

Classification audio

Projet IA en audio et musique

Son, Musique

- Enormes quantités de données
 - Plusieurs dizaines de millions de titres disponibles (streaming)
 - Multiples versions (youtube) : interprétations, reprises, karaoké, etc.
- Plusieurs représentations
 - Vidéo
 - Audio
 - Paroles
 - Partitions
 - Playlistes
 -

Multiples applications

- Reconnaissance d'instruments, interprètes, compositeurs, producteurs, ...
- Tonalité, Tempo, métrique, accords, ...
- Humeur, ambiance, ...
- Recommandation de titres de musique
- Besoins:
 - Labellisation automatique des sons et des musiques

Machine Learning

- Classification
 - Supervisée
 - Non supervisée

Deep Learning

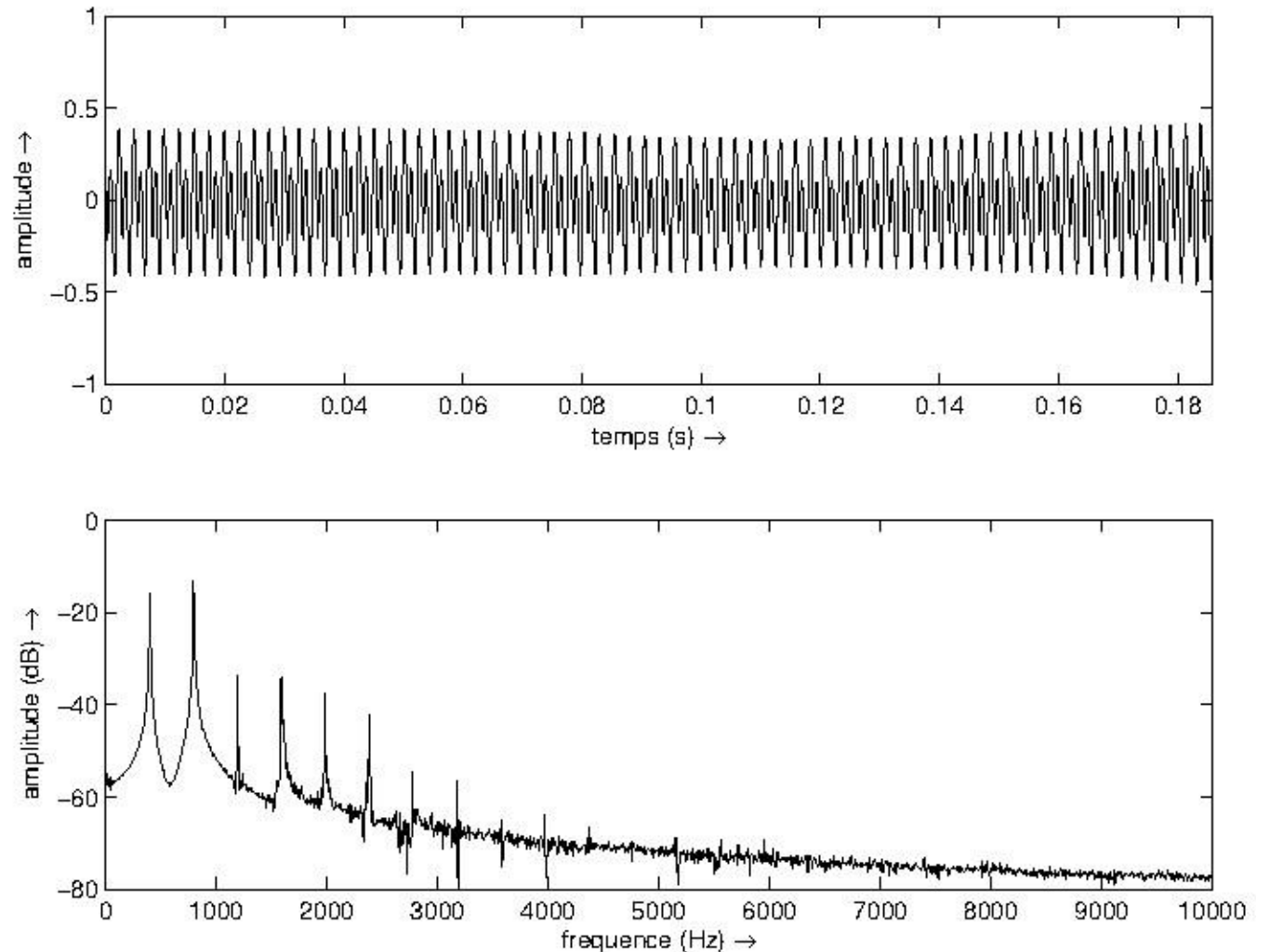
- Etapes
 - Représentations du signal
 - temporelles, spectrales, compressées, ...
 - Une seule valeur, ensemble de valeurs, séquence temporelle, ...
 - Descripteurs (haut niveau, bas niveau)
 - Apprentissage
 - Analyses / Classifications / Evaluations
 - Applications

Représentations

- Bas niveau
 - Représentation temporelle
 - Fréquence échantillonnage (44100Hz ?)
 - Représentation spectrale
 - Echelle linéaire
 - Echelles perception (mel, barks)
- Haut niveau
 - Descripteurs (features)
 - Perceptifs (enveloppe spectrale, énergie, bandes fréquences, ...)
 - Musicaux (hauteurs, volumes, timbre)
 - Autres (MFCC par ex)

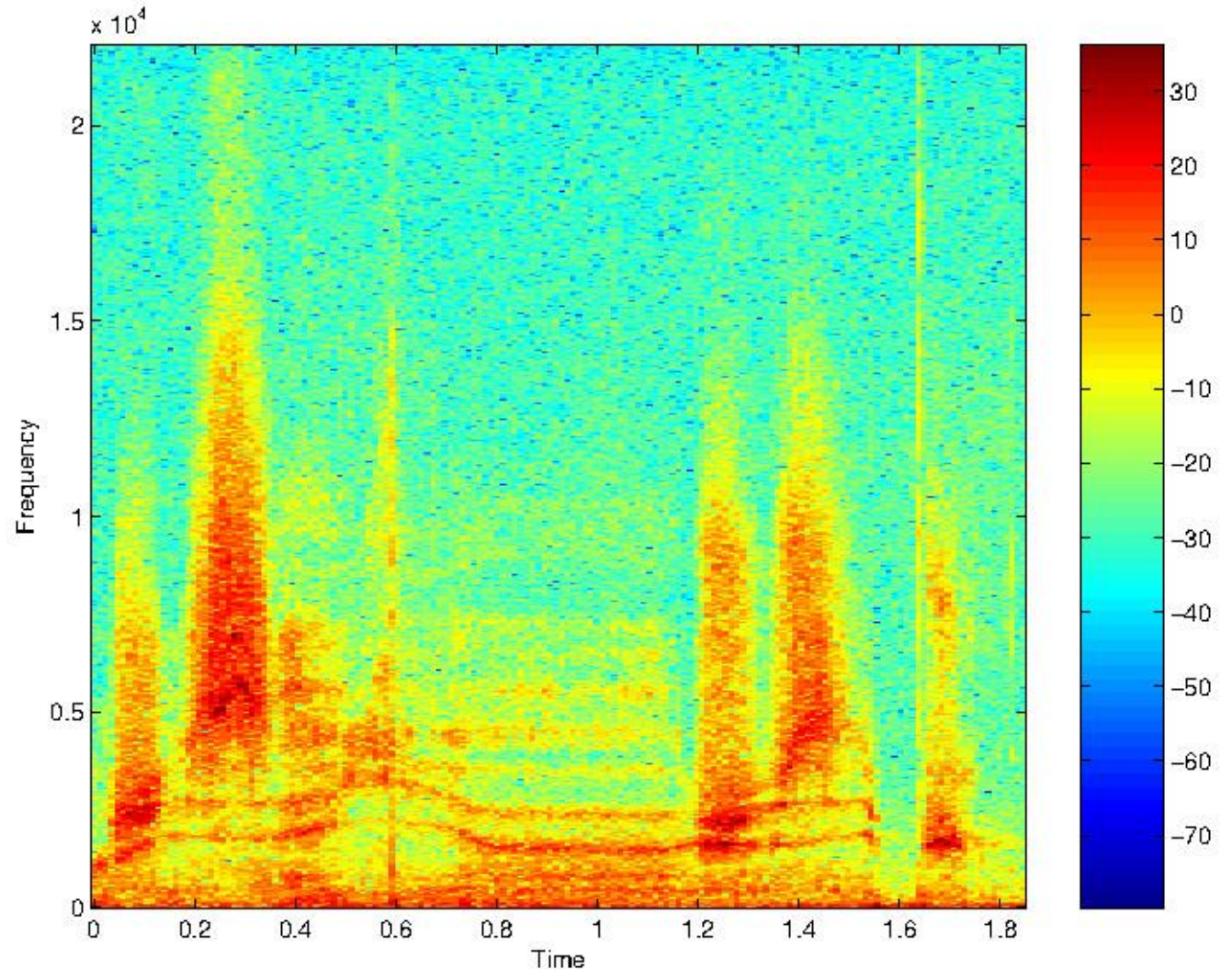
Représentations temporelles/spectrales

- Trames courtes
- Transformée de Fourier
- Pas de perte d'infos
- Deep Learning :
 - A priori pas de différences
 - Adaptation architecture



Représentations 2D

- Echelles différentes
 - Mel, Barks
 - Evolution au cours du tps
- Type Images
 - CNN possibles
 - Formes ?



Python : quelques libs utiles...

- Librosa : traitement du signal audio
- Scikit learn : machine learning
- LightGBM, XGBoost : machine learning (arbres de décision)
- Keras/TensorFlow/Theano : deep learning

Signal : librosa

<https://librosa.github.io/librosa/>

`pip3 install --user scipy`

`pip3 install --user librosa`

`pip3 install --user colorama`

Affichage représentation temporelle d'un son

Affichage d'un spectre d'amplitude

Découpage en trames

Calcul d'un descripteur

Choix des algorithmes ML



Challenge

Classification de morceaux audio selon le genre musical

Challenge

- <https://www.kaggle.com/t/c4d1b6695fe0490b89ba28d2466e91f9>
- 4000 titres répartis en 8 classes de genres
- Format audio (MP3)
- Objectif :
 - Format : challenge kaggle (*il faut s'inscrire*)
 - Tâche : sur un ensemble donné de 4000 titres (sans genre), donner leur genre
 - Evaluation automatique
 - Classement
- Rapport
 - Explications des choix (features, algos, évaluations, métriques)
 - Tests effectués
 - Étude de la précision par classe (ROC curve et AUC pour chaque genre)
 - Matrice de confusion
 - Importance des features
 - Algos Machine Learning, Ensemble ?
 - Décision locale / globale
 - Discussion (commentaires, comment améliorer ?)
- Dataset disponible sur kaggle

Travail demandé

- Choix des descripteurs audio
 - Lib existantes :
 - Librosa <https://librosa.github.io/librosa/feature.html>
 - Essentia https://essentia.upf.edu/documentation/essentia_python_tutorial.html
- Choix des algorithmes de classification
 - Parmi lesquels Deep Learning
 - Ensemble de moteurs
- Choix des paramètres
 - Importance de la démarche
 - AutoML ?
- Augmentation des données ?
- Protocoles d'évaluation
 - Limites sur certains genres, adaptations sur la classif ?

RENDU :

- Fichier CSV sur le format « titre,id_genre », soumis/évalué sur le site kaggle
- Rapport / Code (par email)

Dates

- Rendu le 31 Janvier 2020 (23h59)
- Langages au choix, mais plutôt Python (ou Matlab)

N'hésitez pas à poser des questions.

Ca démarre maintenant !