

# Stage M2

## MERLE – Multimodal Effective Representation Learning of Evolution of birds

### Contexte

Les oiseaux (plus de 10 000 espèces aujourd'hui) constituent la classe de Vertébrés Tétrapodes de loin la plus diversifiée. Ils figurent parmi les organismes les plus étudiés dans tous les domaines de la biologie. Des applications en IA de reconnaissance « automatique » se développent d'ailleurs pour aider à l'identification des espèces, sur critères d'apparence visuelle extérieure [2] ou sur les vocalisations [7, 1]. Dans le même temps, des méthodes statistiques apparaissent pour modéliser l'apparence d'éléments du squelette d'ancêtres communs à des clades de vertébrés, à partir de morphologies actuelles et fossiles bien placées sur des arbres phylogénétiques maintenant bien robustes (ces arbres étant fondés sur la phylogénomique) [4]. Ces analyses morphologiques utilisent des landmarks et semi-landmarks (en nombre mais qui simplifient et fractionnent malgré tout la morphologie de façon arbitraire). Or, jusqu'ici rien n'a été tenté sur la modélisation d'ancêtres de clades entiers d'oiseaux en termes d'apparence visuelle extérieure (et encore moins acoustique). Cela comprend le plumage avec ses patterns et ses couleurs, et les parties nues, ainsi que la silhouette générale (au posé ou en vol) qui délimite l'oiseau. Pourtant cela serait d'un intérêt majeur en évolution. Proposer des reconstitutions d'apparences d'ancêtres de divers groupes (jusqu'à l'ancêtre commun de tous les oiseaux actuels) permettrait ensuite de reconstituer l'évolution le long de certaines lignées (jusqu'à l'actuel) de ces caractéristiques, le tempo de cette évolution, ses continuités et discontinuités, en parallèle avec l'évolution de paramètres extrinsèques (évolution du climat, de la paléogéographie etc. au cours du Cénozoïque), susceptibles d'avoir influé sur ces rythmes d'évolution, et intrinsèques (donnés ostéologiques, et aussi des paramètres simples tels que la masse corporelle, l'écologie générale des taxa). Des résultats inédits en termes d'évolution biologique compartimentée (modulaire) et également de réponse des organismes aux contraintes climatiques et environnementales pourront émerger dans le futur si ces expérimentations sont concluantes. Les plumages à l'intérieur d'un clade d'oiseaux peuvent être extrêmement diversifiés et disparates selon les espèces (couleurs, patterns), et de plus il est difficile de modéliser l'apparence totale d'oiseaux. De rares exemples récents d'études se sont bornés à capter les pourcentages de spectres de couleurs présents dans des dessins d'espèces d'oiseaux d'un groupe donné (e.g. [9]) mais sans traduction des détails. Lorsque les détails de plumage sont pris en compte, c'est toujours jusqu'ici de façon discrétisée; ainsi il peut y avoir plusieurs dizaines de paramètres définis de façon forcément arbitraire et traduisant l'apparence de façon fatalement incomplète et fragmentée. Le but du projet dans lequel s'inscrit le stage est de modéliser l'apparence globale mais précise d'espèces d'oiseaux en utilisant les propriétés d'algorithmes d'apprentissage profond de reconnaissance automatique visuelle. Les analyses seront réalisées en se basant sur des ensembles de dessins d'oiseaux tels qu'accessibles sur Birds of the World. Ces dessins de type « guide d'identification » ont plusieurs avantages : les oiseaux sont dépeints de façon très précise et réaliste, calibrée dans des poses comparables et homologues, sur fond neutre et identique (blanc ou pas); contrairement aux difficultés qui peuvent être rencontrées avec des photos. Ensuite, un but sera de tenter de reconstituer l'apparence d'ancêtres communs exclusifs à des ensembles données d'espèces actuelles, en utilisant les arbres phylogénétiques moléculaires disponibles pour tous les clades à tous les degrés de précision nécessaires. Plus spécifiquement, dans ce projet nous allons concevoir et mettre en œuvre des méthodes d'apprentissage avec des réseaux de neurones génératifs pour modéliser l'évolution des espèces d'oiseaux en se basant sur leur apparence.

## Sujet

Le travail du stage consistera en la constitution d'une base d'apprentissage adaptée à l'apprentissage d'un tel modèle avec des images extraites de référentiels comme Birds of the World (dessins), le choix de codage et normalisation de ces données etc. en collaboration avec l'autre stagiaire sur le projet.

Il existe deux grandes familles de modèles de réseaux de neurones génératifs pour des données en grande dimension, les auto-encodeurs variationnels (Variational Auto-Encoder, VAE) [8] et les GAN (Generative Adversarial Networks) [5, 3]. Nous allons nous concentrer sur la première, car ces modèles sont généralement plus faciles à entraîner, et il est plus simple d'intégrer des contraintes externes et d'influencer la "forme" de l'espace latent, par exemple, en le démêlant (disentanglement) [6]. L'objectif serait alors d'utiliser un VAE convolutif pour apprendre une représentation plus compacte des images, et ensuite d'utiliser cette représentation latente en entrée des modèles existants utilisés par des biologistes (avec lesquels M. Louchart collabore déjà) basés sur l'optimisation sous contraintes par des connaissances du métier (phylogénétiques). La question de la fusion de différentes modalités (vision, audition) ou espaces sémantiques (espèces, caractéristiques) sera également étudiée.

## Profil

Les compétences suivantes sont indispensables :

- master en intelligence artificielle / machine learning ou équivalent
- bonne capacité de programmation (Python, Pytorch/Tensorflow)
- autonomie
- curiosité scientifique
- esprit d'équipe

## Durée

Le stage commencera en Février-Mars pour une durée de 5 mois. La date de début sera soumise aux contraintes administratives (environ 6-8 semaines de délai).

## Gratification

3.9€/h, 35h/semaine (i.e. environ 550€/mois)

## Localisation

Laboratoires LGL et LIRIS, Lyon, France.

## Encadrants

- Mathieu Lefort : MCF au LIRIS (<https://perso.liris.cnrs.fr/mathieu.lefort/>)
- Stefan Duffner : MCF-HDR au LIRIS (<http://duffner-net.de>)
- Antoine Louchart : CR CNRS au LGL (<http://lgltp.e.ens-lyon.fr/annuaire/louchart-antoine>)

## Candidature

Merci d'envoyer un CV, une lettre de motivation et les relevés de notes de l'année en cours et précédente à Mathieu Lefort ([mathieu.lefort@liris.cnrs.fr](mailto:mathieu.lefort@liris.cnrs.fr)) et Stefan Duffner ([stefan.duffner@liris.cnrs.fr](mailto:stefan.duffner@liris.cnrs.fr)).

## Références

- [1] R Anand, T Shanthi, C Dinesh, S Karthikeyan, M Gowtham, and S Veni. Ai based birds sound classification using convolutional neural networks. In *IOP Conference Series : Earth and Environmental Science*, volume 785, page 012015. IOP Publishing, 2021.
- [2] Thomas Berg, Jiongxin Liu, Seung Woo Lee, Michelle L Alexander, David W Jacobs, and Peter N Belhumeur. Birdsnap : Large-scale fine-grained visual categorization of birds. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 2011–2018, 2014.
- [3] Xi Chen, Yan Duan, Rein Houthoofd, John Schulman, Ilya Sutskever, and Pieter Abbeel. Infogan : Interpretable representation learning by information maximizing generative adversarial nets. In *Proceedings of the 30th International Conference on Neural Information Processing Systems*, pages 2180–2188, 2016.
- [4] Ellen J Coombs, Julien Clavel, Travis Park, Morgan Churchill, and Anjali Goswami. Wonky whales : the evolution of cranial asymmetry in cetaceans. *BMC biology*, 18(1) :1–24, 2020.
- [5] Ian Goodfellow, Jean Pouget-Abadie, Mehdi Mirza, Bing Xu, David Warde-Farley, Sherjil Ozair, Aaron Courville, and Yoshua Bengio. Generative adversarial nets. *Advances in neural information processing systems*, 27, 2014.
- [6] Irina Higgins, Loic Matthey, Arka Pal, Christopher Burgess, Xavier Glorot, Matthew Botvinick, Shakir Mohamed, and Alexander Lerchner. beta-vae : Learning basic visual concepts with a constrained variational framework. 2016.
- [7] Peter Jancovic and Münevver Köküer. Bird species recognition using unsupervised modeling of individual vocalization elements. *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, 27(5) :932–947, 2019.
- [8] Diederik P Kingma and Max Welling. Auto-encoding variational bayes. *arXiv preprint arXiv :1312.6114*, 2013.
- [9] Eliot T Miller, Gavin M Leighton, Benjamin G Freeman, Alexander C Lees, and Russell A Ligon. Ecological and geographical overlap drive plumage evolution and mimicry in woodpeckers. *Nature communications*, 10(1) :1–10, 2019.