

Classification de Sons

Clément Gousseau

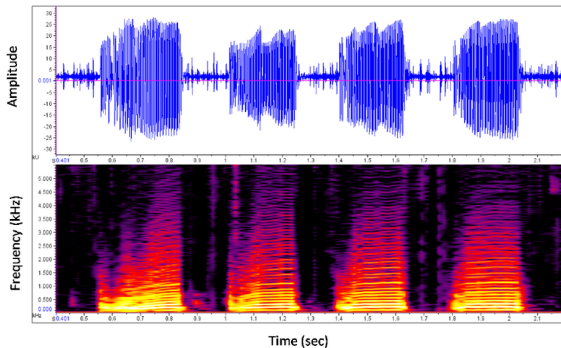
2 août 2019

Plan de l'exposé

- 1 Pre-Processing : construction de log-melspectrogrammes
- 2 Classification

Motivation

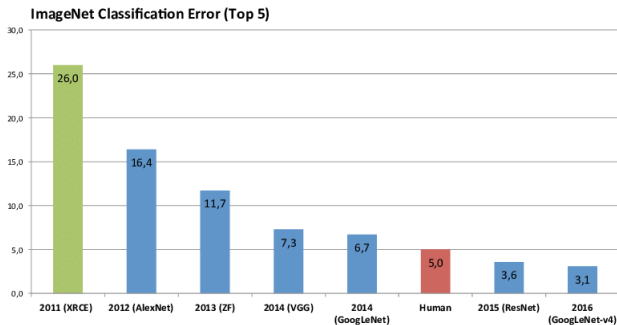
spectrogramme : image (signal 2D) représentant un son (signal 1D)



Pourquoi ?

Motivation

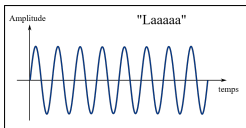
Parce que...



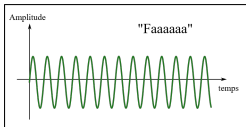
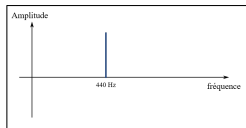
→ très bons résultats des réseaux de neurones artificiels sur la classification d'images

Transformée de Fourier

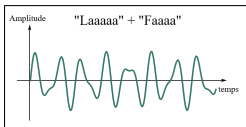
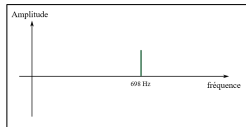
domaine temporel $\xrightarrow{\text{Fourier}}$ domaine fréquentiel



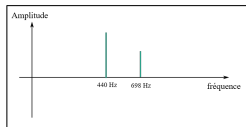
Transformée
de Fourier



Transformée
de Fourier



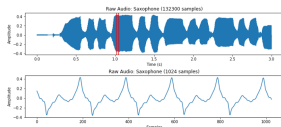
Transformée
de Fourier



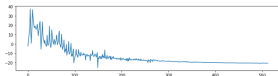
du Signal au Spectrogramme

On extrait un morceau du signal à l'aide d'une fenêtre glissante :

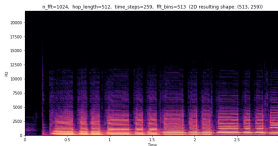
hop length : number of samples between windows ; *nfft* : window length ; *nmels* : résolution fréquentielle.



On calcule le spectre de Fourier sur cet extrait :

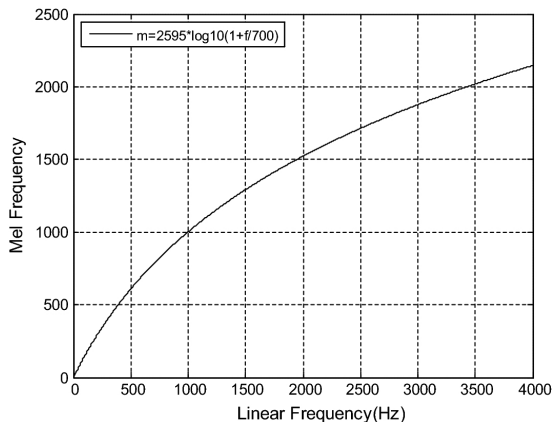


On juxtapose les spectres de Fourier selon l'axe temporel :



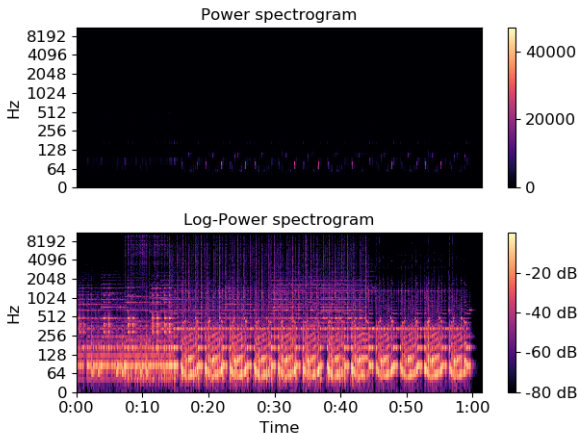
du Spectrogramme au Melspectrogramme

échelle de fréquence basée sur la perception de l'oreille humaine



du Melspectrogramme au log-Melspectrogramme

échelle log pour l'amplitude \rightarrow meilleure représentation



Jeu de données et Tâche

Sons de la base de données UrbanSound8K

8372 sons de moins de 4 secondes appartenant aux classes :

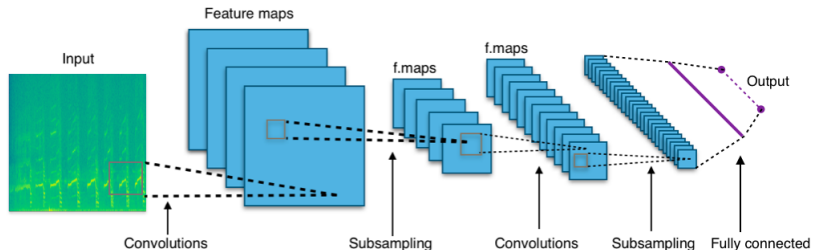
- 0 : air_conditioner
- 1 : car_horn
- 2 : children_playing
- 3 : dog_bark
- 4 : drilling (forage)
- 5 : engine_idling
- 6 : gun_shot
- 7 : jackhammer (marteau-piqueur)
- 8 : siren
- 9 : street_music

Chaque son appartient à une et une seule classe.

Pipeline



Réseaux de neurones artificiels



merci