
Approches frugales pour la classification de scènes audio

Olivier Schwander olivier.schwander@sorbonne-universite.fr

2025

Ce projet s'intéresse à l'une des tâches proposées par le challenge DCASE 2024[1]

L'objectif est de construire des modèles compacts mais performants pour l'analyse de scènes audio. Il s'agit d'une contrainte critique pour les systèmes embarqués (téléphone ou plus petit encore) qui sont par nature limités en puissance de calcul et en consommation énergétique (en particulier sur batterie). Au delà de l'application aux systèmes embarqués, ce type d'approches frugales est également intéressant d'un point de vue écologique pour limiter le coût en ressources (fabrication des composants, consommation électrique, refroidissement des serveurs, etc) des modèles et donc l'impact des méthodes d'apprentissage sur la planète.

Après une analyse des méthodes de l'état de l'art qui ne se préoccupent pas de la taille des modèles, nous nous intéresserons à des modèles de plus en plus compacts et nous ferons en sorte de maintenir leur bonne performance. Pour ce faire, nous étudierons entre autre des méthodes de distillation et d'élagage. Ce projet permettra également de mettre en œuvre un protocole expérimental qui ne s'intéresse pas que aux performances mais classe les modèles en fonction d'un compromis avec l'occupation mémoire et le temps de calcul.

En fonction du temps disponible et de l'avancement, nous pourrons envisager une participation au challenge 2025 (qui n'est pas encore publié pour le moment).

[1] <https://dcase.community/challenge2024/task-data-efficient-low-complexity-acoustic-scene-classification>