Identification des cris d’oiseaux Cornell :

Les oiseaux jouent un rôle très important dans la nature , ils sont d’excellents indicateurs de la détérioration de la qualité de l’habitat et de la pollution de l’environnement .

Il est plus facile d’entendre un oiseau que de le voir , en se basant sur ça beaucoup de recherches et de projets ont été fait et sont entrain d’être réalisé en travaillant sur les sons , avec un détection et une classification appropriée du son les chercheurs seraient en capacité de comprendre les facteurs de la qualité de vie en fonction de l’évolution de la population d’oiseaux .

Ce projet de deep Learning sera d’identifier une grande quantité de vocalisations d’oiseaux dans les enregistrements de paysages sonores , la complexité de ce projet est qu’ il y a des étiquettes faibles , ou des sons anthropiques , ou encore des sons d’appels d’oiseaux ou de non oiseaux .

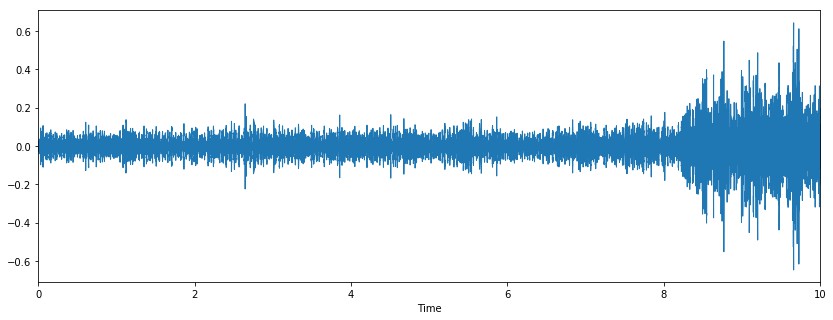
L’objectif de la compétition est de trouver des détecteurs et des classificateurs efficaces pour mieux analyser des enregistrements sonores complexes .

Pour mener à bien ce concours les participants disposent de deux dossiers train\_audio et test\_audio qui contiennent des courts enregistrements pour l’entrainement et le test ainsi que plusieurs fichiers csv qui nous donnent des informations sur les audios .

Lorsque on étudie des audios il est très difficile d’analyser des données dans un état brute , car on les informations sont brouillées .

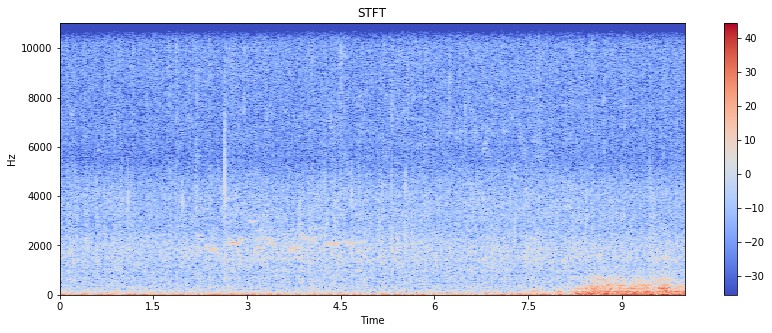
On a utilisé la bibliothèque Librosa pour extraire différentes formes de l’audio ,

Pour représenter ce dont on parle on prend un audio à l’état brute :

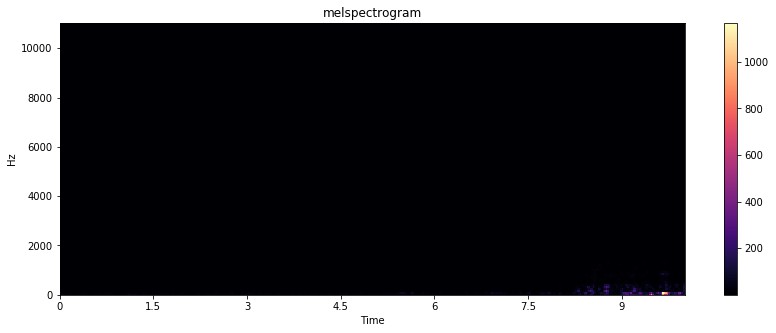


Sous cette forme il est compliqué de faire de bonnes analyses , ça nous permet pas d’identifier et de savoir si le bruit est celui d’un oiseau ou d’une autre chose .

Il est donc nécessaire comme tout autre système d’apprentissage automatique de reconnaissance vocale , d’extraire les caractéristiques , il en existe plusieurs mais on sait intéresser sur 3 en particuliers :



La représentation au-dessus représente la représentation STFT : Short time Fourier Transformation c’est une transformation très connue .



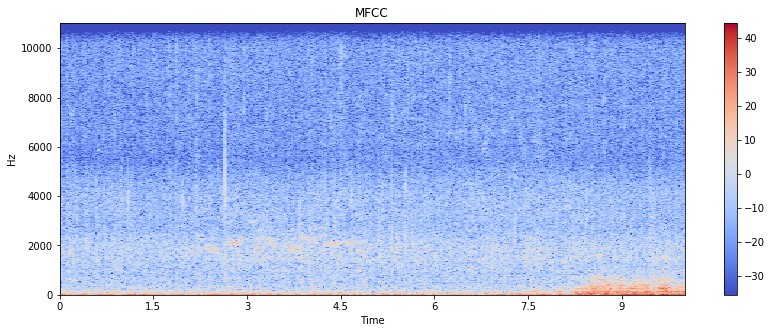
La représentation au-dessus représente la représentation Mel Spectogramme :

**Mel Spectogramme:**

* + C’est une représentation temps-fréquence d’un son.Il est échantillonné en un certain nombre de points autour de temps ti et de fréquences fj équidistants (sur une échelle de fréquence Mel).

La représentation ci-dessous représente la représentation MFCC :

* **Mfcc (Mel-frequency cepstral coefficients)** calculé de la manière suivante:
  + On prends la transformée de Fourier du signal
  + Ensuite on cartographie les puissances du spectre obtenues ci-dessus sur l'échelle mel
  + On prends les logs des puissances de chaque fréquence mel.
  + On effectue la transformation en cosinus discrète de la liste des “log puissances mel”, comme s’il s’agissait d’un signal.
  + Les MFCC sont les amplitudes du spectre résultant



En regardant d’autres projets dans le même domaine , on peut remarquer que la représentation en mel spectogramme est la représentation qui obtient de meilleurs résultats .