



## **Réseaux de neurones profonds en auto-apprentissage pour la détection d'événements sonores dans un contexte semi-supervisé.**

### **Information générales:**

Thématique : Intelligence Artificielle, Apprentissage profond, Traitement du signal  
Poste : Stage de fin d'études (M2 ou Ingénieur)  
Durée : 6 mois  
Début : Février / Mars  
Lieu : IRIT - UPS - 118 Route de Narbonne, Toulouse

### **Contexte**

L'équipe SAMoVA travaille sur l'analyse, l'indexation et la structuration de documents vidéo et audio. Elle a proposé différents indicateurs acoustiques pour l'analyse sonore dans des milieux bruités. Ces dernières années, elle s'intéresse également au Deep Learning et met en œuvre des réseaux de neurones profonds pour des tâches variées. Le projet ANR LUDAU met l'accent sur les techniques d'apprentissage semi-supervisée. Ces techniques permettent l'utilisation d'un grand volume de donnée qui est impossible à annoter entièrement.

### **Enjeux et objectifs du stage**

Dans ce stage, nous nous intéressons à la détection d'évènements sonores (par exemple des klaxons ou des chants d'oiseaux) enregistrés dans un environnement réel. Ces données audio proviennent de vidéos Youtube auxquelles des extraits de 10 secondes sont sélectionnées. Le volume de données provenant de Youtube est bien trop grand pour être entièrement annoté. De ce fait, les données peuvent être divisés en deux catégories, annotées et non-annotées.

L'objectif du stage est de construire un système de détection d'évènements sonores en utilisant une stratégie d'auto-apprentissage dans un contexte semi-supervisé. Les étapes du stage pourraient être :

- Faire un état de l'art des méthodes d'apprentissage semi-supervisé existantes.
- Concevoir et entraîner un réseau de neurones profond de détection d'événements sonores simple et de manière entièrement supervisé.
- Modifier le réseau pour appliquer une stratégie d'auto-apprentissage de manière supervisée pour commencer, puis de manière semi-supervisée.
- Participer à une compétition internationale (DCASE 2020).

## Encadrement

Thomas Pellegrini

[thomas.pellegrini@irit.fr](mailto:thomas.pellegrini@irit.fr)

Léo Cances

[leo.cances@irit.fr](mailto:leo.cances@irit.fr)

## Références

1. Lu, L., Sun, E., Gong, Y. "Self-Teaching Networks". Proc. Interspeech 2019,
2. C. Leo, P. Thomas, G. Patrice "Multi-task learning and post-processing optimization for sound event detection", TechReport, DCASE 2019