

Stage master recherche

Apprentissage profond: propagation & bioacoustique

7 Novembre 2020

Contacts:

Antoine Liutkus (antoine.liutkus@inria.fr)

Ricard Marxer (ricard.marxer@lis-lab.fr)

Hervé Glotin (glotin@univ-tln.fr)

Contexte

Ce stage se situe dans le cadre d'une collaboration entre le LIS (**Toulon, équipe DYN1**¹) et Inria (**Montpellier, équipe Zenith**²). Alors que l'équipe DYN1 s'intéresse à l'applications de méthodes d'IA pour la bioacoustique, l'équipe Zenith travaille sur la gestion de données scientifiques à grande échelle. Les deux équipes ont beaucoup collaboré par le passé, notamment dans le cadre des campagnes internationales d'évaluation LifeClef. Le stage sera co-encadré par Ricard Marxer³, Antoine Liutkus⁴ et Hervé Glotin⁵.

Ce stage de recherche invite à son prolongement naturel en thèse avec la même équipe financée par la chaire d'excellence IA d'Hervé Glotin (ADSIL: ADvanced underSea Intelligent Listening), dont l'objectif est la modernisation des méthodes d'acoustique sous-marine, appliquée notamment un meilleur suivi de la biodiversité des océans, dont la mégafaune, les cétacés, dont DYN1 est expert [10, 1, 2, 3, 4].

¹ <https://www.lis-lab.fr/dyni/>

² <https://team.inria.fr/zenith/>

³ <https://scholar.google.com/citations?user=EloO3uwAAAAJ&hl=en>

⁴ <https://scholar.google.fr/citations?user=0s7V4FAAAAAJ&hl=fr>

⁵ <https://scholar.google.com/citations?user=DqjeizcAAAAJ&hl=en>

1. **Un cadre pluridisciplinaire:** Le projet, nouveau, comprend des prolongements en acoustique sous-marine, en apprentissage profond, gestion de données.
2. **Des applications pour la défense et la biodiversité:** La chaire ADSIL vise à l'exploitation de techniques innovantes pour l'acoustique sous-marine, et notamment pour un meilleur suivi de la biodiversité.

Sujet

Ce stage se focalise sur l'apprentissage profond pour le traitement de signaux issus de l'acoustique sous-marine. Il s'intègre dans la logique du projet ADSIL, qui comprend trois volets principaux, sur lesquels l'équipe DYNI a déjà mené des travaux prometteurs :

1. **Modélisation de propagation bio-acoustique** : mise au point de modèles profonds capables d'approximer de manière convaincante les simulations en éléments finis, exactes mais coûteuses. Une piste prometteuse : les Neural Ordinary Differential Equations.
2. **Classification et détection multicanales** : où il s'agit d'exploiter des enregistrements audio (hydrophones) très bruités pour le suivi de cétacés par acoustique passive en intégrant les modèles de propagation (1).

Dans le cadre de ce stage, le but est d'améliorer les performances sur (2). Cela se fera par l'utilisation de meilleurs modèles profonds et techniques d'apprentissage, pour les deux tâches.

Les enjeux de ce stage sont:

- **L'utilisation de modèle deep en Neural Ordinary Differential Equations [11]** pour investir la question de l'IA en modèle de propagation sous-marine bioacoustique.
- **L'utilisation de réseaux attentionnels (Transformers)** pour la classification. De tels réseaux ont déjà montré leur énorme potentiel en traitement d'image et de la parole, et l'objectif est ici de les utiliser comme de bons candidats pour le débruitage et la classification d'enregistrements pour le suivi des cétacés.
- **L'utilisation de techniques d'auto-apprentissage et de pré-apprentissage** pour la classification des enregistrements, qui ont récemment montré qu'elles permettaient la réduction drastique des quantités de données annotées nécessaires.

Sur chacun de ces points, le but est d'introduire l'acoustique sous-marine auprès de la communauté du Machine Learning comme une tâche ambitieuse aux multiples spécificités. Le travail pourra se traduire par la diffusion de bibliothèques Python dédiées, de corpus publics d'apprentissage, de baselines de systèmes, et enfin bien sûr d'articles de recherche.

Le corpus pour valider les approches existe, est issu des mission Sphyrna Odyssey [10, 12]. Il est composé de trajectoires de plusieurs animaux (cachalots) évoluant de -300 m à -1500 m de profondeur, de 500 m à 3000 m de distance de l'antenne acoustique. Ils émettent donc leurs sources biosonar à plusieurs profondeurs et diverses conditions, données permettant de tester un modèle de propagation et classification jointe.

Les pistes de prolongement de ce stage en thèse sont multiples et dépendent des affinités personnelles du candidat. Le financement de thèse pour ce sujet existe dans la chaire IA ADSIL et attend son/sa candidat/e pour septembre 2021.

Profil

Le candidat doit avoir le profil suivant :

- Être en master 2 recherche ou assimilé, dans un des champs suivants:
machine learning, computer science, mathématiques appliquées, statistiques, traitement du signal.
- Bonne pratique de l'anglais oral et écrit.
- Intérêt pour la programmation. Expérience Python souhaitable.

Les éléments suivants seront appréciés:

- Expérience en PyTorch, ou à défaut en TensorFlow.
- Notions de machine learning et ou de traitement du signal.
- Une implication dans la communauté Machine Learning (fournir compte github).
- **Un intérêt fort pour la science IA appliquée à la biodiversité.**

Déroulement pratique du stage à discuter avec le/la stagiaire.

Références

1. P. Best, et al. "Deep Learning and Domain Transfer for Orca Vocalization Detection." *International joint conference on neural networks* (2020).
2. M. Ferrari, et al. "DOCC10: Open access dataset of marine mammal transient studies and end-to-end CNN classification." *Int. Joint Conference on Neural Networks (IJCNN)*. IEEE, (2020).
3. M. Ferrari, et al. "Wave Propagation in the Biosonar Organ of sperm whales using a Finite Difference Time Domain method." *VIHAR* (2019).
4. M. Poupard et al. "Passive acoustics to monitor flagship species near boat traffic in the unesco world heritage natural reserve of scandola." *Input Academy: International Conference on Innovation in Urban and regional planning* (2019).
5. A. Nugraha et al. "Multichannel audio source separation with deep neural networks." *IEEE/ACM Transactions on Audio, Speech, and Language Processing* 24.9 (2016): 1652-1664.
6. F. Stöter et al. "Open-unmix-a reference implementation for music source separation." (2019).
7. Q. Xu et al. "Self-training and Pre-training are Complementary for Speech Recognition." *arXiv preprint arXiv:2010.11430* (2020).
8. A. Saeed et al. "Contrastive Learning of General-Purpose Audio Representations." *arXiv preprint arXiv:2010.10915* (2020).
9. C. Subakan et al. "Attention is All You Need in Speech Separation." *arXiv:2010.13154* (2020).
10. H. Glotin, "Sphyrna Odyssey Missions, Research Report.", 200p, <http://sabiod.org/pub/SO1.pdf> (2020).
11. T. Ricky et al, "Neural Ordinary Differential Equations." <https://arxiv.org/pdf/1806.07366.pdf> (2019).
12. H. Glotin et al, Démonstration des pistages de cachalots en 3D sur données réelles: http://sabiod.univ-tln.fr/pub/SPHYRNA/3D/SO_Glotin_Thellier_etal_PhyseterAlliance_Monaco_20200114_3DtracksX_Y_Z.mp4 (2020).
13. M. Ferrari, et al. "Wave propagation in the biosonar organ of sperm whales using a finite difference time domain method", *PeerJ Preprints* 7, e27995v1 (2020).
14. M. Poupard et al, "Real-time passive acoustic 3d tracking of deep diving cetacean by small non-uniform mobile surface antenna", *ICASSP IEEE* (2019).