

Classification de scènes audio - Participation au challenge DCASE

Olivier Schwander olivier.schwander@sorbonne-universite.fr

2024

Malgré les progrès récents en apprentissage profond, la classification de données audio reste un problème difficile et non résolu de façon satisfaisante. Une partie des limites actuelles vient du manque de supervision, ce qui impose d'utiliser des modèles adaptés à de faibles quantités de données supervisés. Les techniques possibles sont par exemple :

- l'augmentation de données (transformations simples ou GAN),
- les connaissances a priori et les pré-traitement adaptés (représentation temps-fréquence),
- l'apprentissage faiblement ou semi supervisé (mélange de données étiquetées et non-étiquetées, ou étiquetage moins précis),
- modèles plus robustes exploitant mieux la structure des données (réseaux de neurones du second ordre).

Le challenge DCASE¹ offre un ensemble de tâches centrées autour de la classification de données audio. Le stagiaire pour mettre un œuvre à la fois des techniques classiques de deep learning et des méthodes plus originales et prometteuses pour ces données.

1. <https://dcase.community/workshop2024/index>