

Pré-Rapport de Projet

Détection et classification de sons
à l'aide de Spiking Neural Networks

COURREGE Téo
GANDEEL Lo'aï

Date de remise : November 15, 2023

1 Introduction

Les réseaux neuronaux à impulsions ou "spiking neural networks" (SNN), sont une classe de modèles neuronaux bio-inspirés qui cherchent à reproduire le fonctionnement du cerveau en utilisant des impulsions électriques discrètes. Contrairement aux réseaux de neurones artificiels traditionnels, qui traitent l'information sous forme de valeurs continues, les réseaux neuronaux à impulsion tentent de capturer la dynamique des signaux neuronaux réels. Ces modèles se sont avérés particulièrement pertinents pour des applications liées à la modélisation de systèmes biologiques, à l'apprentissage supervisé et non supervisé, à la reconnaissance de motifs, à la robotique et à d'autres domaines de l'intelligence artificielle. L'intérêt croissant pour les réseaux neuronaux à décharge repose sur leur capacité à traiter l'information de manière asynchrone, à une échelle temporelle, ce qui les rend adaptés pour la résolution de problèmes complexes qui impliquent la perception sensorielle en temps réel et le traitement de données temporelles.

2 Présentation des objectifs

Nos principaux objectifs seront de comprendre leur fonctionnement ainsi que de construire l'un de ces réseaux afin de pouvoir détecter et classifier des sons, et évaluerons les différences entre eux (nous nous focaliserons en particulier sur les différences en puissance de calcul et utilisation énergétique).

3 Plan et Planification

- Dans un premier temps, il sera nécessaire de décrire le fonctionnement d'un SNN, les avantages et inconvénients qu'il comporte.
- Le dataset principal utilisé pour ce projet sera "Audioset," un dataset se basant sur des vidéos YouTube (2 millions de vidéos, 527 labels). Le dataset étant très grand, nous devons donc récupérer juste une partie des vidéos pour les labels que l'on choisira, puis extraire la partie audio qui nous intéresse à partir des vidéos.
- Il s'agira ensuite de créer, puis d'entraîner notre SNN en détaillant nos choix lors de sa construction. Nous essaierons aussi de trouver un modèle pré-entraîné afin de le comparer au nôtre.
- Nous comparerons les différents résultats entre notre SNN et d'autres modèles de réseaux de neurones.

4 Références

- YAMnet - <https://github.com/tensorflow/models/tree/master/research/audioset/yamnet>
- Dataset Audioset - <http://research.google.com/audioset/index.html>
- Copyright pour l'utilisation de vidéos YouTube - <https://www.quora.com/Is-it-legal-to-use-YouTube-videos>

- Téléchargement audio pour la préparation d'un dataset de reconnaissance vocale - <https://mlearning.ai/prepare-speech-recognition-dataset-using-youtube-videos-5361ae3d5077>
- Modèle pré-existant : Reconnaissance audio - https://link.springer.com/chapter/10.1007/978-981-10-5230-9_57
- Speech recognition (un peu éloigné mais peut donner des idées) - <https://arxiv.org/pdf/1911.08373.pdf>
- Autre - <https://core.ac.uk/download/pdf/288002964.pdf>
- Autre - <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6987407/>