

## Proposition de Stage de Recherche M2/fin de cycle Ingénieur 2019-2020

Lieu : IRIT – Toulouse

Collaboration avec le LPP-Université Paris 3 CNRS et CLILLAC–ARP université Paris Diderot)

<https://www.irit.fr>

<http://www.clillac-arp.univ-paris-diderot.fr/>

<http://lpp.in2p3.fr/>

**Titre** : Techniques de visualisation pour les réseaux de neurones convolutifs appliquées à la phonétique acoustique pour la comparaison de voix

Responsables du stage : Thomas Pellegrini (Université Toulouse III Paul Sabatier, IRIT), Lionel Pibre (Université Toulouse III Paul Sabatier, IRIT), Emmanuel Ferragne (Université Paris 7, CLILLAC-ARP) et Cédric Gendrot (Université Paris 3, LPP)

### Description

Les réseaux de neurones convolutifs ont récemment montré leur efficacité dans le domaine de la reconnaissance du locuteur. Outre leurs performances très élevées, ils s'avèrent novateurs car ils extraient eux-mêmes les descripteurs pertinents, contrairement à des systèmes plus classiques (UBM-GMM, I-vectors, à qui l'utilisateur fournit des paramètres connus pour représenter la variation dans la parole, ces paramètres restant assez opaques dans leur interprétation physiologique (MFCC, LFCC). L'intérêt de l'utilisation de réseaux de neurones convolutifs pour la phonétique acoustique réside donc dans la possibilité de tirer profit de ces paramètres extraits de façon autonome par le réseau en leur fournissant des données très « brutes » (e.g. sous forme d'images : spectrogrammes) afin de mieux comprendre ce qui se trouve dans le signal acoustique et d'identifier la spécificité d'un locuteur.

L'objectif de ce stage consiste à affiner différentes méthodes de visualisation (type CAM, GradCAM, etc.) dans le but de les appliquer à la description phonétique acoustique pour la comparaison de voix.

Ce stage sera intégré au projet ANR Voxcrim Comparaison de voix appliquée à la criminalistique (<https://voxcrim.univ-avignon.fr>). Le projet ANR VoxCrim concerne l'identification des individus par leur voix dans le domaine criminalistique et, plus exactement, la « comparaison de voix » pour la sécurité nationale et l'expertise judiciaire.

Le stage est rémunéré et se déroulera à l'IRIT site UPS (Toulouse).

Profil :

- Master 2 en Informatique
- connaissances avancées en machine learning et en deep learning
- langages : Python et Matlab

Durée : 6 mois maximum

Rémunération : le/la stagiaire recevra la gratification standard (environ 580 euros par mois).

Contacts :

Emmanuel Ferragne : [emmanuel.ferragne@univ-paris-diderot.fr](mailto:emmanuel.ferragne@univ-paris-diderot.fr)

Lionel Pibre : [lionel.pibre@irit.fr](mailto:lionel.pibre@irit.fr)