustique :

- 1) l'implémentation d'algorithmes de réseaux de neurones profonds pour la détection et la reconnaissance d'espèces animales variées et leur exploitation en environnement naturel,
- 2) l'apprentissage faiblement supervisé d'auto-encodeurs CNN et de « représentations démêlées » pour permettre de mieux structurer l'analyse d'environnements sonores complexes (par exemple séparer : activité animale, activité géophonique, et activité humaine),
- 3) l'interprétation de la sortie des algorithmes avec application à l'étude des paysages sonores : par exemple, la réalisation de cartes de saillance acoustique des différents événements sonores, l'implémentation d'indices acoustiques pour la cartographie des paysages sonores, et l'étude de leur distribution spatiale et temporelle à partir de l'analyse de données issus des sites étudiés pendant la thèse.

. Objectif principal

Le projet de thèse aura pour objectif principal de développer, de tester et d'appliquer des mesures d'écoacoustique permettant de séparer par IA les sous-ensembles des paysages sonores sensibles aux changements climatiques. Plus précisément, le projet de thèse visera la réalisation d'algorithmes d'apprentissage profonds faiblement supervisés pour la reconnaissance d'espèces animales et de composants d'un paysage sonore. Cette stratégie permettra de renseigner l'évolution relative des différents ensembles des paysages sonores, information essentielle pour une bonne compréhension et anticipation des effets des changements climatiques.

. Cadre de travail

Les données sonores seront issues de deux projets MNHN long-terme financés mis en place dans deux environnements particulièrement sensibles aux changements climatiques: (1) projet dB@Jura qui, en collaboration avec le Parc Naturel Régional du Haut-Jura, a pour objectif de suivre le paysage sonore du massif forestier du Risoux particulièrement sensible à un potentiel réchauffement climatique en raison d'un micro-climat très froid, (2) projet dB@Guyane monté en collaboration avec le CNRS Guyane qui a pour objectif de suivre la forêt tropicale sur le long terme, forêt également sensible aux écarts thermiques en raison d'une homéostasie thermique (faible amplitude thermique). Les techniques d'analyse seront développées au sein de l'Institut d'Alembert et l'IRCAM, instituts experts en analyse de données audio de toutes natures (vocales, musicales, architecturales).

. Résultats attendus

Ce projet doctoral devrait aboutir à la mise à disposition de nouveaux outils d'analyse des environnements sonores naturels applicables aux deux cas forestiers particuliers. Leur utilisation devrait permettre de documenter pour la première fois les patrons de variations temporels au niveau journalier et saisonnier, servant ainsi de point de référence pour des analyses futures sur plusieurs années. Ces techniques pourront être utilisées dans d'autres contextes écologiques traités par le MNHN mais également par d'autres laboratoires en France et à l'étranger. Le développement d'indices simples de diversité acoustiques par le MNHN a suscité un fort engouement international laissant préfigurer un fort succès pour des outils permettant d'accéder à une information plus fine de la diversité sonore animale. Les nouveaux algorithmes, et indices, permettront de caractériser de manière plus fiable et plus informative les paysages sonores naturels. Les enregistrements et les codes seront mis à disposition. Les résultats seront soumis à des revues internationales d'apprentissage machine, d'acoustique ou d'écologie.

. Faisabilité

Ce projet doctoral repose sur un travail conceptuel, théorique et technique relativement difficile mais tout à fait faisable grâce à un encadrement complémentaire et à la mise à disposition de bases de données déjà acquises ou déjà financées. Les principales difficultés pourraient émerger du stockage des données audio et des temps de calcul. Cependant, des leviers techniques ont été anticipés avec un programme de sauvegarde des données via la sonothèque du MNHN et un accès ouvert au centre de calcul de l'IN2P3 du CNRS.

. Environnement financier

Ce projet s'adosse au projet dB@Jura financé à hauteur de 20 k€/an par le PNR du Haut Jura (convention PNR-MNHN) et au projet dB@Guyane financé par le CNRS Guyane. Par ailleurs, le projet sera soutenu par les crédits récurrents et potentiellement par les appels à projets internes (ISYEB, MNHN, LAM, IRCAM).

. Échéancier

- *Tâche 1 : Collecte de données (T0 – T0+36)* - Tout au long de la thèse, le doctorant participera activement à la collecte de données sur le terrain dans le Jura (échantillonnage), notamment pour enrichir les bases de données qui seront utilisées pour l'apprentissage.

- *Tâche 2 : Étude bibliographique (T0 – T0+6)* - Les six premiers mois se concentreront sur la réalisation d'une étude bibliographique, avec implémentation d'algorithmes état de l'art et premières expérimentations en reconnaissance supervisée d'espèces animales. Un séminaire sera organisé à T0+6 pour présenter l'état de l'art et les directions de recherche envisagées.

- *Tâche 3 : Approche supervisée (T0+6 – T0+16)* - Le premier objectif est d'implémenter une approche supervisée par réseau de neurones convolutifs (CNN) pour la détection et reconnaissance d'événements sonores, avec application au suivi d'espèces emblématiques du Jura (Lynx, Loup, Grand Tétra). La tâche consistera notamment à la réalisation d'un algorithme de détection et de reconnaissance à partir des données récoltées dans le projet dB@Jura. Les défis à résoudre seront la capacité d'apprendre des algorithmes robustes à partir de données accessibles restreintes sur des espèces spécifiques (par exemple par pré-apprentissage sur des bases variées collectées sur internet) et d'éprouver l'efficacité de ces algorithmes sur des données collectées in situ dans des environnements naturels.

- *Tâche 4 : Approche faiblement supervisée (T0+16 – T0+34)* - Le deuxième objectif est de proposer une approche d'apprentissage faiblement supervisée (avec peu ou pas de données annotées) pour la caractérisation d'ensemble sonores flous (activité animale, sons géophoniques, activité humaine) avec application à l'extraction d'indices acoustiques fiables. Cette approche prolongera la tâche supervisée en reprenant des architectures CNN et en apprenant des codes acoustiques de faible dimensionnalité permettant la représentation des paysages sonores (par exemple, avec des architectures de type auto-encodeur). Dans un premier temps, un travail sera mené sur l'interprétation des codes acoustiques appris de manière totalement non-supervisée par exemple en proposant des représentations visuelles de la distribution des ensembles visés dans l'espace des codes acoustiques appris. Dans un second temps, un approfondissement sera réalisé pour contraindre l'apprentissage des codes pour tendre vers l'extraction d'indices acoustiques fiables : par exemple, les codes acoustiques appris pourront être adaptés par « transfert learning » sur des petites bases contenant des exemples des ensembles sonores à décrire.

. Encadrement

Ce projet doctoral bénéficiera d'un triple encadrement complémentaire avec une expertise en suivi de la biodiversité et écologie pour la partie MNHN et une double expertise en analyse du signal et séparation de sources par IA pour la partie INSTITUT D'ALEMBERT/IRCAM. Cette association de compétences en sciences de la vie et traitement du signal audio devrait permettre de réaliser des avancées notables totalement impossibles sans transdisciplinarité.

. Profil attendu

Le candidat devra montrer de fortes compétences en apprentissage machine, traitement du signal, programmation Python et un intérêt majeur pour les questions écologiques.

. Sélection de références récentes des encadrants

Benaroya L, **Obin N**, Liuni M, Roebel A, Rauml W, Argenti S (2018) Binaural Localization of Multiple Sound Sources by Non-Negative Tensor Factorization. *IEEE/ACM Transactions on Audio, Speech and Language Processing*.

Desjonquères C, Rybak F, Castella E, Llusia D, **Sueur J** (2018) Acoustic communities reflects lateral hydrological connectivity in riverine floodplain similarly to macroinvertebrate communities. *Scientific Reports*, 8, 14387.

Doh Y, Delfour D, Augier E, Glotin H, Graff C, **Adam O** (2018). Bottlenose dolphin (*Tursiops truncatus*) sonar slacks off before touching a non-alimentary target. *Behavioural Processes*, 157, 337-345.

Doh Y, **Adam O**, Nolibé G (2018). Medium-term stereophonic recording of humpback whales in Sainte Marie channel, Madagascar: daily variation on whale density. *WIO Journal of Marine Science*, special issue 1/2018, 53-63.

Lopez Marulanda J, Roynette N, Blanchard T, **Adam O**, Delfour, F (2019). Description of a bottlenose dolphin calf's acoustic and visual exploratory behavior towards a non-alimentary complex object. *International Journal of Comparative Psychology*, 32.

Obin N, Benaroya L, Jacques C (2019) Learning structured representation of audio signals with stacked convolutional NMF, *IEEE Journal of Selected Topics in Signal Processing, special issue on Machine Learning for Audio Signal Processing*.

Ulloa JS, Aubin T, Llusia D, Bouveyron C, **Sueur J** (2018) Estimating animal acoustic diversity in tropical environments using unsupervised multiresolution analysis. *Ecological Indicators*, 90, 346 - 355.