

# ISEN

ALL IS DIGITAL!

LILLE



yncréa



Projet M1 - 2017

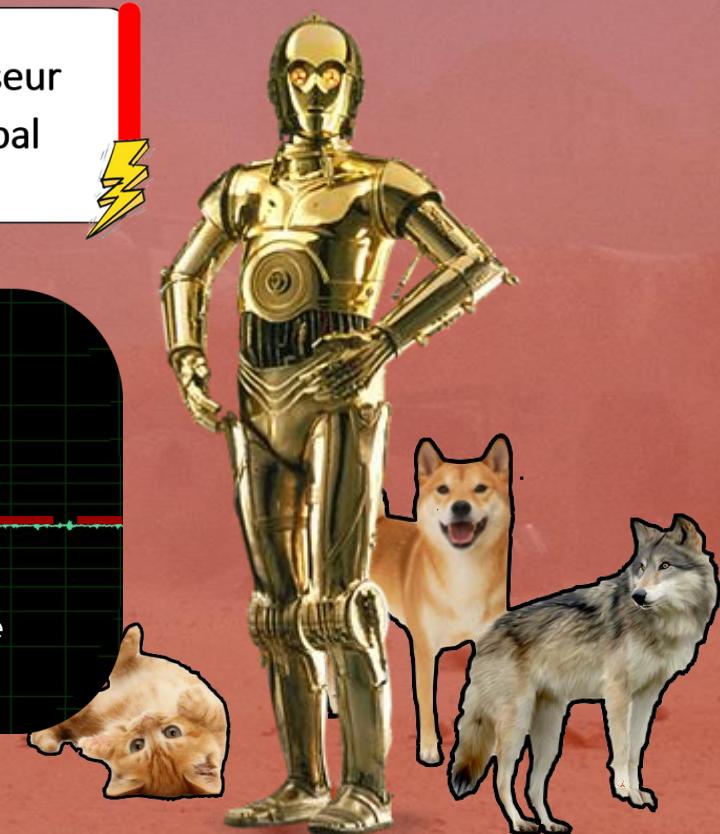
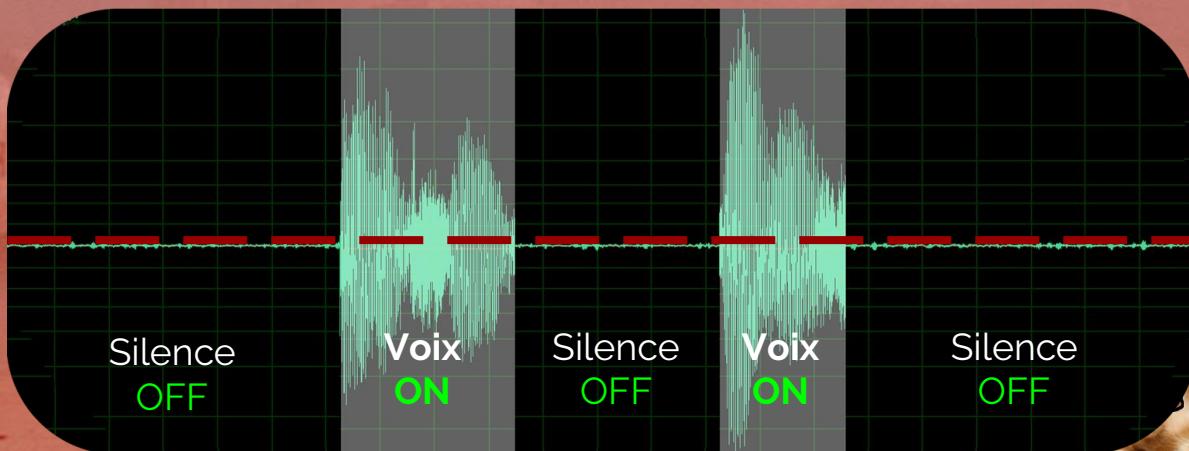
# Étude de la détection automatique de caractéristiques sur un signal audio

Le Bellego Jean & Lecoq Simon

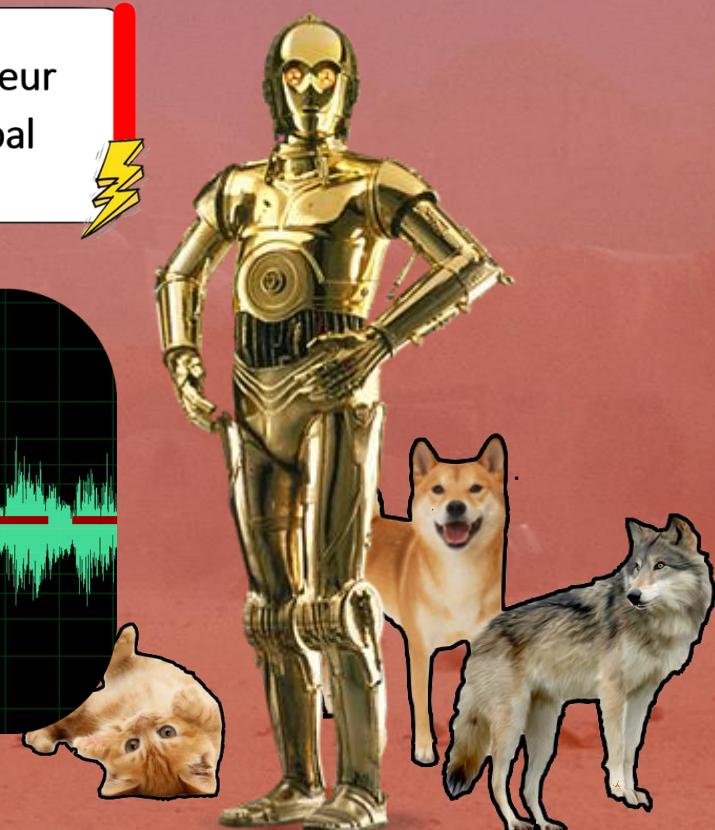
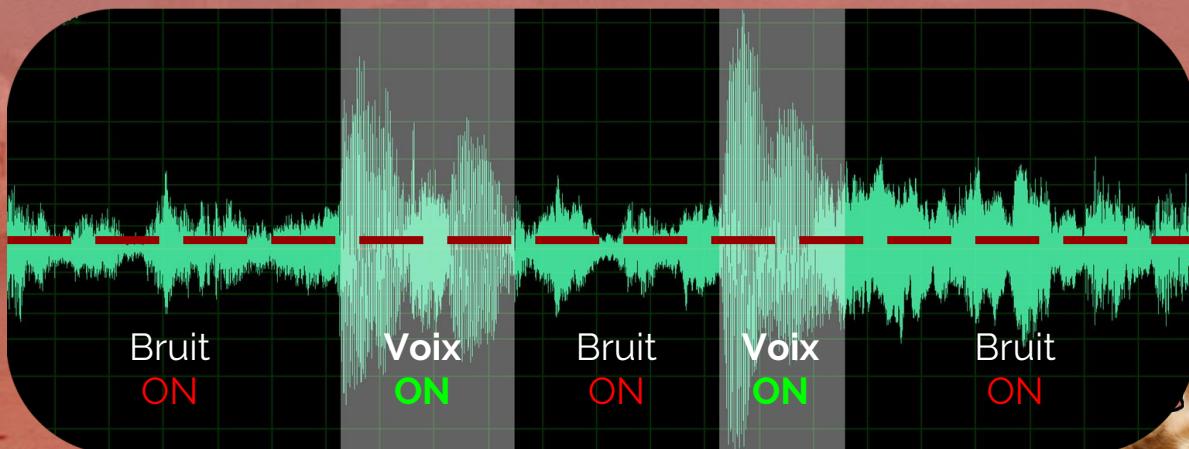
# Pourquoi ?



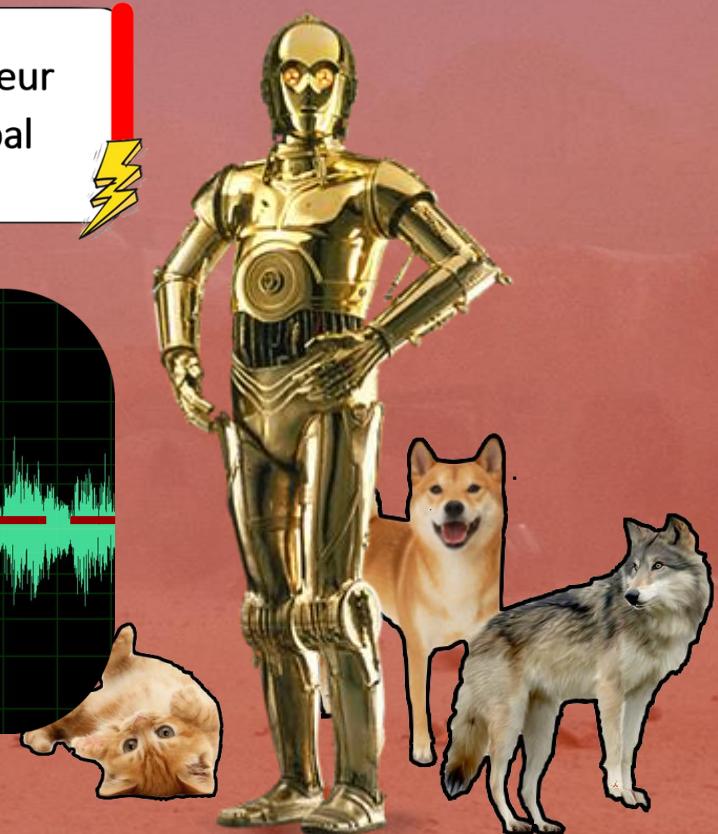
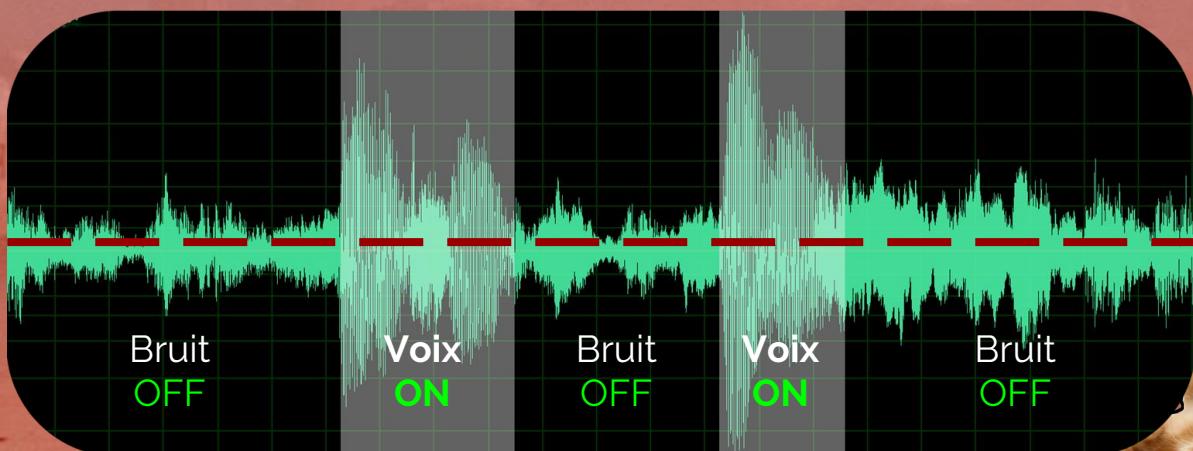
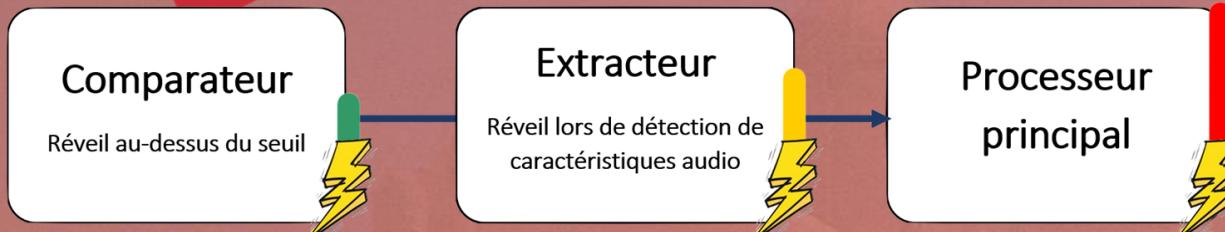
# Wake on threshold



# Wake on threshold



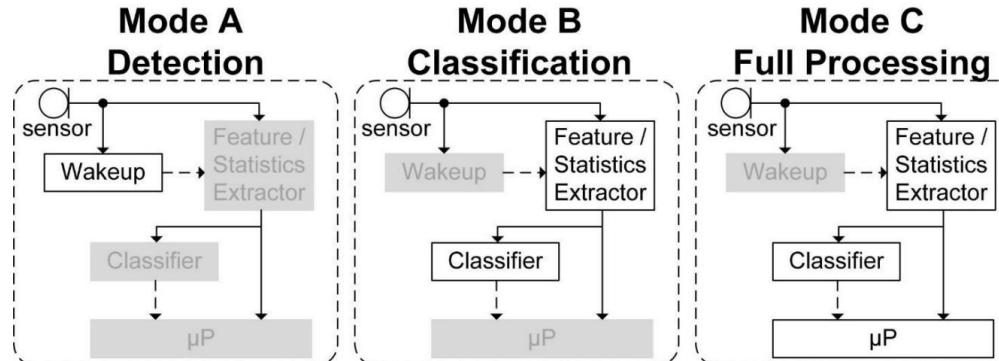
# Wake on feature



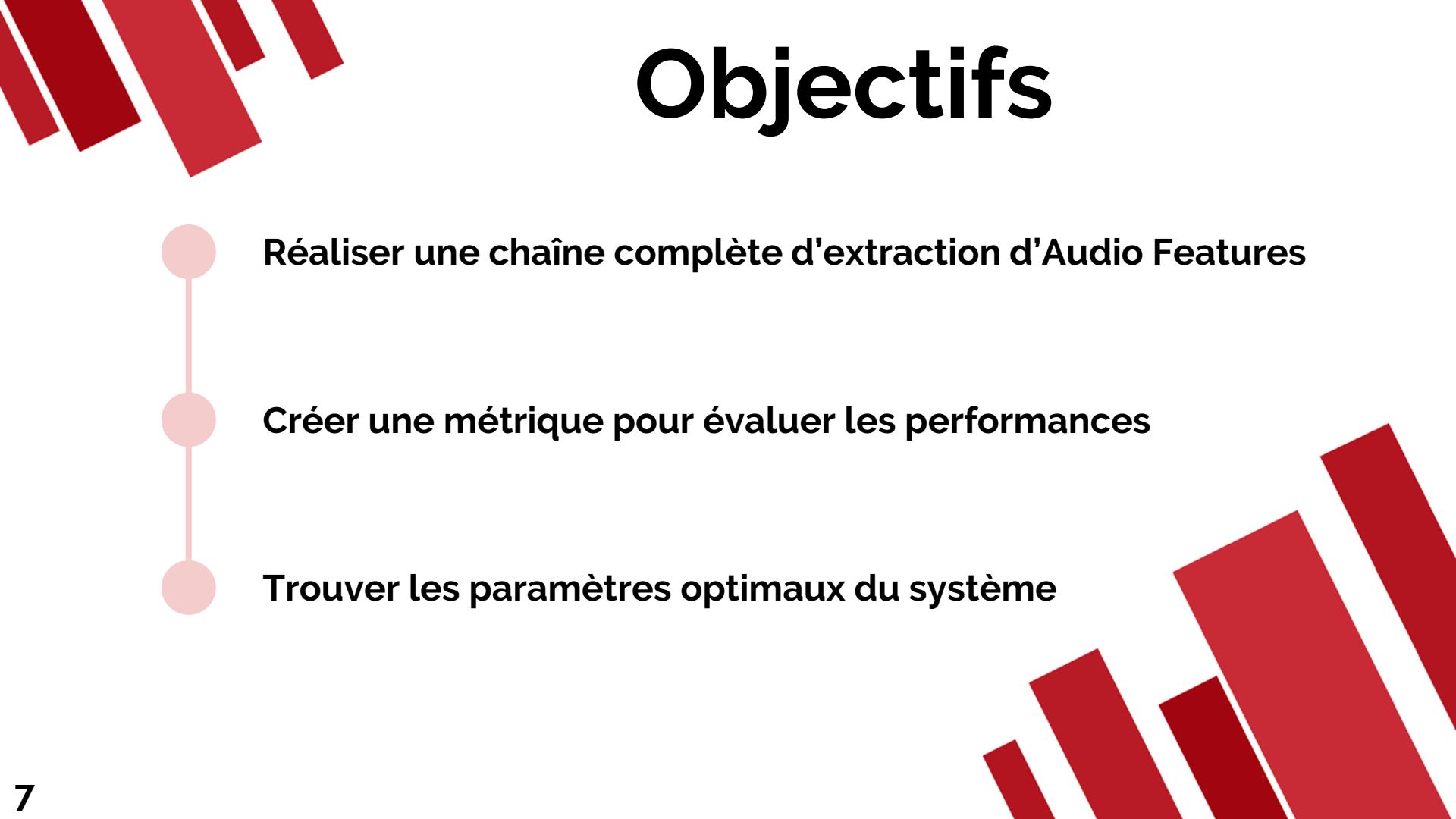
# État de l'art



The enabler of low-power Systems-on-Chip



# Objectifs

- 
- Réaliser une chaîne complète d'extraction d'Audio Features
  - Créer une métrique pour évaluer les performances
  - Trouver les paramètres optimaux du système

# Sommaire



Contexte et état de l'art

Objectifs

Présentation du système

Recherche des paramètres optimaux

Gestion et organisation du projet

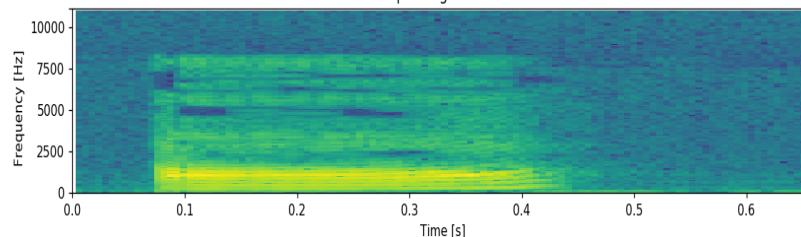
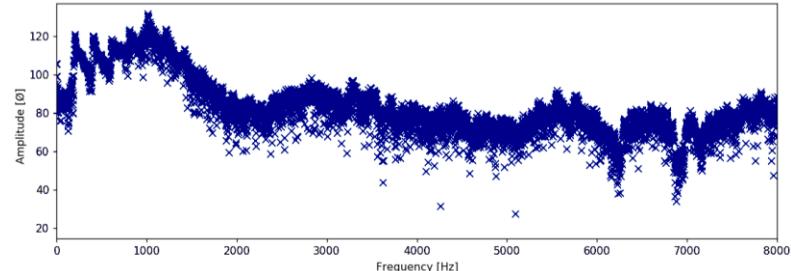
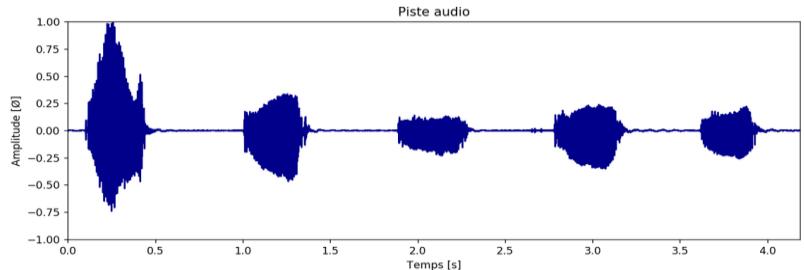
Démonstration

# Présentation du système

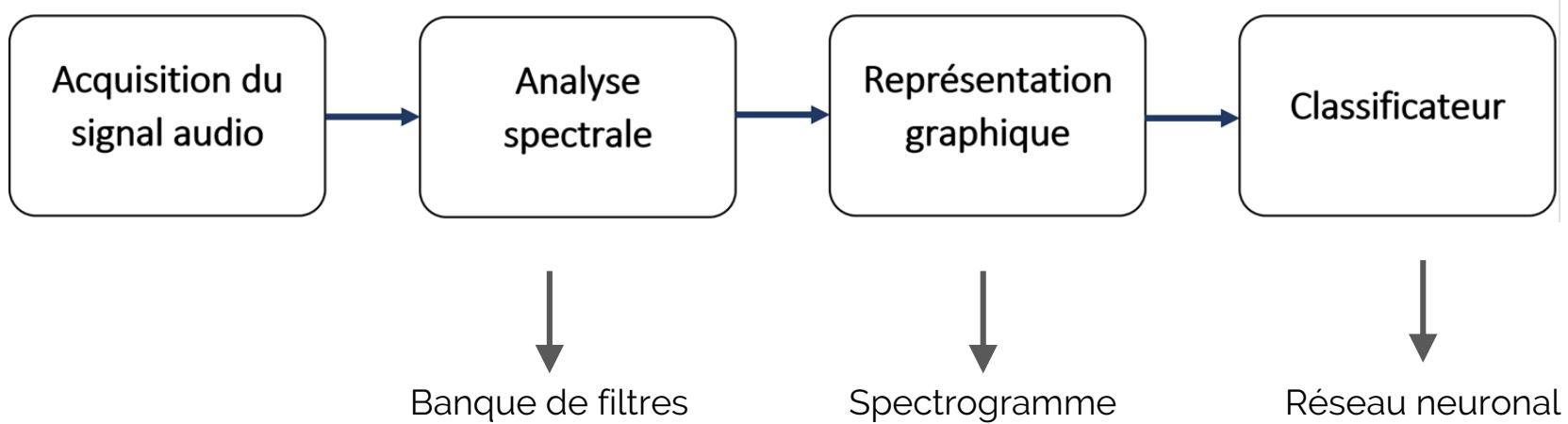
# Les signaux audios

Pour identifier des signaux audio il faut :

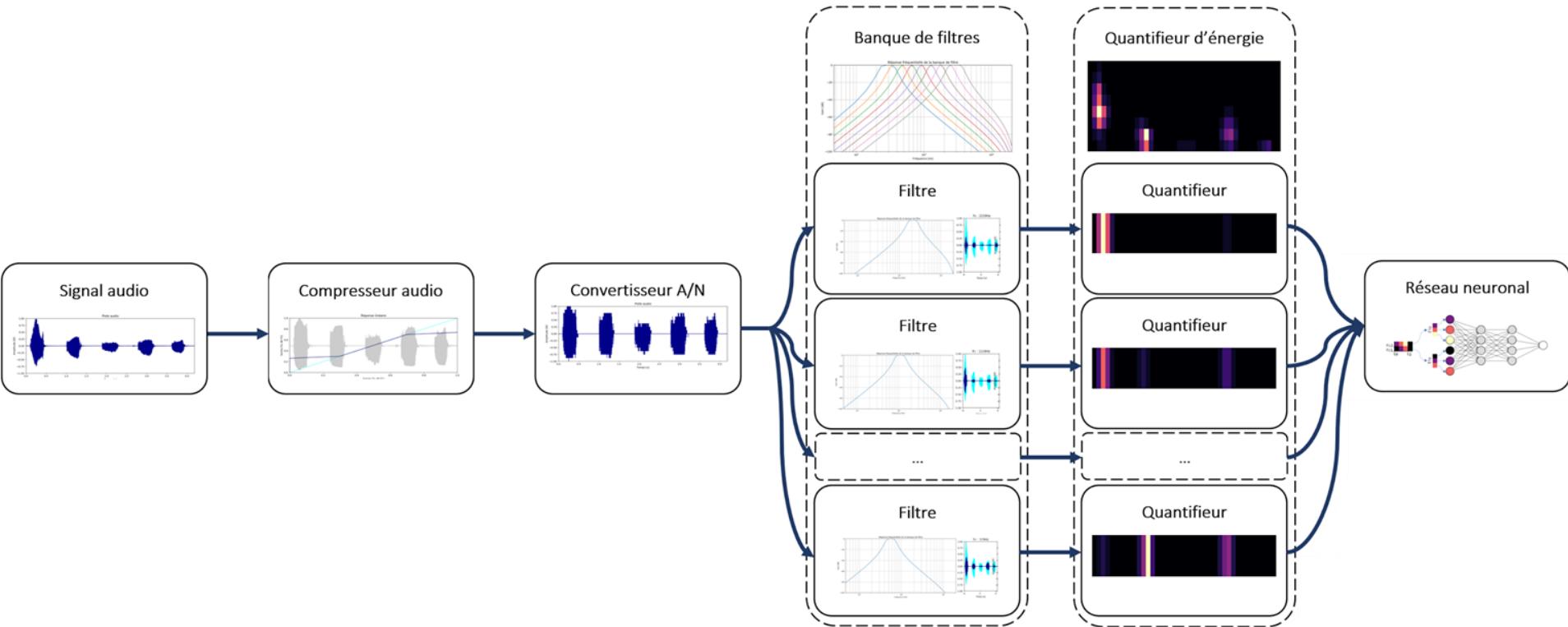
- Informations fréquentielles
- Outil d'analyse graphique
- Quantification de la pertinence des résultats



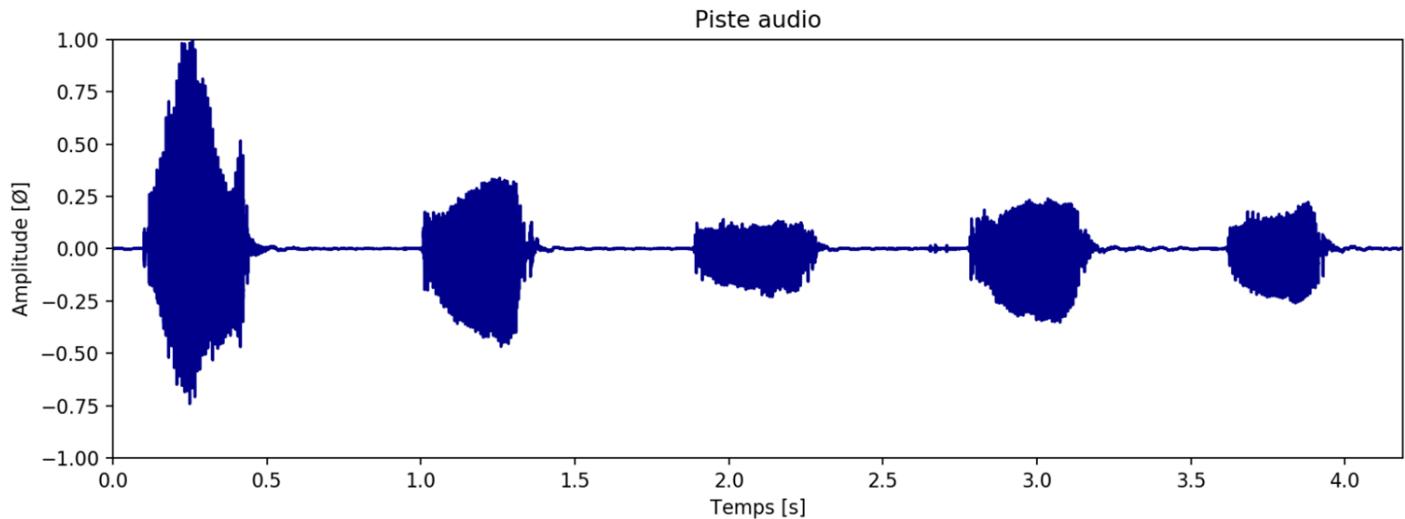
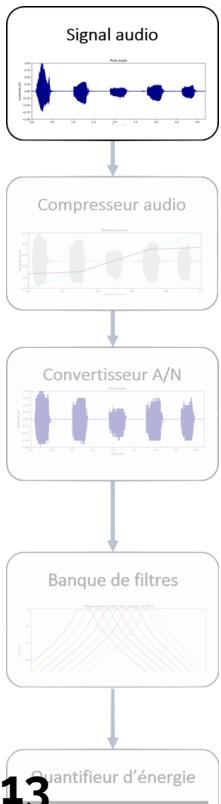
# Nos choix



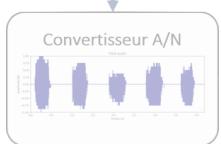
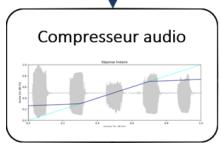
# Aperçu du système



# Acquisition du signal



# Comresseur audio

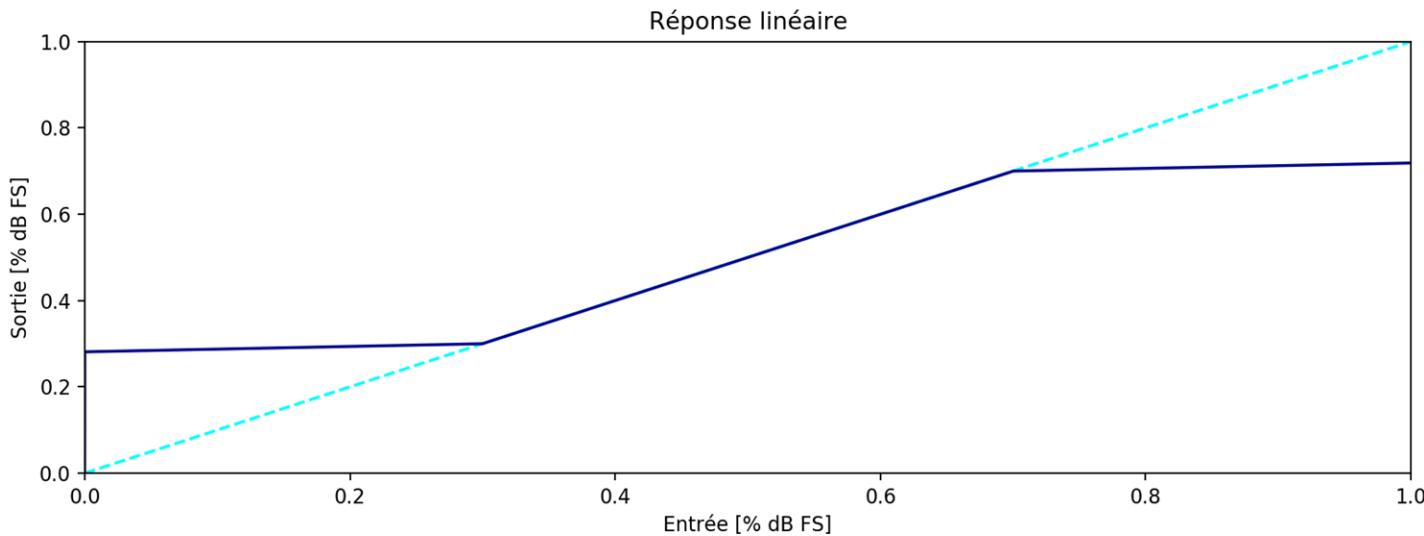


$$y < seuil_{bas}$$

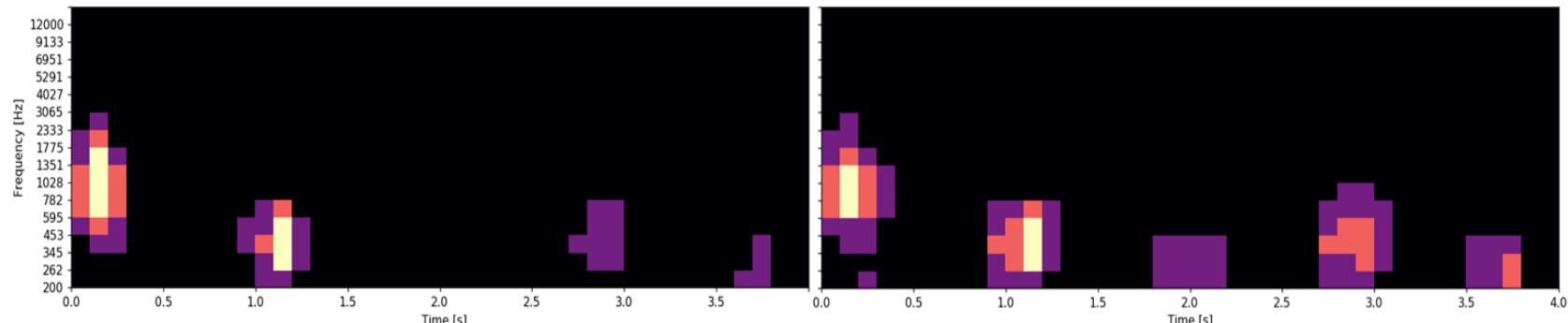
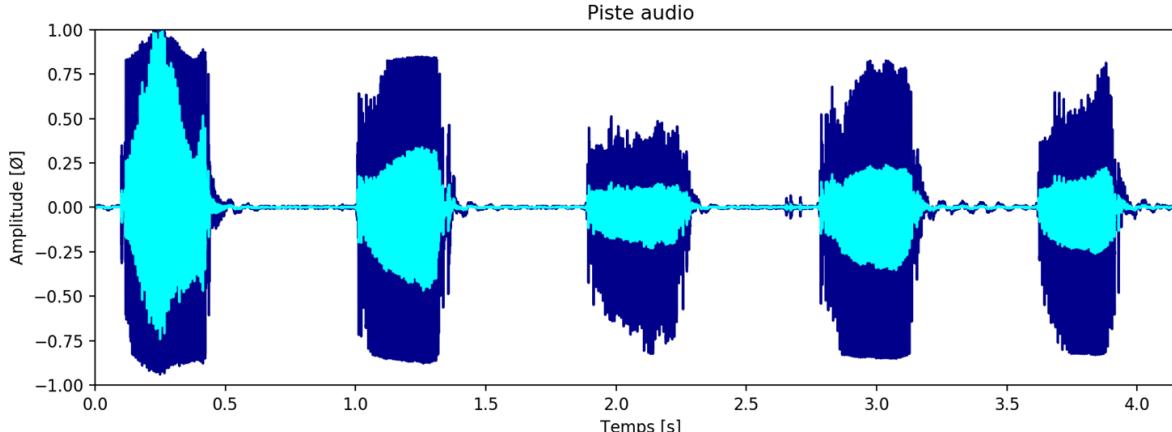
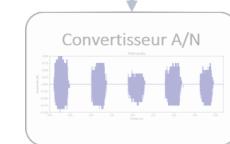
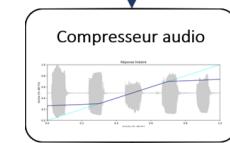
Augmentation du gain de sortie  
Amplifie le signal d'entrée

$$y > seuil_{haut}$$

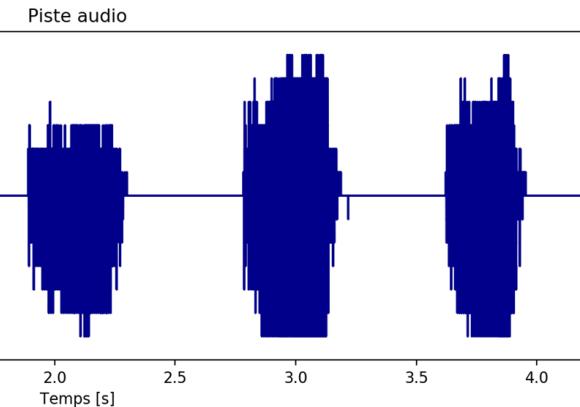
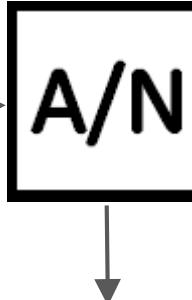
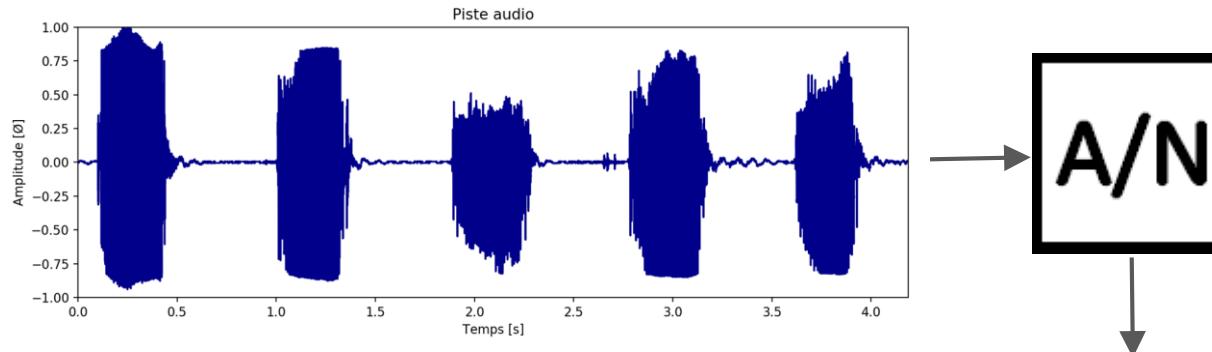
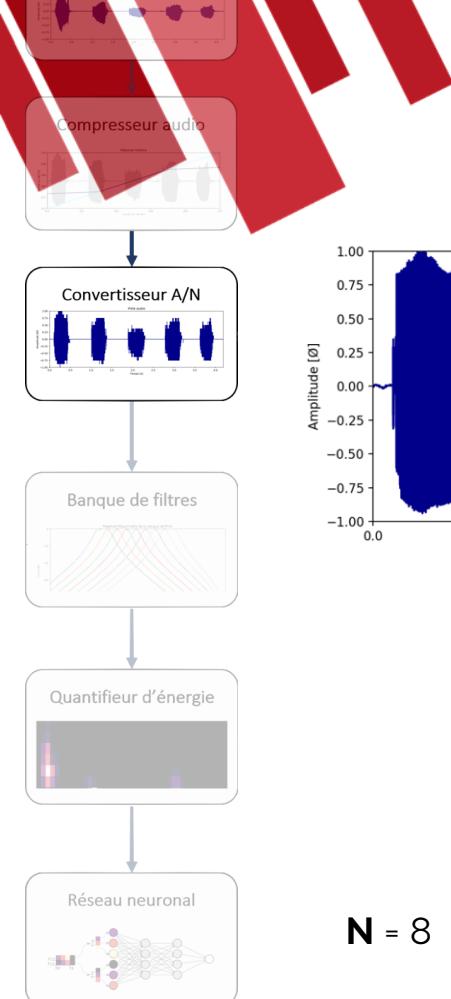
Réduction du gain de sortie  
Atténue le signal d'entrée



# Comresseur audio

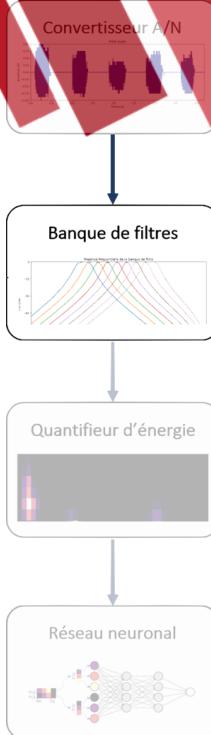


# Convertisseur A/N



$N = 8$

# Banque de filtres



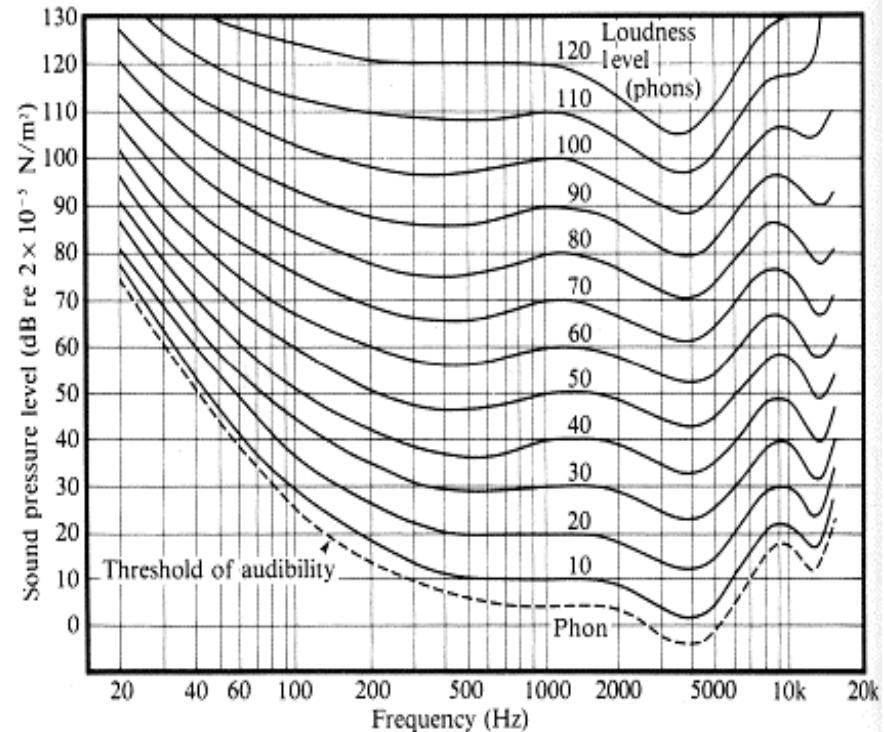
**choix des filtres :**

-passe-bande

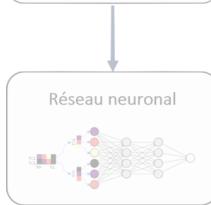
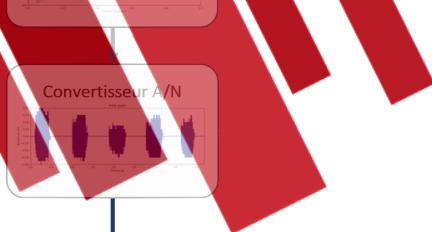
-ordre 3

-espacement entre  
les fréquences centrales ?

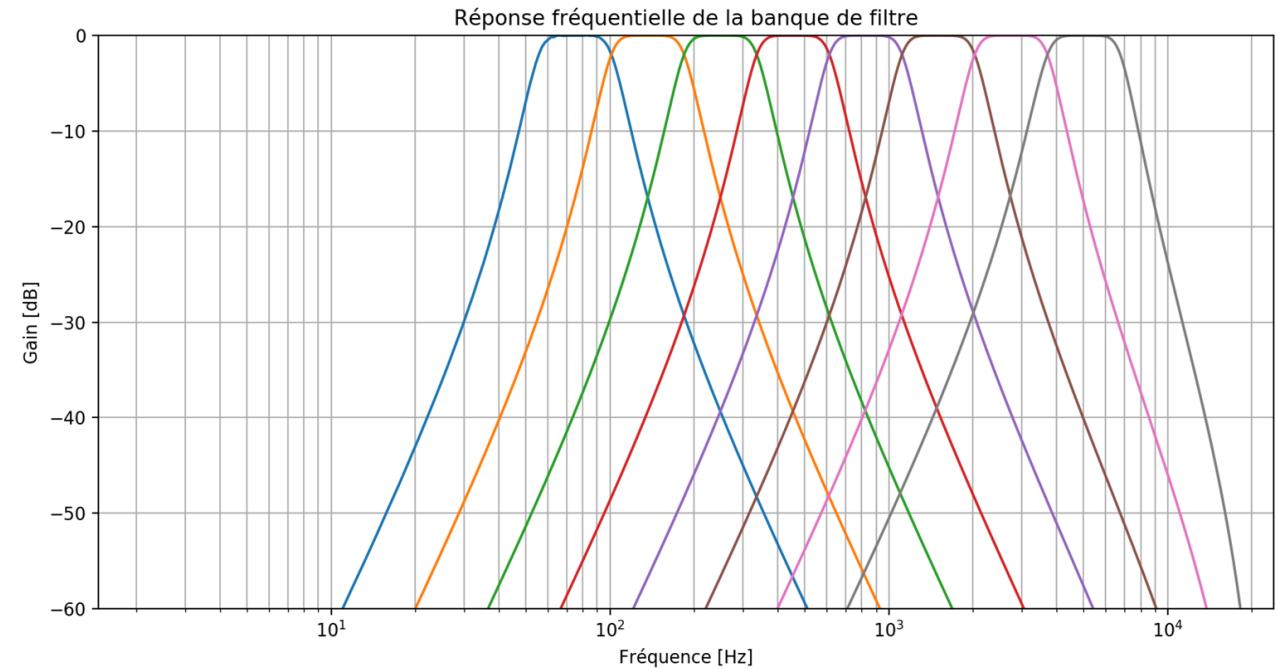
$$f_{coupure} = f_{centrale} \cdot \left( \sqrt{1 + \frac{1}{4Q^2}} \pm \frac{1}{2Q} \right)$$



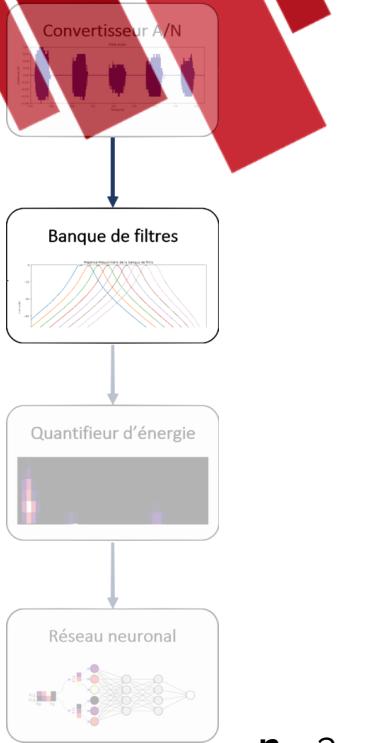
# Banque de filtres



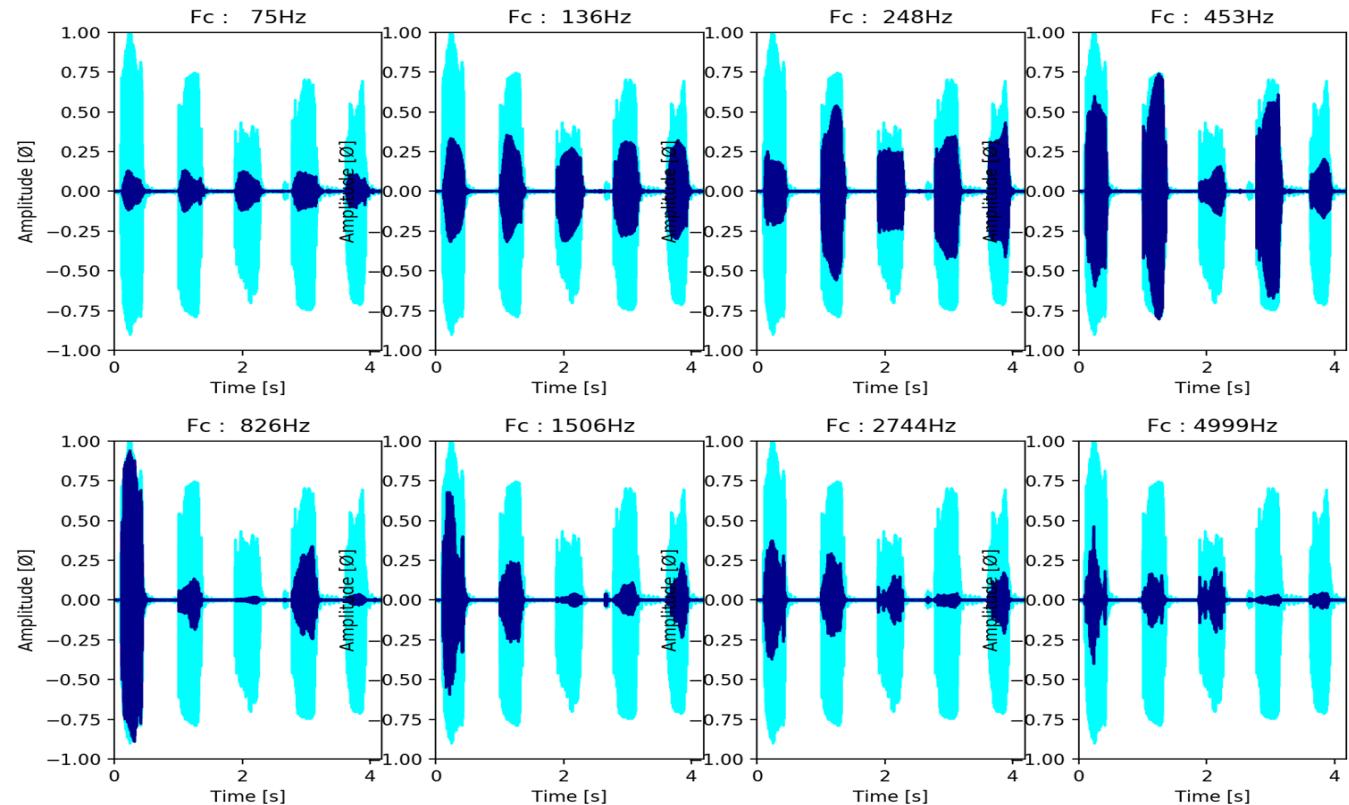
**n** = 3  
**q** = 1.5  
**nb\_filters** = 8  
**fmin** = 75  
**fmax** = 5000  
**scale** = "log"



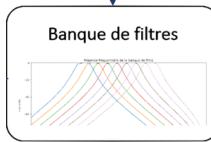
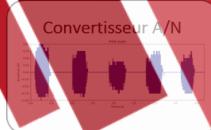
# Banque de filtres



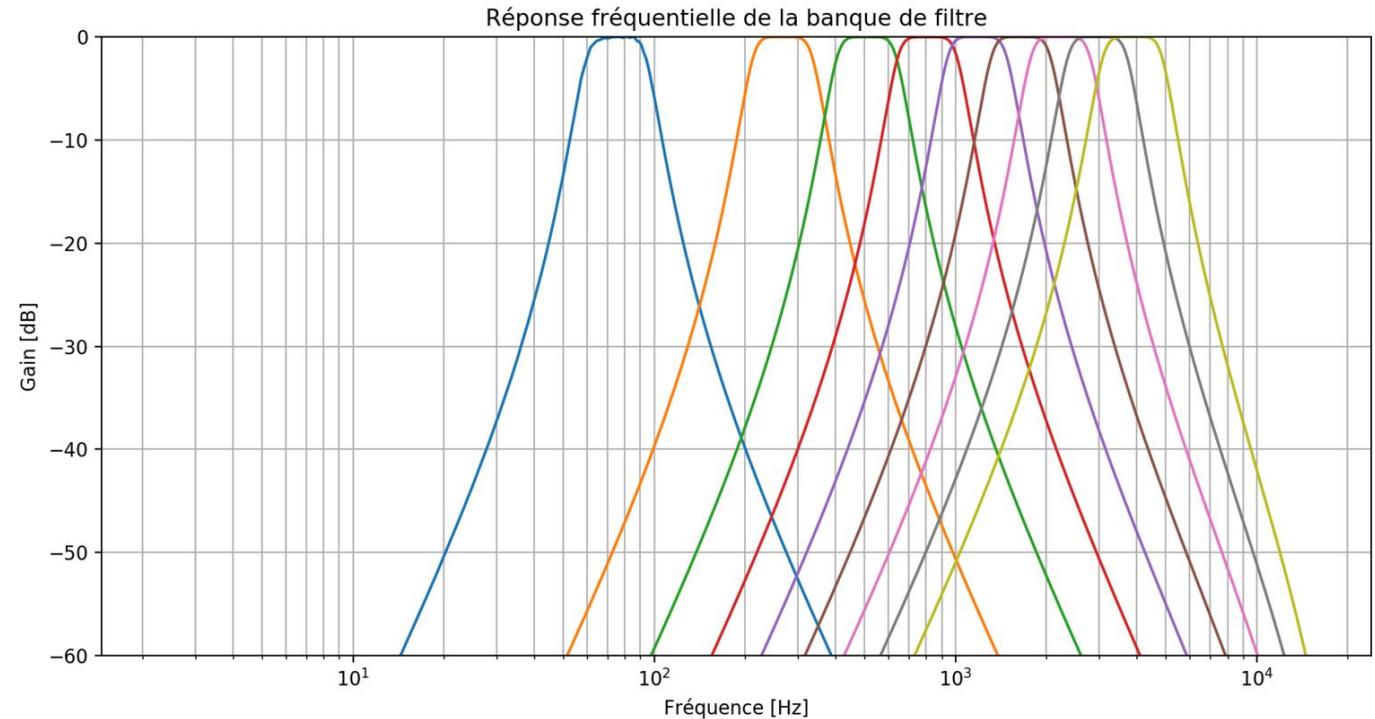
**n** = 3  
**q** = 1.5  
**nb\_filters** = 8  
**fmin** = 75  
**fmax** = 5000  
**scale** = "log"



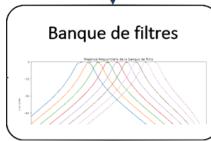
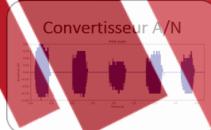
# Banque de filtres



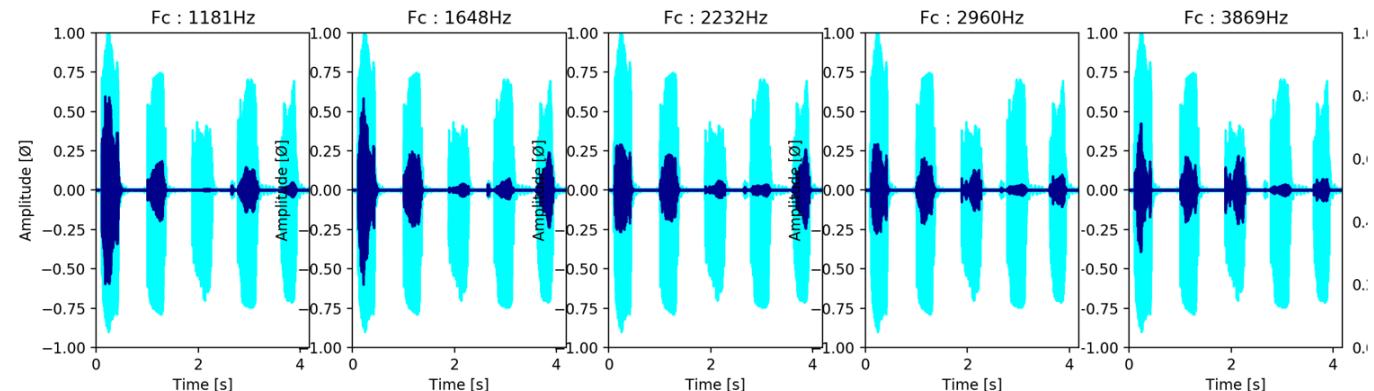
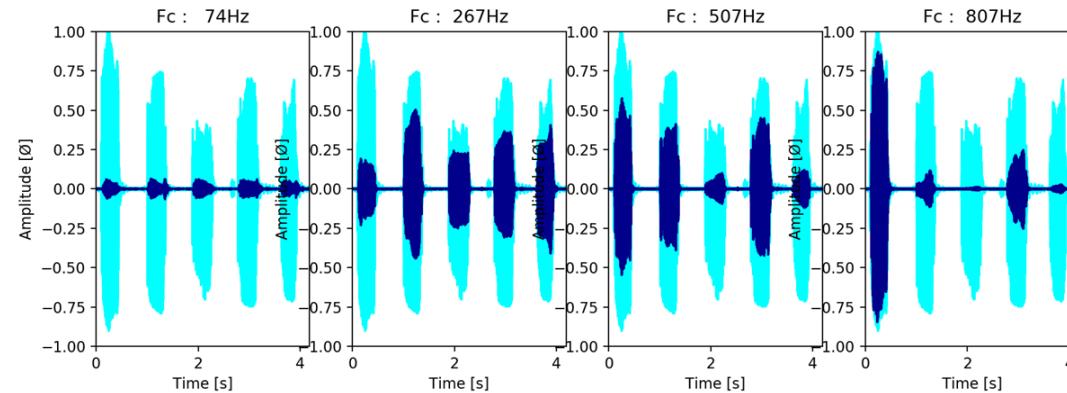
**n** = 3  
**q** = 2  
**nb\_filters** = 8  
**fmin** = 75  
**fmax** = 5000  
**scale** = "mel"



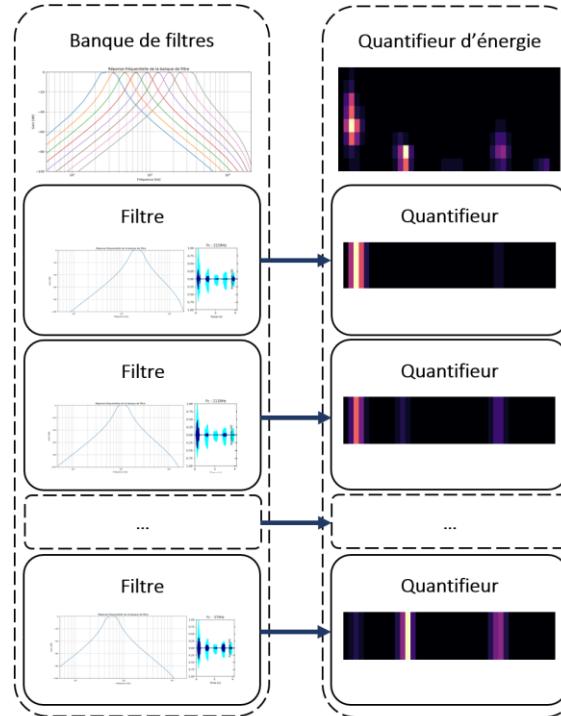
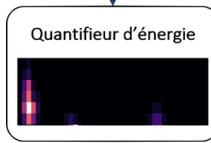
# Banque de filtres



**n** = 3  
**q** = 2  
**nb\_filters** = 8  
**fmin** = 75  
**fmax** = 5000  
**scale** = "mel"



# Quantifieurs d'énergie



## Quantification de l'énergie

$$E = \sum_{t=t_0}^{t_0+dt} |y(t)|^2$$

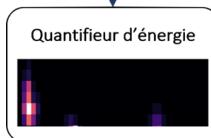
## Spectrogramme sur mesure

Fréquence ciblée

Fenêtre temporelle

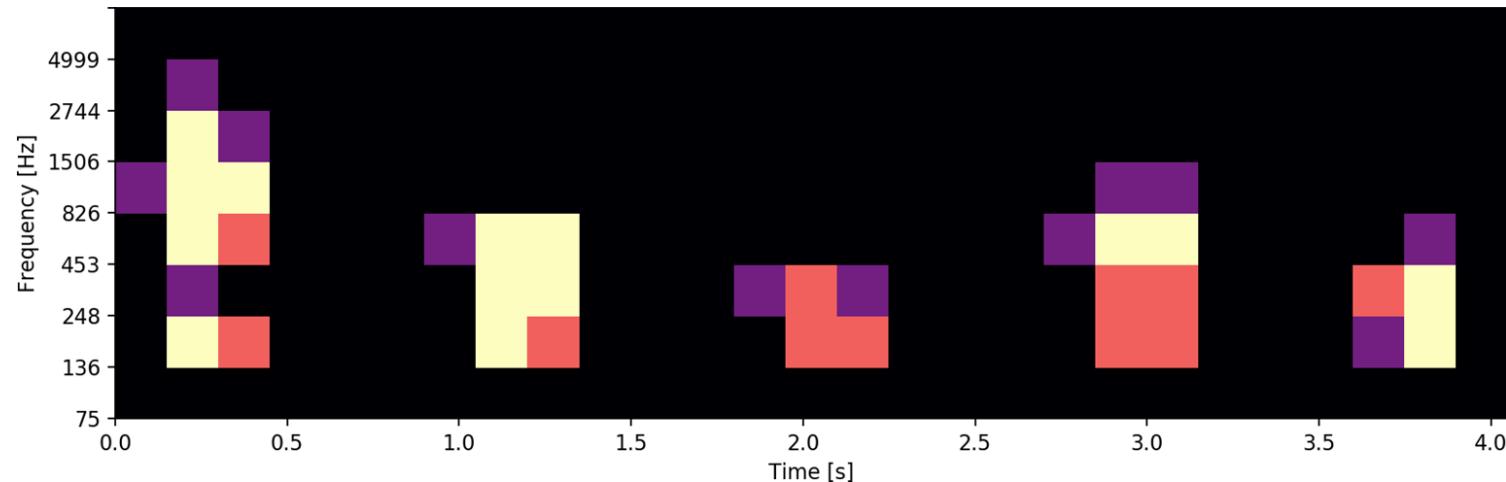
Résolution de l'amplitude

# Quantifieurs d'énergie

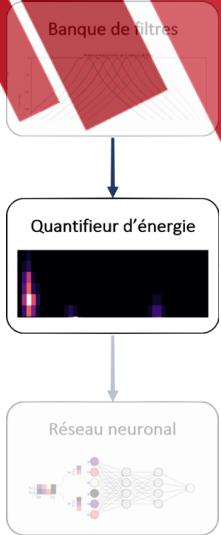


Banque de filtre  
Échelle logarithmique

**time\_res** = 150 ms  
**amp\_res** = 2 bits (4 valeurs)

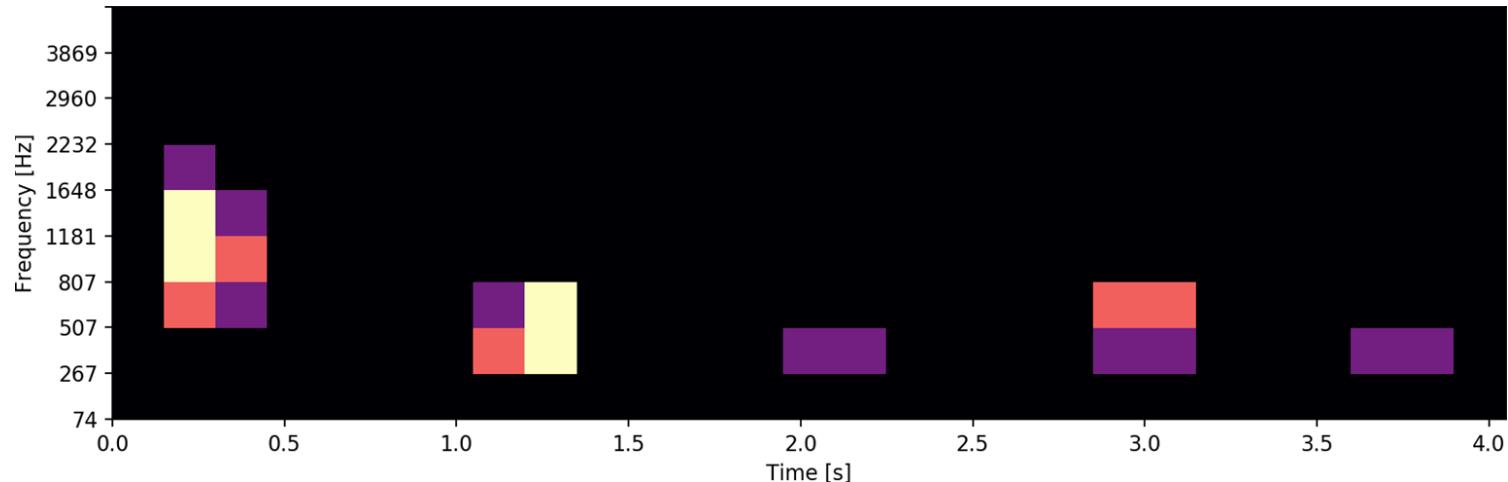


# Quantifieurs d'énergie

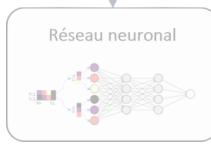
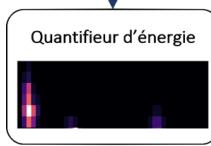


Banque de filtre  
Échelle de Mel

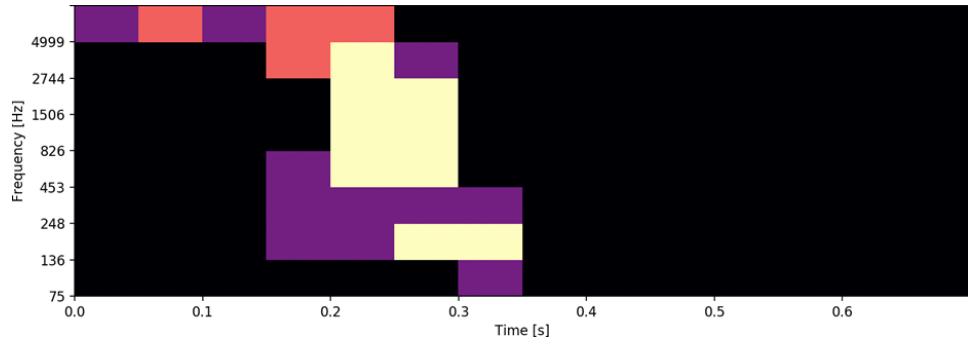
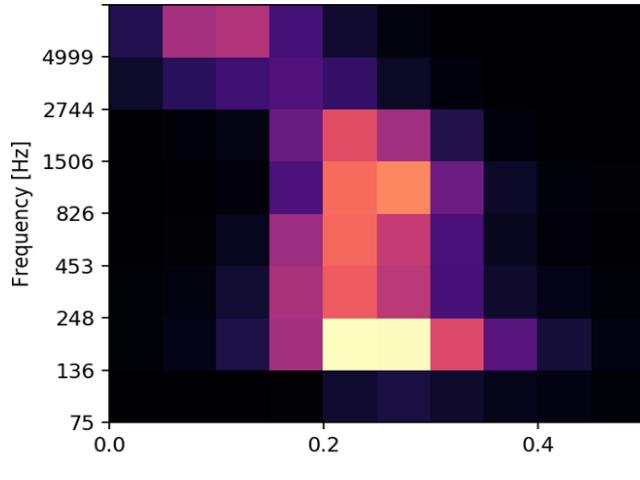
**time\_res** = 150 ms  
**amp\_res** = 2 bits (4 valeurs)

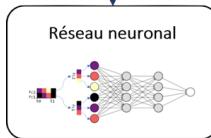


# Quantifieurs d'énergie



Spectrogramme moyen  
du mot “Chien”





# Réseau de neurones

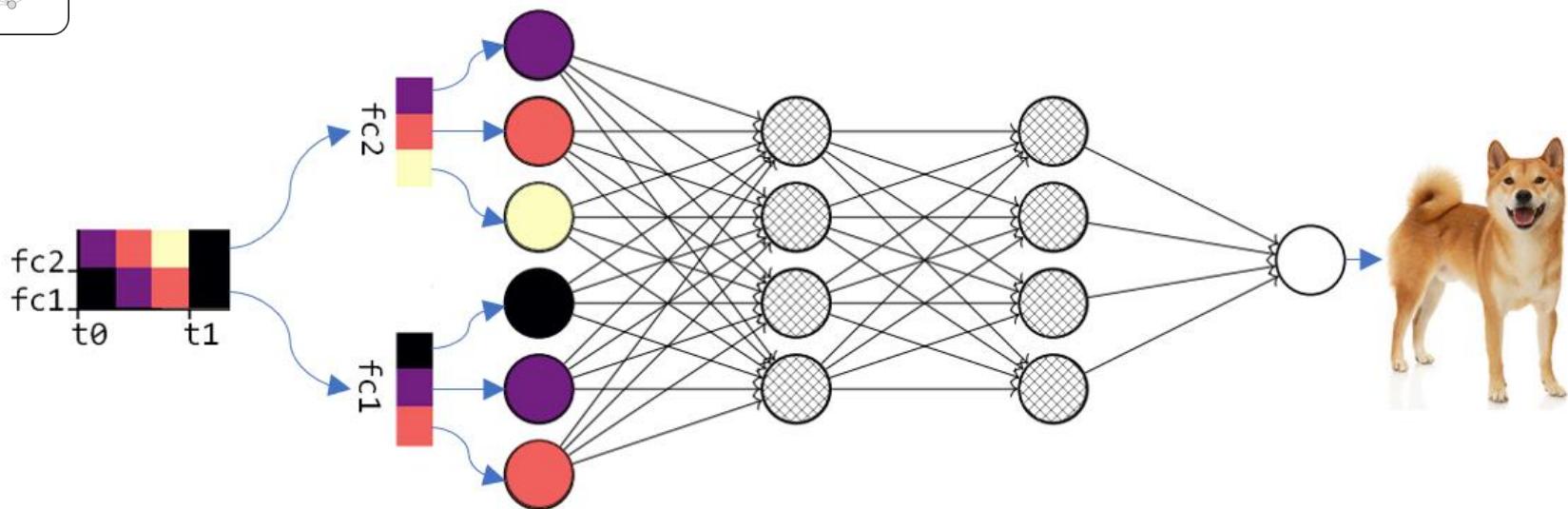
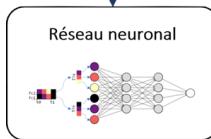
## Évaluation des performances

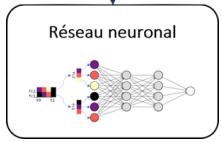
### Classification

Voix contre bruit  
Homme ou femme  
Mot-clé



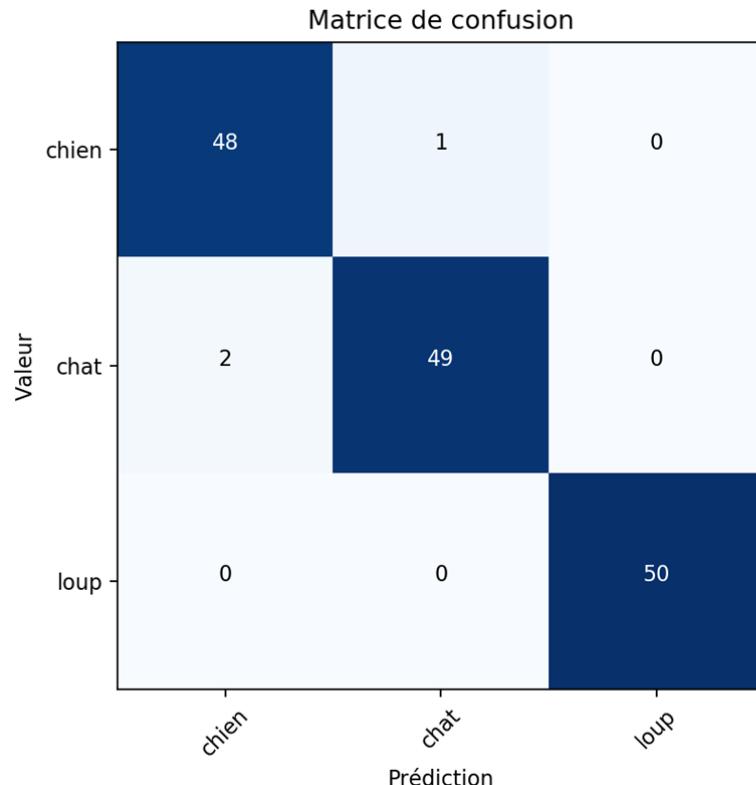
# Réseau de neurones



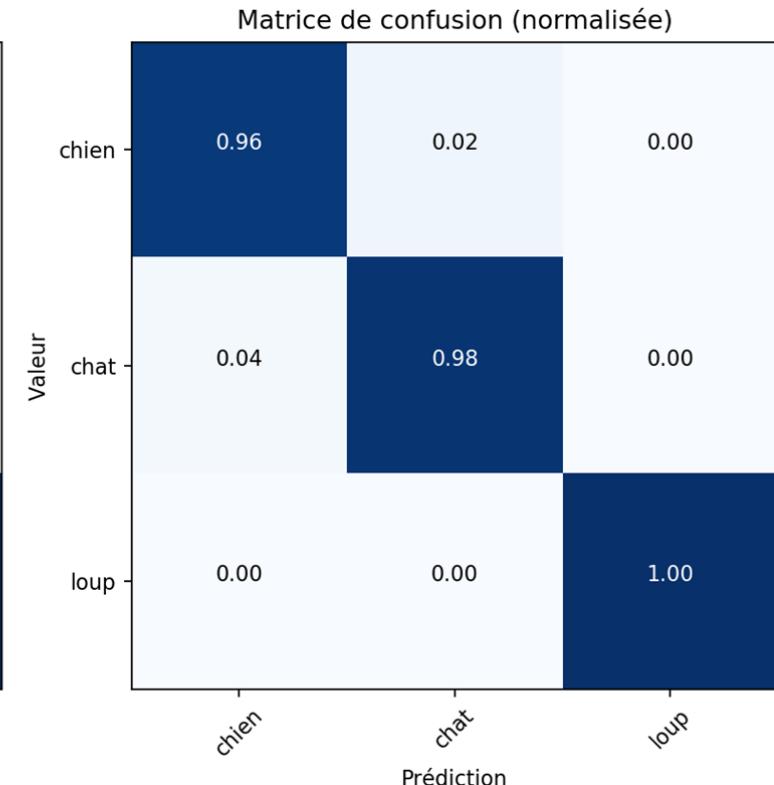


# Réseau de neurones

Matrices de confusion

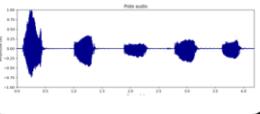


**neurons** = (30)  
**learn** = 70 (x 3)  
**test** = 50 (x 3)

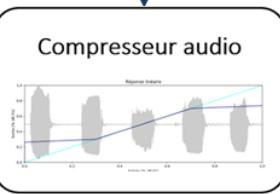


# Recherche des paramètres optimaux

Signal audio

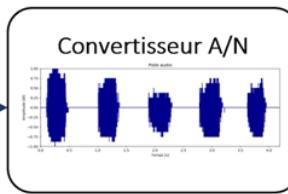


Compresseur audio



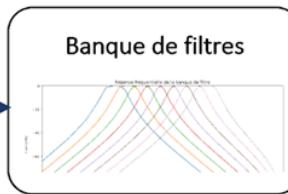
Seuil bas  
Seuil haut  
Ratio

Convertisseur A/N



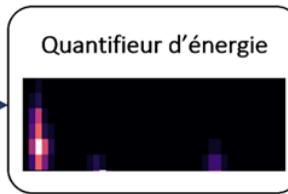
Résolution  
Fréq. échant.

Banque de filtres



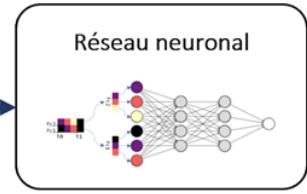
Nb. filtres  
Ordre n  
Facteur Q  
Fréq. min.  
Fréq. max  
Fréq. centrales

Quantifieur d'énergie



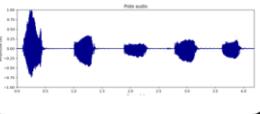
Rés. temp.  
Rés. amp.

Réseau neuronal

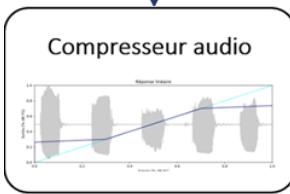


Nb. neurones  
Nb. couches

Signal audio

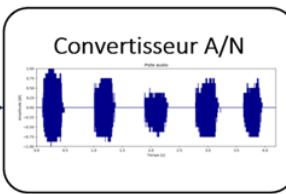


Compresseur audio



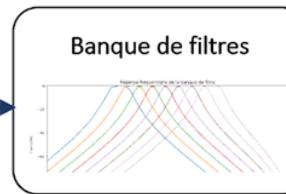
Seuil bas  
Seuil haut  
Ratio

Convertisseur A/N



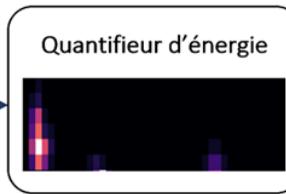
Résolution  
Fréq. échant.

Banque de filtres



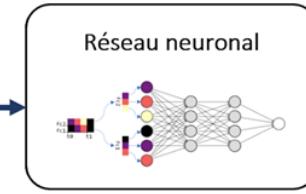
Nb. filtres  
Ordre n  
Facteur Q  
Fréq. min.  
Fréq. max  
Fréq. centrales

Quantifieur d'énergie



Rés. temp.  
Rés. amp.

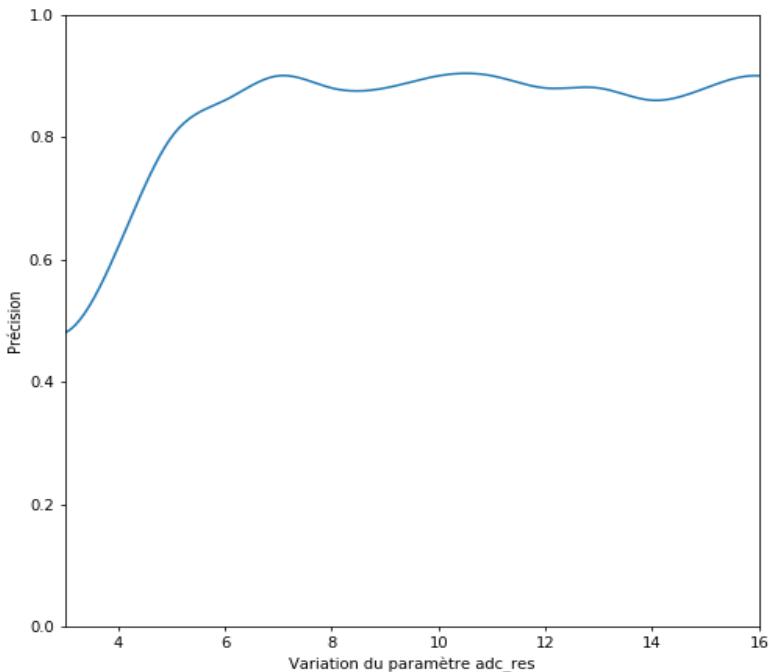
Réseau neuronal



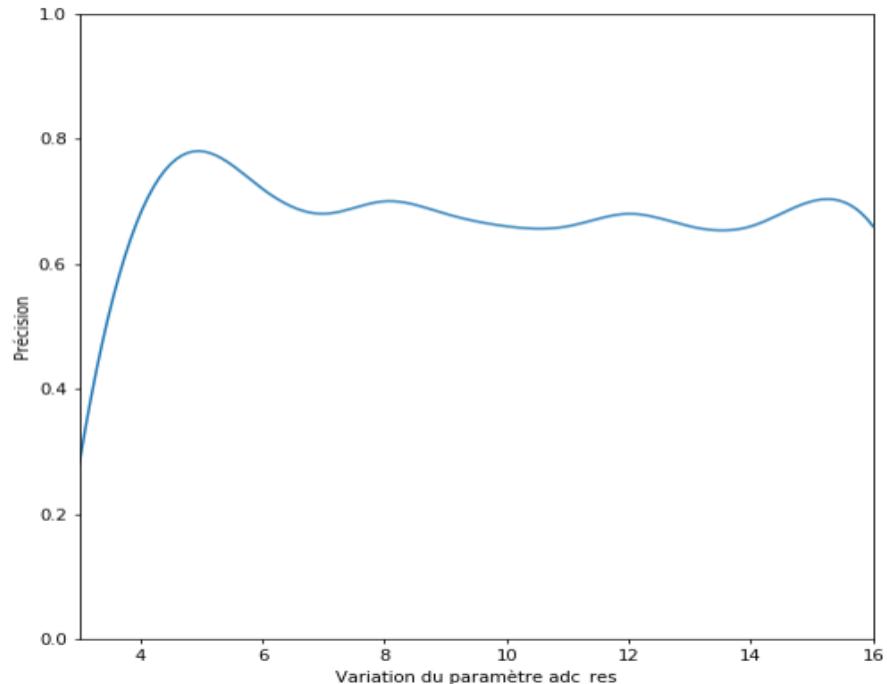
Nb. neurones  
Nb. couches

# Banc de test

Evolution de la précision en fonction de la **résolution de l'ADC**



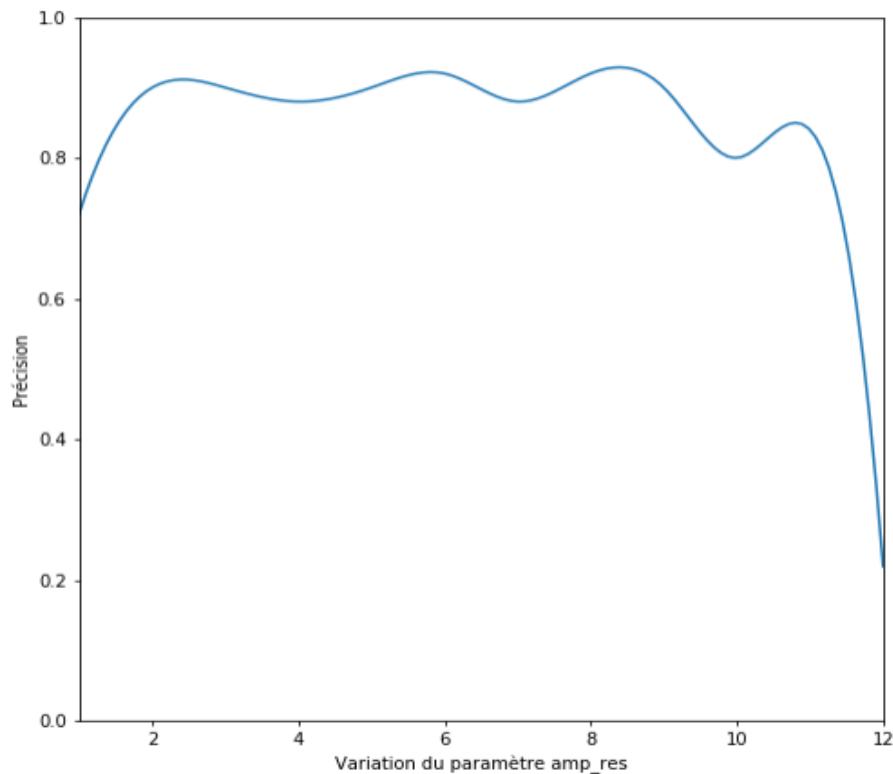
**Q= 1, nb\_filters= 6, amp\_res= 2**



**Q= 3, nb\_filters= 16, amp\_res= 2**

# Banc de test

Evolution de la précision en fonction de la **résolution de l'énergie**



**Q= 1, nb\_filters= 6, adc\_res= 8**

# Gestion et organisation du projet

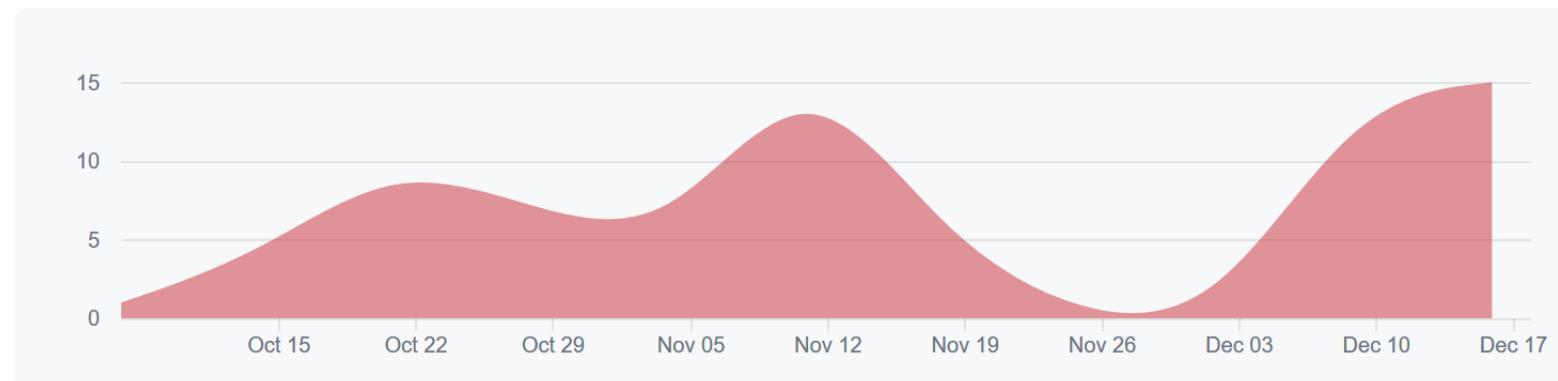
# Travail régulier

(hors semaine des partiels)

Oct 8, 2017 – Dec 19, 2017

Contributions: Commits ▾

Contributions to master, excluding merge commits



# Travail organisé

## Documentation

Sound

Navigation

A

adl  
bandpass  
bookmark  
compute  
cut  
dec  
dec  
energies  
energy  
gen\_data  
gen\_filtered  
gen\_filters  
gen\_sine  
learning\_files  
learning  
live\_record  
player  
plot\_arygram  
plot\_data  
plot\_datagram  
plot\_dbs  
plot\_energies  
plot\_filtered  
plot\_formants  
plot\_freq  
plot\_nccrogram  
plot\_spectrogram

Index

A | B | C | D | E | F | G | H | I | J | K | L | M | N | O | P | Q | R | S | T

adl (fichier du module bin.adl)

B

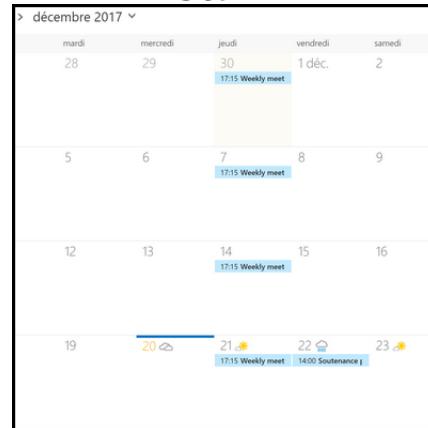
bandpass0 (fichier du module bin.bandpass)  
bin  
bin\_bandpass0 (fichier du module bin.bandpass)  
bin\_benchmark3  
bin\_adl (fichier)  
bin\_bandpass (fichier) [1]  
bin\_benchmark (fichier)  
bin\_compute (fichier)  
bin\_conpute (fichier)  
bin\_cut (fichier)  
bin\_drc (fichier)  
bin\_drc (fichier)  
bin\_formants (fichier)  
bin\_envelope (fichier)  
bin\_envelope (fichier)  
bin\_gen\_data (fichier)  
bin\_gen\_filtered (fichier)  
bin\_gen\_filters (fichier)  
bin\_gen\_sine (fichier)  
bin\_learning (fichier)

C

## Rapport interactif



## Réunions et suivi



# Objectifs



## Réaliser une chaîne complète d'extraction d'Audio Features

Chaîne composée de six blocs

Grande modularité au niveau des paramètres

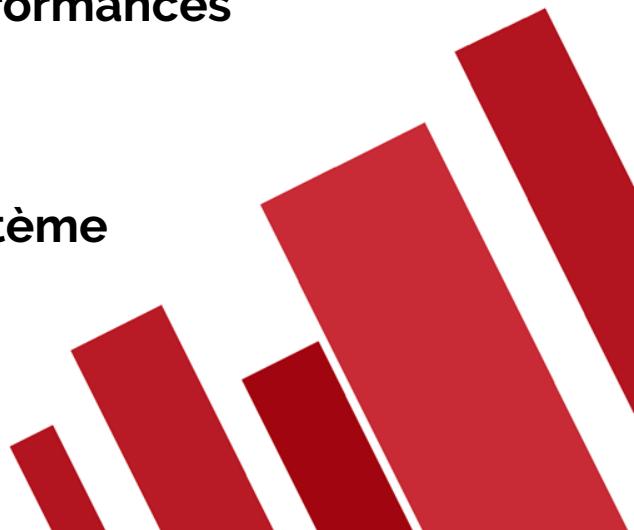
## Créer une métrique pour évaluer les performances

Utilisation d'un réseau de neurones artificiels

Précision et matrices de confusion

## Trouver les paramètres optimaux du système

Réalisation d'un banc de test





OK  
ISEN