

## **Offre de stage:**

### **Augmentation de données par simulations physiques pour la reconnaissance d'images, application au diagnostic médical.**

L'augmentation de données est une méthode largement utilisée en *machine learning* : au cours de l'apprentissage automatique de l'algorithme, on lui présente non seulement les données réelles, mais également des données simulées obtenues en modifiant celles-ci. Ainsi, on enrichit la base d'apprentissage, en couvrant mieux la diversité des cas rencontrés en pratique, et on peut au bout du compte améliorer les performances de l'algorithme. Cette technique est particulièrement pertinente quand le jeu de données n'est pas de grande dimension.

En reconnaissance d'images (*computer vision*), les méthodes classiques d'augmentation de données consistent en des modifications assez simples de l'image originale, telles que des translations, rotations ou dilatations.

Motivé par une application en diagnostic médical, l'objectif du stage sera de construire un générateur d'images artificielles plus riche, intégrant une modélisation physique de la réponse de l'instrument quand sa configuration évolue. Ce générateur sera intégré à un flux de travail de *machine learning* utilisant des réseaux de neurones convolutifs (CNN) et appliqué à des jeux de données réels, générés dans le cadre d'une collaboration entre bioMérieux et plusieurs partenaires académiques.

Contact : Nicolas Faure

Profil : niveau Master 2 ou Ecole d'Ingénieur. Spécialité en physique, en sciences de la donnée, ou idéalement à l'interface entre les deux.

Compétences clé : Sciences de la donnée/machine learning ; optique physique; programmation en Python. La connaissance des bibliothèques tensorflow et keras est un plus ; ainsi que la programmation en Matlab.

Localisation : Grenoble, Centre Christophe Mérieux, plusieurs déplacements à Lyon à prévoir.

Durée et période : environ 6 mois, printemps-été 2020.