

Réseaux de neurones du second ordre pour l'analyse de scènes audio

Olivier Schwander et Matthieu Cord

11 décembre 2018

Localisation et encadrement

Localisation :

- Équipe MLIA, LIP6, Sorbonne Université

Encadrants :

- Olivier Schwander <olivier.schwander@lip6.fr>
- Matthieu Cord <matthieu.cord@lip6.fr>

Contexte

Comme beaucoup d'applications en apprentissage automatique, l'analyse de données audio a beaucoup bénéficié du développement des méthodes utilisant l'apprentissage profond. Cependant, ce bénéfice n'est vraiment valable que lorsque suffisamment de données étiquetées sont disponibles. Ce qui est peut-être un facteur limitant pour beaucoup d'applications. On s'intéresse ici à un cadre différent des réseaux de neurones classiques, baptisé réseaux de neurones du second ordre[3]. Ces réseaux s'intéressent à des statistiques du second d'ordre, des matrices de covariances calculées sur les données d'entrées. Les différentes couches de ces réseaux exploitent la géométrie riemannienne de ces matrices pour réaliser des transformations pertinentes. Ce genre de méthode a montré une bonne robustesse aux faibles quantités de données dans différentes applications (EEG[1] et radar). Or l'analyse de scènes audio est justement confronté à des faibles quantités de données, ou à un étiquetage limité qui oblige à se placer dans un contexte semi ou faiblement supervisé.

Objectifs

L'objectif de ce stage est d'étudier des architectures du second ordre adaptées pour la classification de scènes audio. Une extension sera de s'intéresser à la localisation et à la segmentation d'évènements en s'inspirant des techniques de

vision[2]. On visera notamment une participation à l'édition 2019 du défi DCASE (voir ¹ pour l'édition 2018).

Références

- [1] Marco CONGEDO, Alexandre BARACHANT et Rajendra BHATIA. “Riemannian geometry for EEG-based brain-computer interfaces ; a primer and a review”. In : *Brain-Computer Interfaces* 4.3 (2017), p. 155-174.
- [2] Thibaut DURAND et al. “WILDCAT : Weakly Supervised Learning of Deep ConvNets for Image Classification, Pointwise Localization and Segmentation”. In : *The IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. 2017.
- [3] Zhiwu HUANG et Luc J. VAN GOOL. “A Riemannian Network for SPD Matrix Learning.” In : *AAAI*. T. 1. 2017, p. 3.

1. <http://dcase.community/challenge2018/>