

Séparation de sources en musique par modèles adverses

Mickael Chen et Ludovic Denoyer

2 janvier 2017

Dans la communauté du *Deep Learning*, plusieurs approches récentes s'intéressent au développement de modèles génératifs, c'est-à-dire des modèles capables de capturer une densité complexe à partir de données d'apprentissage, et ce de façon non supervisée. En parallèle de cela, plusieurs variantes de modèles de réseaux de neurones pour la génération de son et de musique (Wavenet par exemple) sont apparus et permettent aujourd'hui de générer des sons d'une qualité proche de la qualité réelle aussi bien en musique qu'en génération de la parole.

A l'intersection de ces deux familles d'approches, nous proposons de nous intéresser à la problématique classique de la séparation de sources dans des données audio, et particulièrement dans des données musicales. L'idée est de concevoir un algorithme permettant d'extraire d'un morceau de musique l'ensemble des phrases sonores issues des différents instruments de musique impliqués. Pour cela, nous nous baserons sur des modèles récents mêlant auto-encodeurs et generative adversarial neural networks (GANs) afin de les appliquer aux données musicales.

Le travail s'organisera en trois étapes :

1. La compréhension des méthodes génératives de type adversarial neural networks, et de leur application à la séparation de sources
2. La prise en main de ces modèles et leur implémentation à l'aide d'une plateforme de développement de modèles de réseaux de neurones profonds (de type Torch – en python)
3. L'implémentation et l'évaluation des modèles proposés par les encadrants

Prérequis : Solide motivation pour la découverte des modèles de Deep Learning, grande capacité de travail.

Contact : ludovic.denoyer@lip6.fr