Classification acoustique des drones :

Approche combinant analyse temps-fréquence et Deep Learning

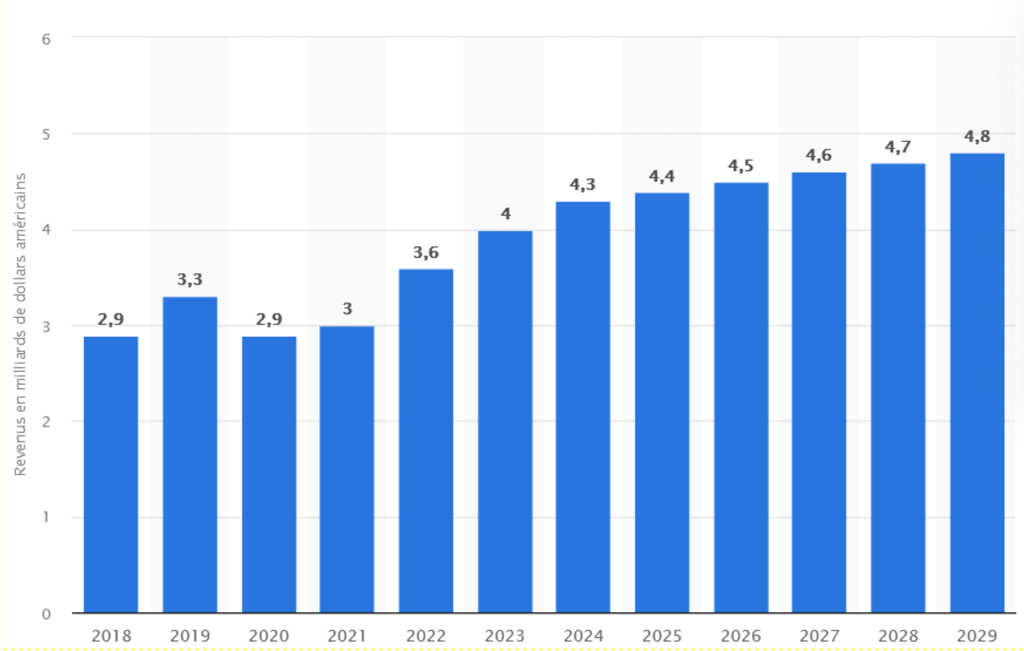
PROJET SYSTÈME –SOIA 2025

RAPHAËL GRISEL

Contexte

En France :

- 103 000 drones vendus en 2020
- Parc de drones de loisirs en France estimé à 2,5 millions d'appareils



Marché mondial du drone grand public [1]

[1] Maxime Gautier, 14 janv. 2025 , https://fr.statista.com/statistiques/607544/croissance-prevue-marche-drones-professionnels-monde/

Contexte

Attaque sur le président Venezuelien Maduro en 2018 [2]

Crash de drone imprévisible

[2] T. Guardian, “P. daniels, venezuela’s nicolas maduro survives apparent assassination attempt,” https://www.theguardian.com/world/2018/ aug / 04/ nicolas - maduros - speech - cut - short - while -soldiers-scatter., 2018.

Sommaire

- Etat de l'art:
  - Méthodes de classification
  - Méthodes TFT et Deep Learning

- Projet
  - Le Dataset
  - Le Modèle
  - Détection de drones
  - Classification
  - Représentation en spectrogramme de Mel

- Retour sur le projet

Etat de l'art : méthodes de machine learning

Détection et classification de drones

Système d'observation :

- Acoustique
- Radar
- Vision
- Radiofréquence

Modèle de machine Leaning :

- Convolutional Neural Network (CNN)
- Recurrent Neural Network (RNN)
- Recurent Convolutional Neural Network (RCNN)
- Support Vector Machine (SVM)

Etat de l'art : méthodes de machine learning

Méthode	Avantages	Inconvénients
Acoustique	Etude de la signature sonore Méthode peu intrusive	Portée de détection limitée Sensible au bruit ambiant
Radar	Longue portée Peu sensible aux conditions météorologiques	Pas adapté aux drones de faible SER
Vision	Caméras peu coûteuses Fournit des informations visuelles	Dépend de la ligne de mire Sensible aux conditions météorologiques
Radiofréquence	Localiser les contrôleurs du drone. Longue portée	Inutile si le drone est en mode autonome Sensible au brouillage et aux interférences

Acquisition de données pour la détection et la classification de drones [3]

[3] B. TAHA and A. SHOUFAN, "Machine learning-based drone detection and classification: State-of-the-art in research," IEE, 2019.

Etat de l'art:

Méthode acoustique :

- Peu coûteux
- Peu technique

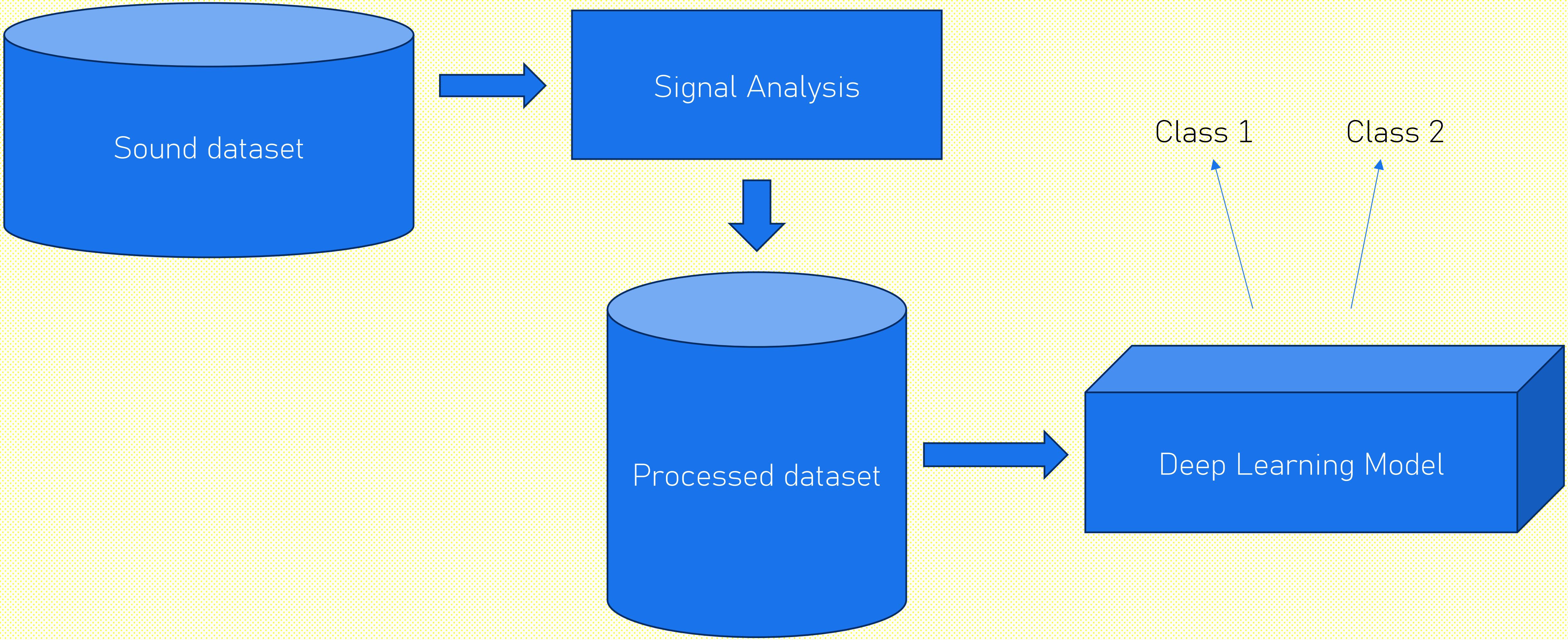
MAIS

- Manque de données d'entraînement
- Prise en compte du bruit ambiant

[4] A. M. A. A.-A. Sara Al-Emadi, Abdulla Al-Ali, "Audio based drone detection and identification using deep learning," 2019 15th International Wireless Communications Mobile Computing Conference (IWCMC),2019

[5] Sungho Jeon1, Jong-Woo Shin, Young-Jun Lee, Woong-Hee Kim, YoungHyoun Kwon, and Hae-Yong Yang "Empirical Study of Drone Sound Detection in Real-Life Environment with Deep Neural Networks", 2017arXiv

Etat de l'art : chaîne de traitement



Etat de l’art : analyse temps fréquence

Représentation conjointe du signal dans les domaines temporel et fréquentiel

- Analyse des signaux non stationnaires
- Extraction des caractéristiques pertinentes
- Classification basée sur l’imagerie

Etat de l’art :

Audio Based Drone Detection and Identification using Deep Learning

Sara Al-Emadi, Abdulla Al-Ali, Amr Mohammad, Abdulaziz Al-Ali
Department of Computer Science and Engineering
Qatar University
Doha, Qatar
saraalemadi@ieee.org, {abdulla.alali, amrm, a.alali}@qu.edu.qa

Mis à disposition d’un dataset d’enregistrement de drones

Détection et classification sur des spectrogrammes d’enregistrement sonores de drones

Accuracy des modèles pour les tâches de détection et classification [4]

	Détection	Classification
CNN	96,38 %	92,94%
RNN	75,00%	57,16%
CRNN	94,72%	92,22%

CNN est le modèle le plus performant pour la détection et la classification

4] A. M. A. A.-A. Sara Al-Emadi, Abdulla Al-Ali, “Audio based drone detection and identification using deep learning,” 2019 15th International Wireless Communications Mobile Computing Conference (IWCMC),2019

Objectifs du projet

- Implémenter un algorithme de détection et de classification acoustiques sur le dataset proposé par [4]

- Etudier l'influence de différentes Transformations Temps-Fréquence sur les performances du modèle de Deep Learning

4] A. M. A. A.-A. Sara Al-Emadi, Abdulla Al-Ali, "Audio based drone detection and identification using deep learning," 2019 15th International Wireless Communications Mobile Computing Conference (IWCMC),2019

Le Dataset

Bebop [6]

Mambo [6]

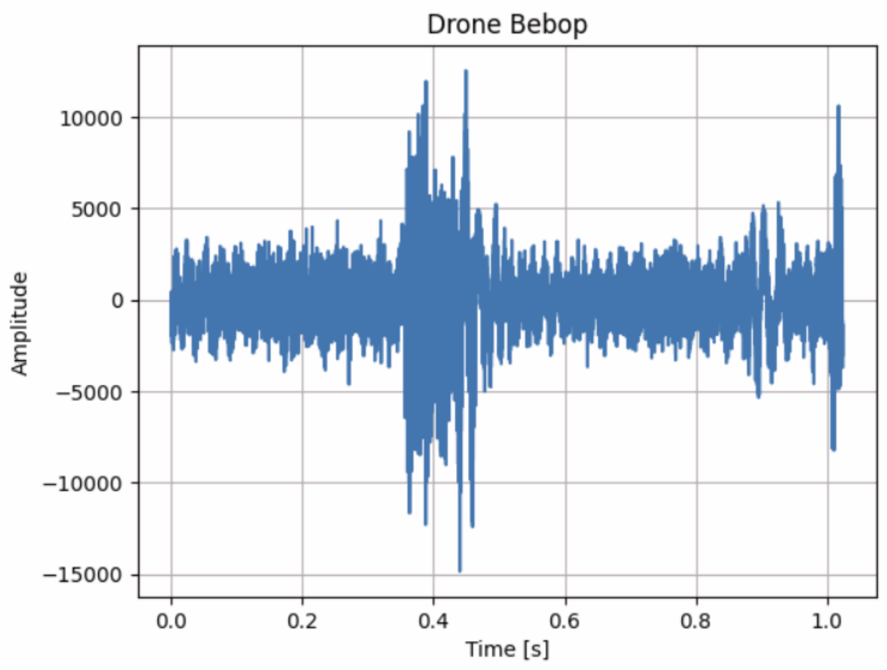
Environnement
sonore

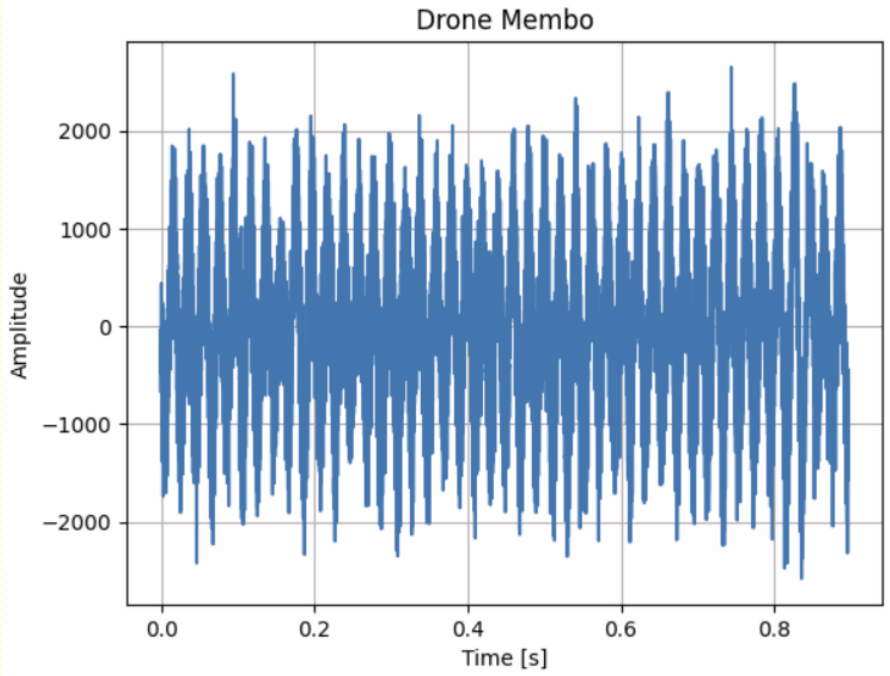
[6] parrot.com

Le Dataset

- Acquisition :
  - microphone d'un smartphone
  - environnement intérieur calme
  - 11 minutes et 6 seconds par drone
- Augmentation :
  - Simuler enregistrements réalistes
  - Sons drones + sons environnements
- Mise en forme :
  - Etiquetage
  - Découpage en segment de 1 seconde
  - Dataset Détection : 2 classes
  - Dataset Classification : 3 classes

Analyse des enregistrements



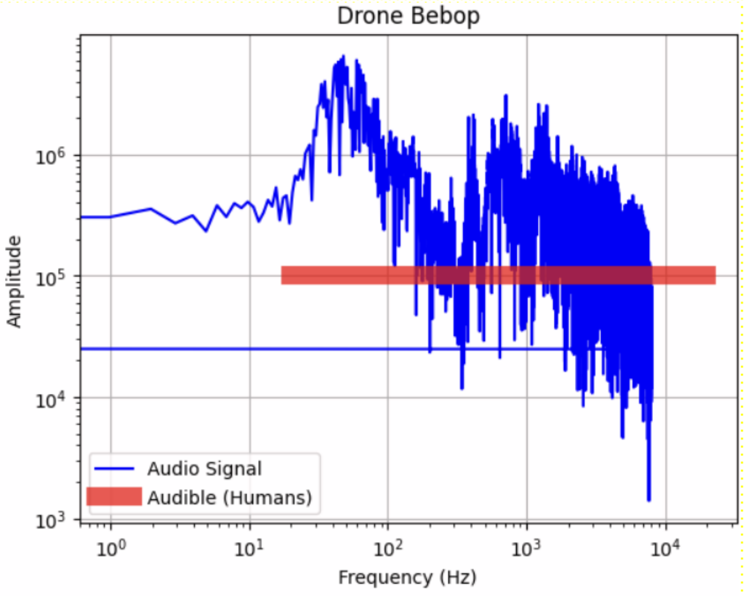


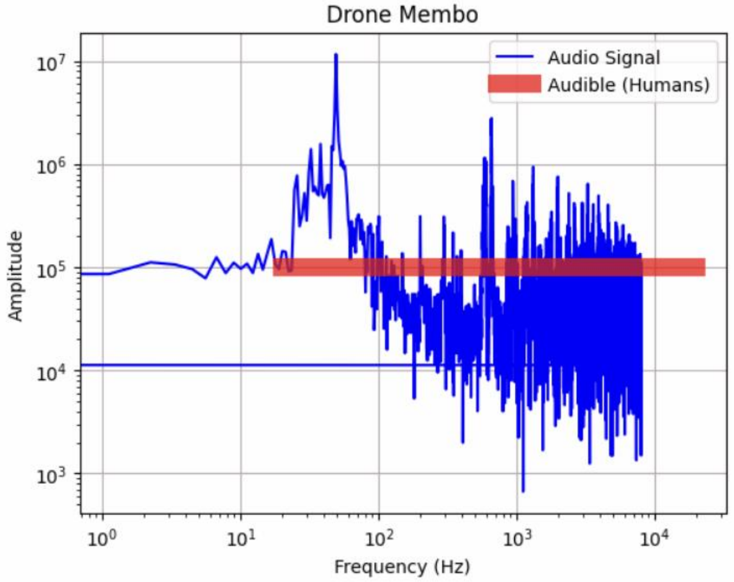
Visualisations temporelles des enregistrements sonores

Analyse des enregistrements

Fast Fourier Transform :

$$X[k] = \sum_{n=0}^{N-1} x[n] e^{-j\frac{2\pi kn}{N}}, \quad k = 0, 1, 2, \ldots, N-1$$





Analyse des enregistrements

- Spectrogrammes :
  - Transformée de Fourier à court terme:

$$X(m,k) = \sum_{n=0}^{N-1} x[n]w[n-mR]e^{-j\frac{2\pi kn}{N}}$$

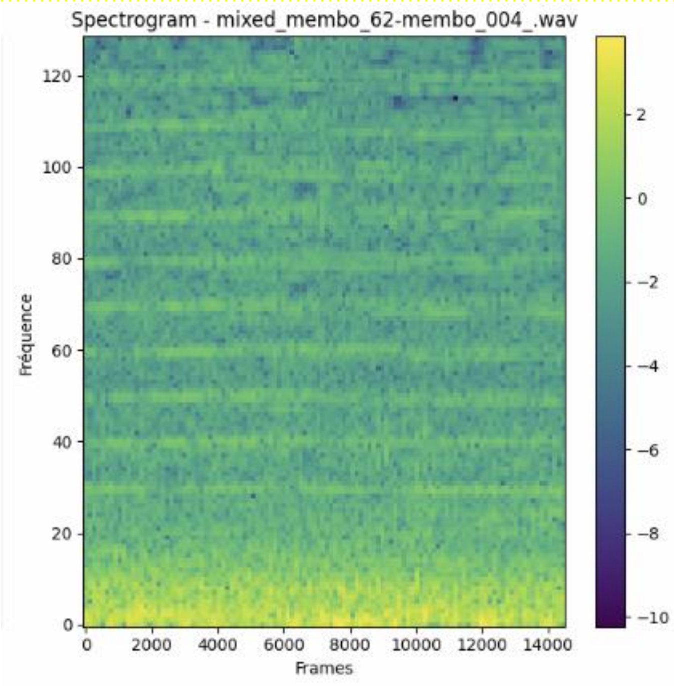
  - Spectre en amplitude :

$$S(m,k) = |X(m,k)|^2$$

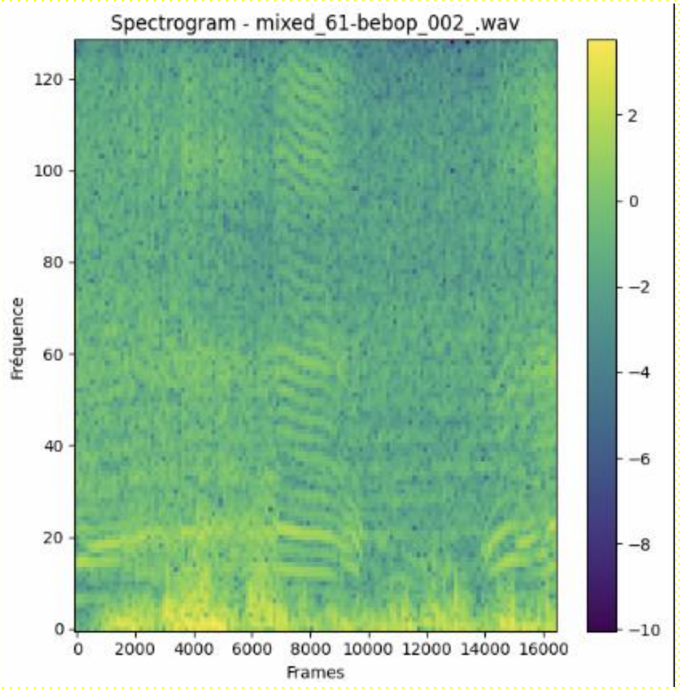
  - Conversion en dB :

$$S_{\mathrm{dB}}(m,k) = 10\log_{10} S(m,k)$$

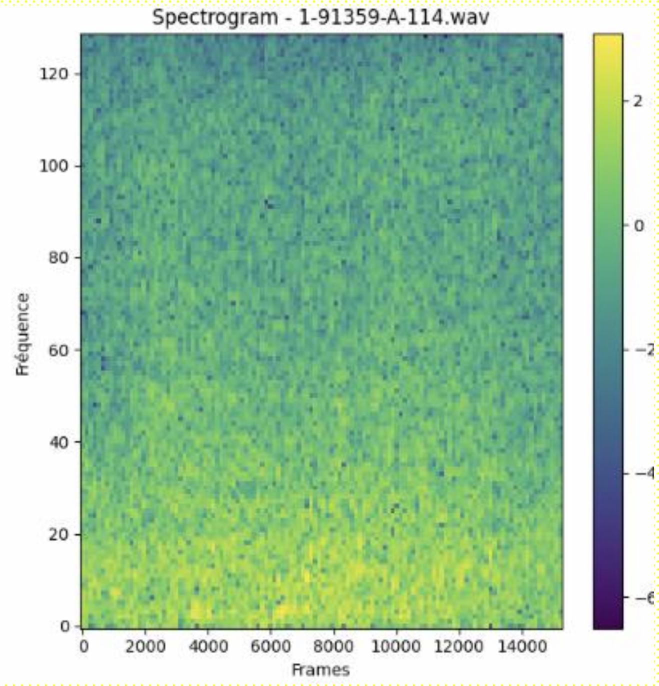
Analyse des enregistrements



Mambo

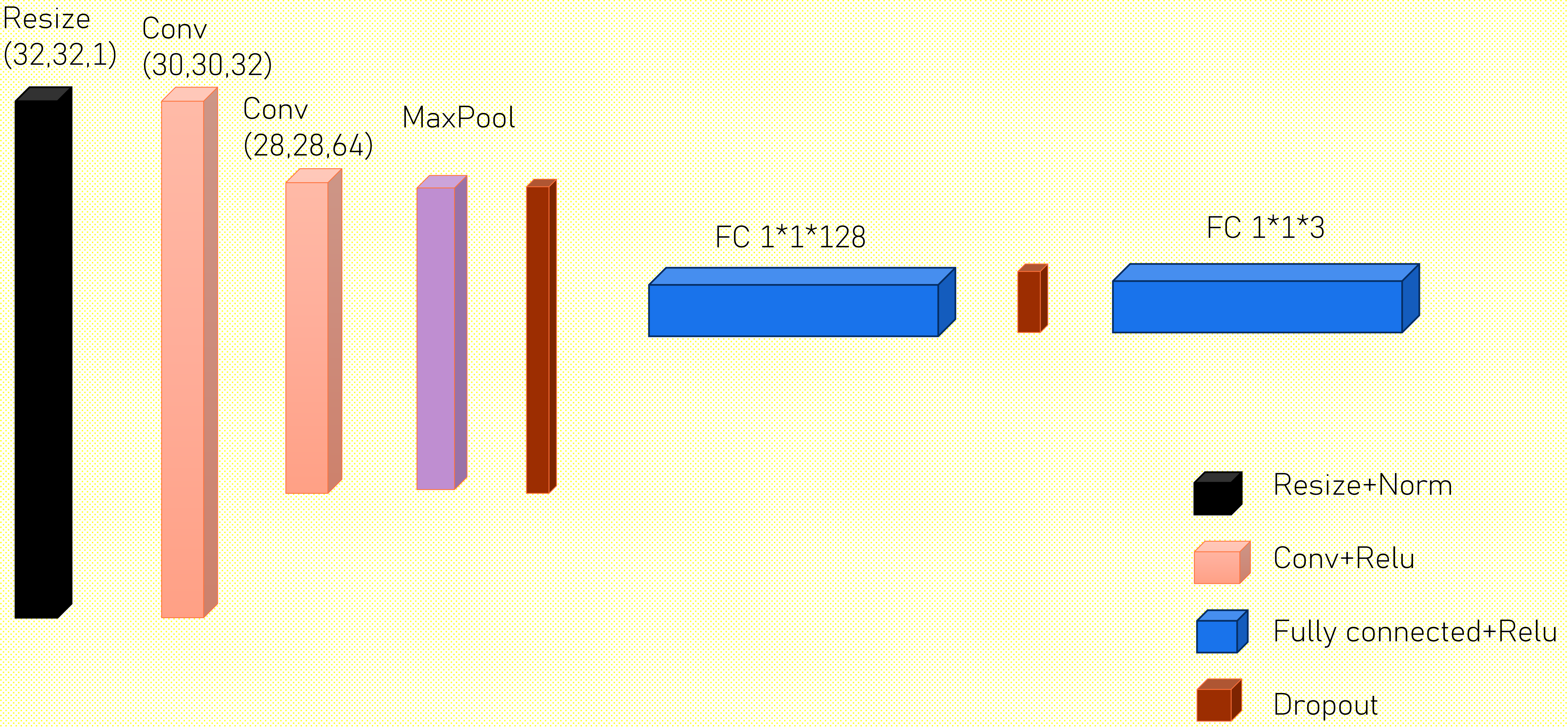


Bebop



Environnement

Modèle de Deep Learning : CNN



Modèle de Deep Learning : CNN

Processus d'entraînement :

- Optimiseur : Adam
- Learning rate : Exponential Decay Scheduler
- Loss : Sparse Categorical Crossentropy :

$$L = -\frac{1}{N}\sum_{i=1}^{N}\log P(y_i)$$

- Early stopping : après 5 epochs consecutive sans amélioration de la loss
- Entrainement sur 50 epochs

Modèle de Deep Learning : CNN

Processus d'évaluation :

- ❖ Accuracy :

$$Acc = \frac{TP + TN}{TP + TN + FP + FN}$$

- ❖ F1-score:

$$F1 = 2 \times \frac{P \times R}{P + R}$$

$$P = \frac{TP}{TP + FP}$$

$$R = \frac{TP}{TP + FN}$$

- ❖ Matrice de confusion:

Performances évaluées sur 10 entraînements de 50 epochs

Détection

Entraînement :

- Avec scheduler et early stopping : convergence du modèle

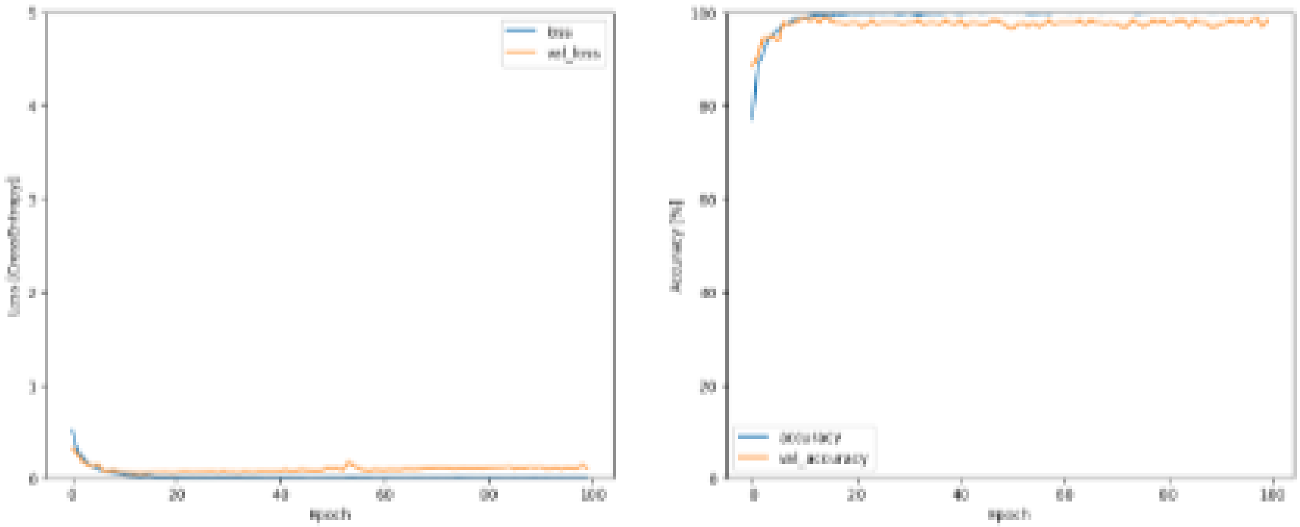
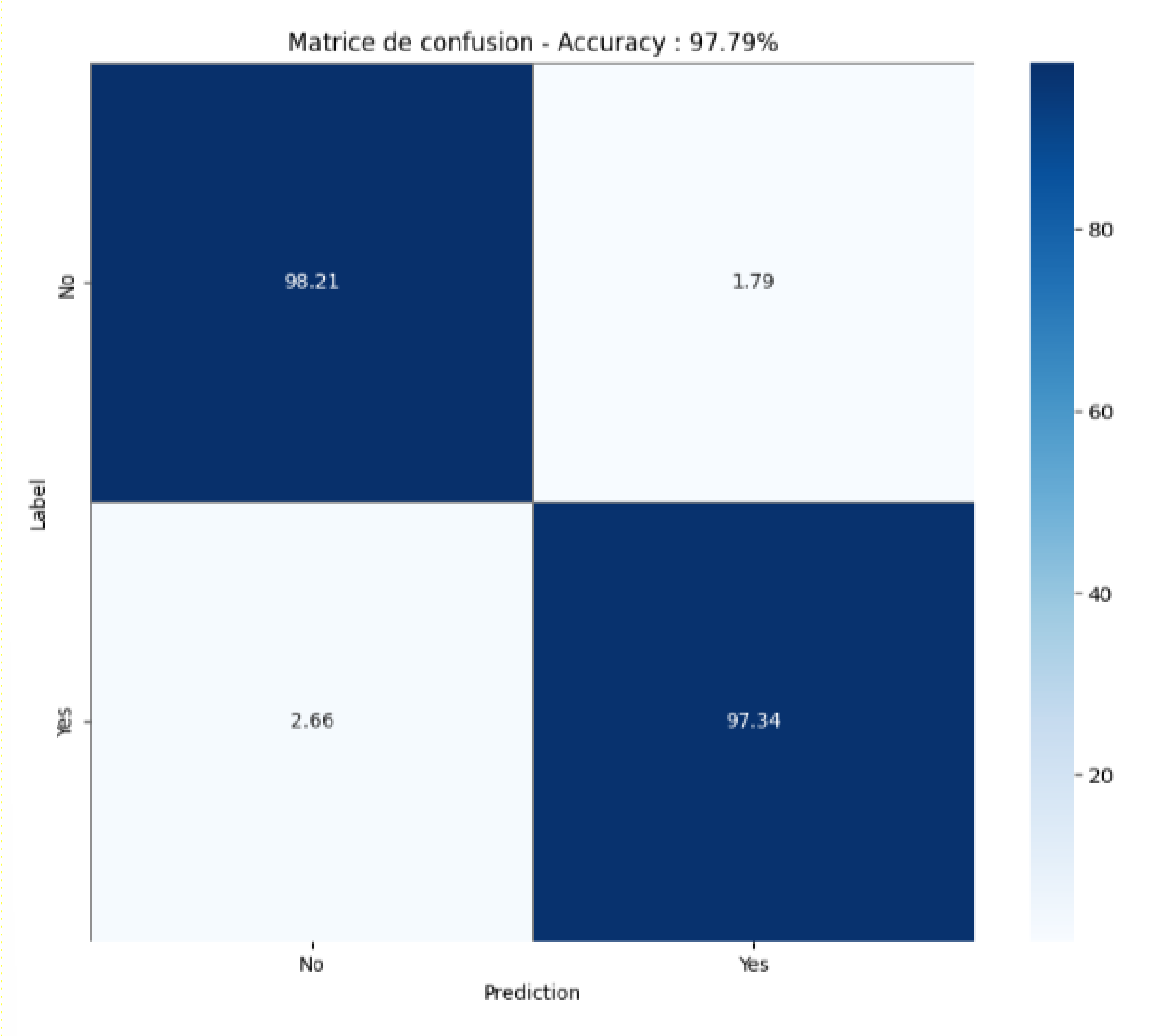


Fig. 6: Training loss and accuracy over epochs

Détection

Evaluation du modèle:



- Accuracy : 97,62%
- F1-score : 95,90%

Meilleures performances que le modèle de détection du papier [4]

Classification multiclass

Evaluation du modèle:

- Accuracy : 98,46%
- F1-score : 95,90%

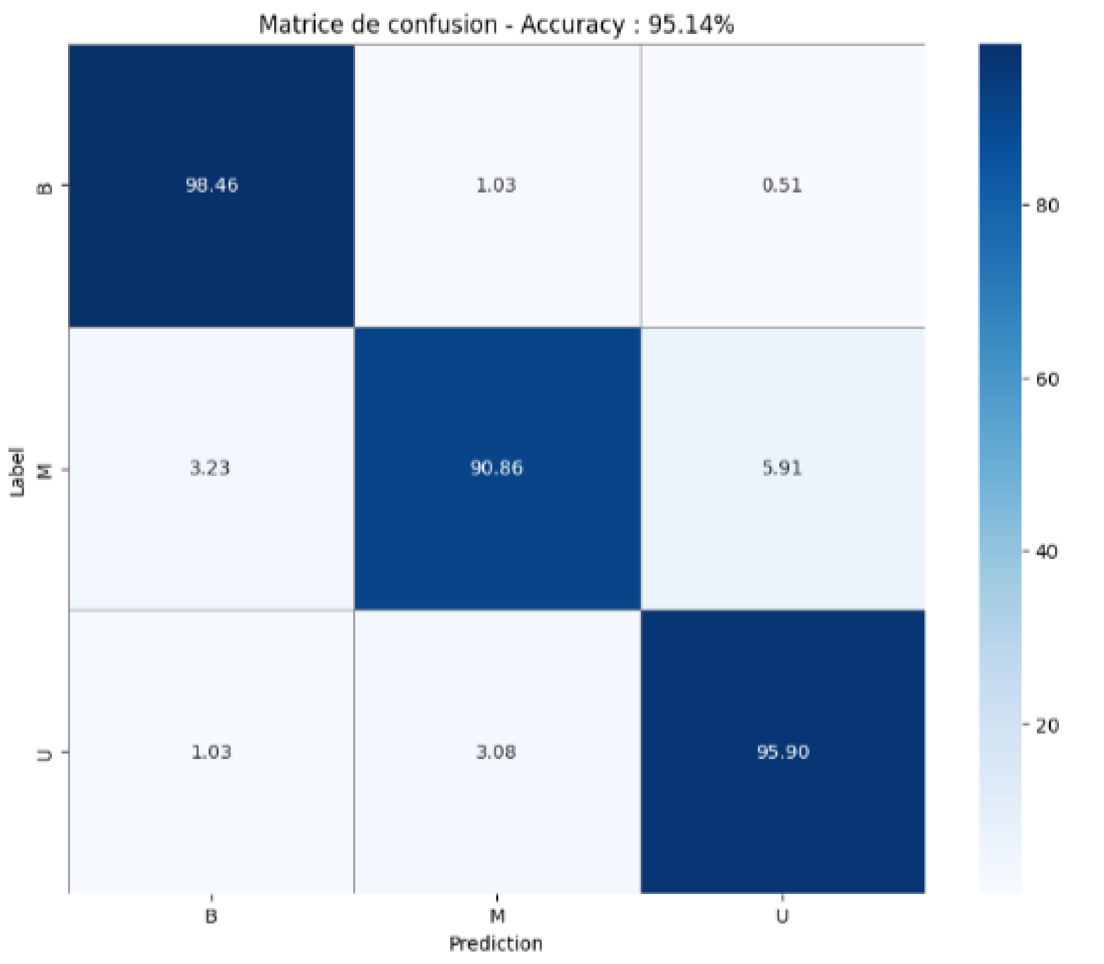


Fig. 8: Confusion matrix for the classification task

Meilleures performances que le modèle de classification du papier [4]

Interprétation des résultats

	Detection	Classification
Our model	97,62%	98,46%
Paper's model [4]	96,38%	92,94%

Amélioration significative de l'accuracy pour la classification multiclass

- ❖ Equilibrage des classes
- ❖ Normalisation des spectrogrammes
- ❖ Choix de l'architecture du model (Dropout, Conv, Dense)
- ❖ Choix des hyperparamètres (LR, Loss, EearkyStopping)

Spectrogrammes de Mel

Représentation proche des perceptions humaines , diminue les détails des hautes fréquences

Spectrogrammes Mel :

- Transformée de Fourier à court terme:

$$X(m,k) = \sum_{n=0}^{N-1} x[n]w[n-mR]e^{-j\frac{2\pi kn}{N}}$$

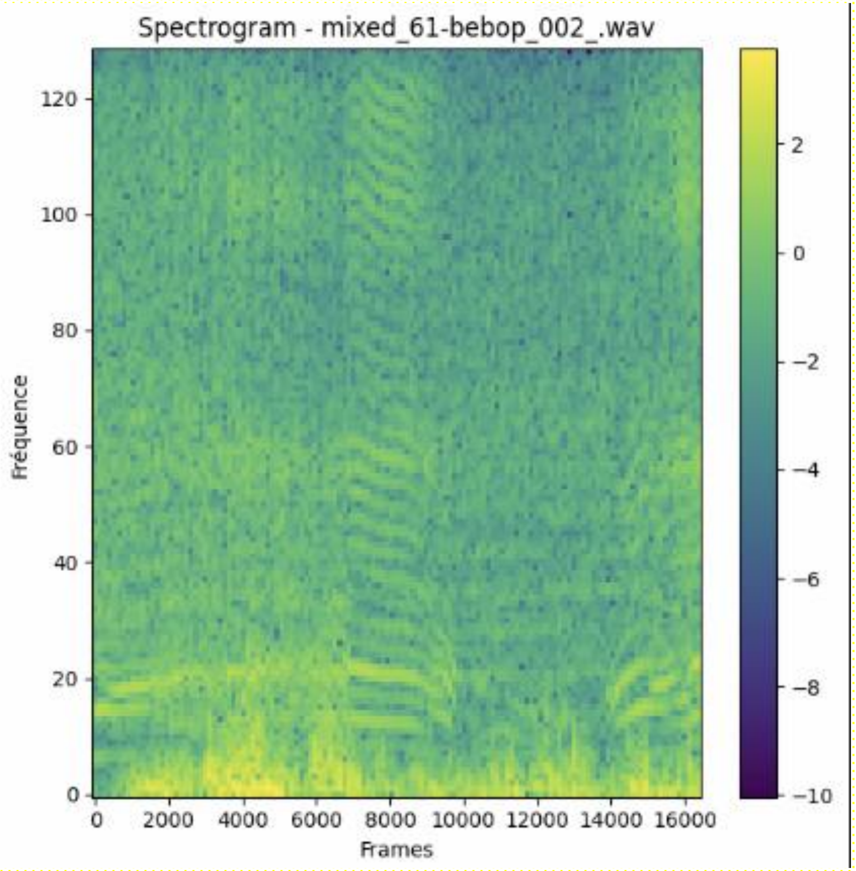
- Spectre en amplitude :

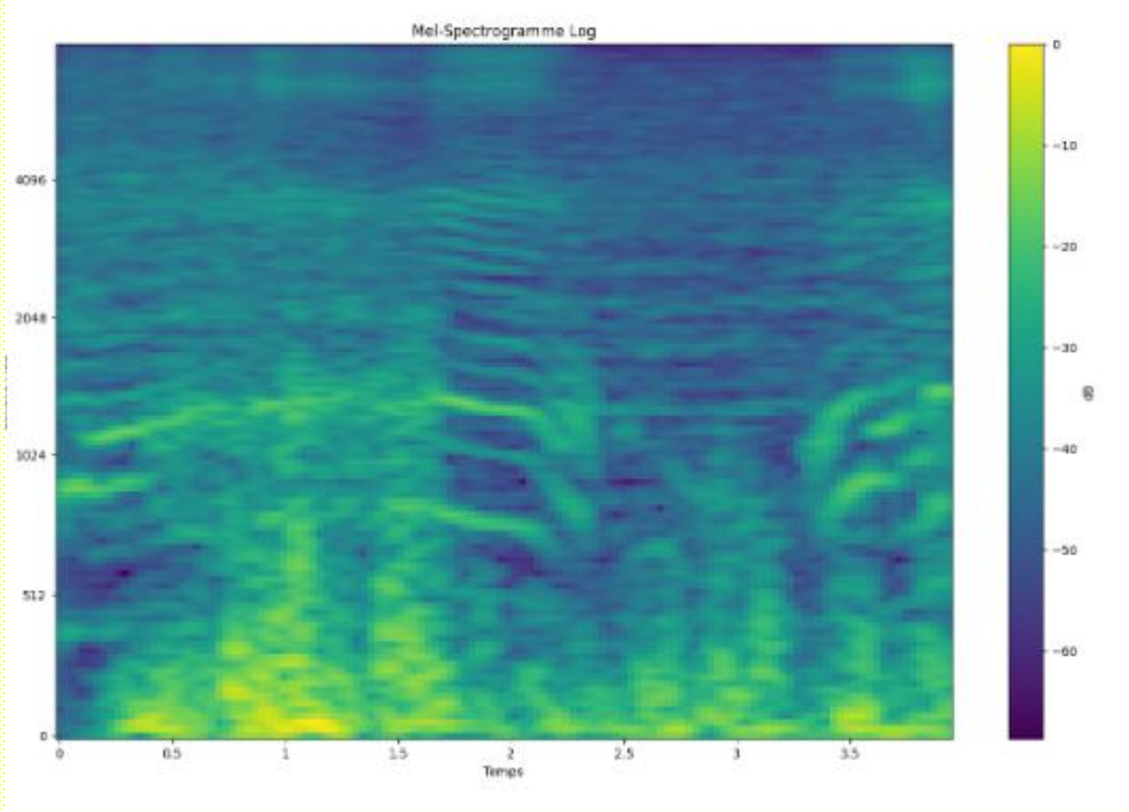
$$S(m,k) = |X(m,k)|^2$$

- Conversion des fréquences linéaire en échelle de Mel:

$$f_{\text{mel}} = 2595 \log_{10}\left(1 + \frac{f}{700}\right)$$

Spectrogrammes de Mel





Spectrogramme et Spectrogramme de Mel de Bebop

Détection et Classification sur spectrogrammes Mel

	Détection	Classification
Spec + CNN	97,62%	98,46%
Mel-Spec + CNN	97,14%	87,4%

Chute significative de performances pour la classification avec les spectrogrammes de Mel

Interprétation des résultats

- Contenu des plus hautes fréquences différencie les drones entre eux
- Contenu des plus basses fréquences discrimine les drones et l'environnement

- Filtre de Mel adoucie la représentation fréquence : perte de détails

Retour sur le projet

- Implémentation d'un algorithme adaptable de détection et de classification acoustique de drones

- Modèle plus performant que le papier de référence [4]

- Etudier les spécificités des représentation et Spectrogrammes et spectrogrammes de Mel et leur impacte sur les performances du modèle

Conclusion

- Méthode combinant représentation Temps-fréquence et Deep Learning montre des bons résultats
- Le choix de la TFT impacte les performances du modèle
- Faible coût , peu intrusive, modèle performant : classification acoustique combinant TFT et deep learning est une solution efficace aux problématiques soulevées par l'essor des drones de loisir

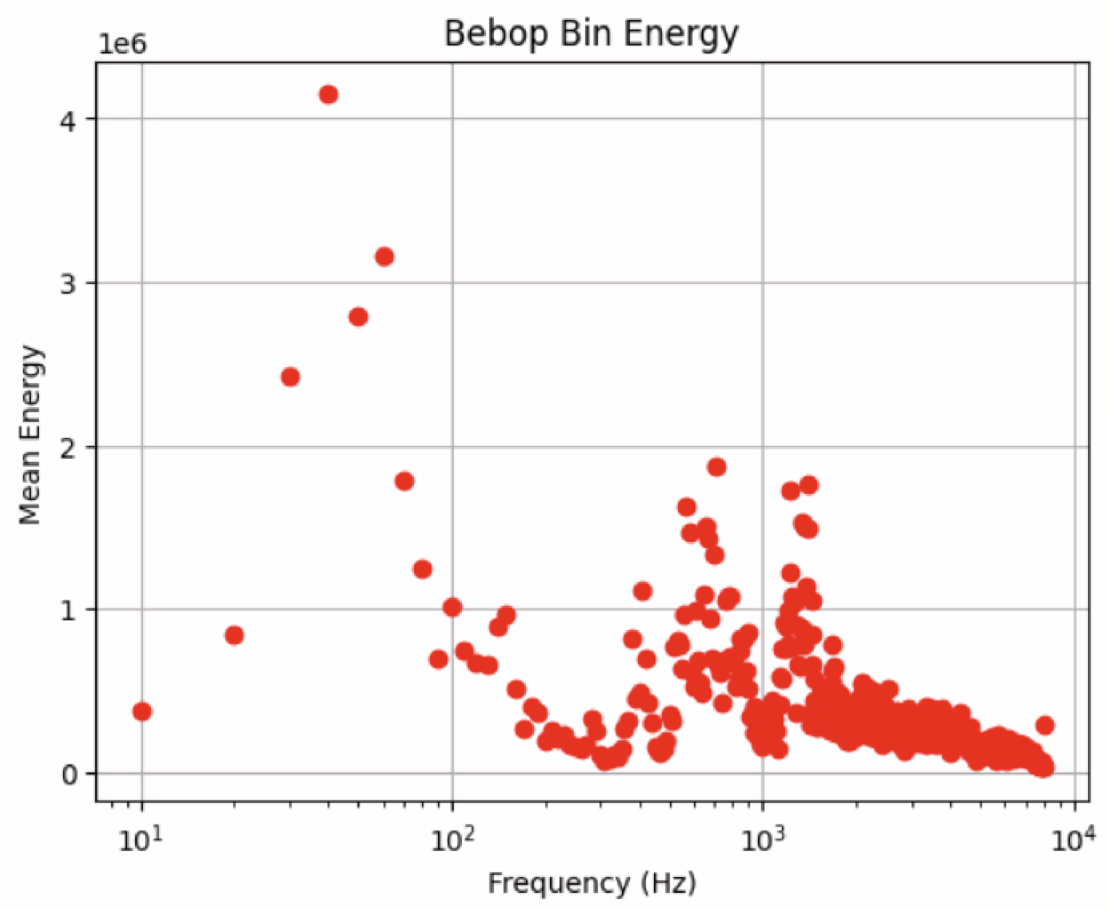
Références Bibliographiques

- [1] Maxime Gautier, 14 janv. 2025 , https://fr.statista.com/statistiques/607544/croissance-prevue-marche-drones-professionnels-monde/
- [2] T. Guardian, “P. daniels, venezuela’s nicolas maduro survives apparent assassination attempt,” https://www.theguardian.com/world/2018/ aug / 04/ nicolas - maduros - speech - cut - short - while -soldiers-scatter., 2018.
- [3] B. TAHA and A. SHOUFAN, “Machine learning-based drone detection and classification: State-of-the-art in research,” IEE, 2019.
- [4] A. M. A. A.-A. Sara Al-Emadi, Abdulla Al-Ali, “Audio based drone detection and identification using deep learning,” 2019 15th International Wireless Communications Mobile Computing Conference (IWCMC),2019
- [5] Sungho Jeon1, Jong-Woo Shin, Young-Jun Lee, Woong-Hee Kim, YoungHyoun Kwon, and Hae-Yong Yang "Empirical Study of Drone Sound Detection in Real-Life Environment with Deep Neural Networks", 2017arXiv

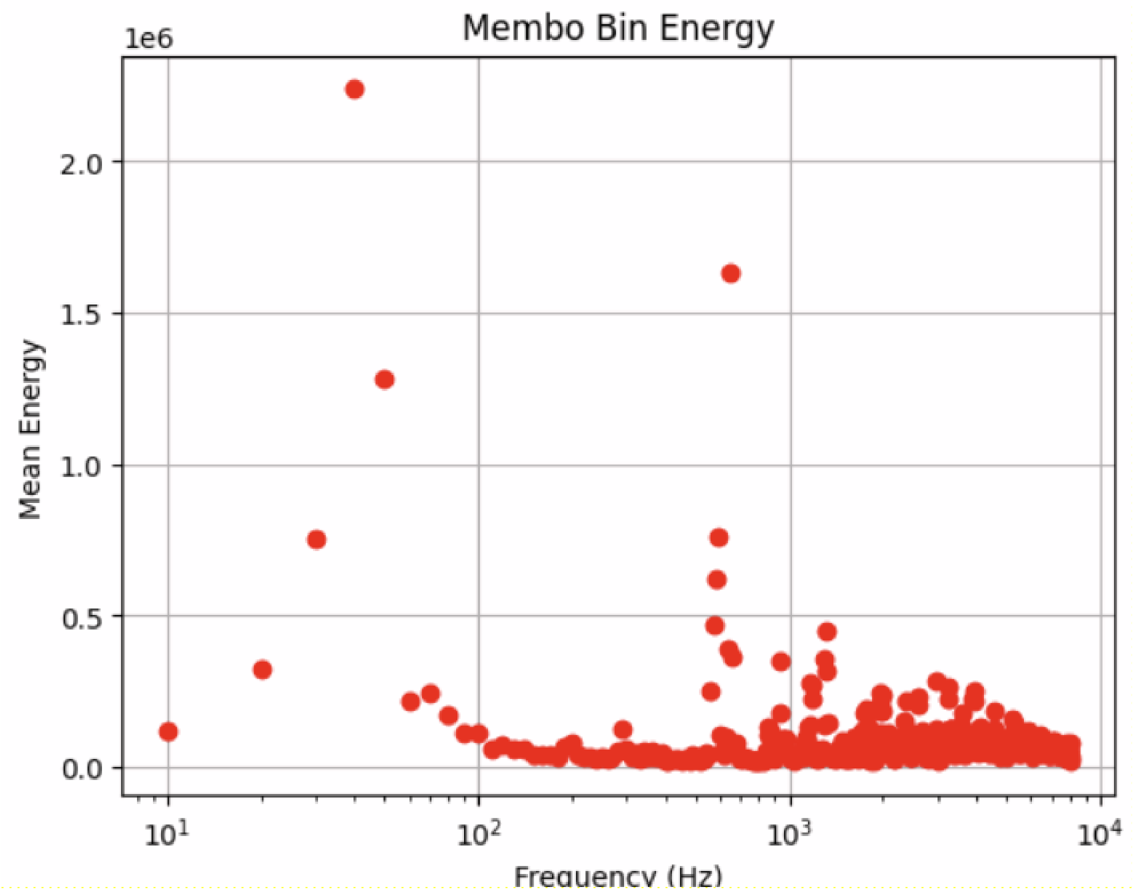
Références Bibliographiques

- [6] parrot.com
- [7] V. C. Georgopoulos, “Advanced time-frequency analysis and machine learning fo pathological voice detection,” 12th International Symposium on Communication Systems, Networks and Digital Signal Processing (CSNDSP), 2020.
- [8] R. B. Constantin CONSTANTINESCU, “An overview on sound features
- in time and frequency domain,” Sciendo, 2023.

Annexe



Bebop Bin Energy
1e6
Mean Energy
Frequency (Hz)



Membo Bin Energy
1e6
Mean Energy
Frequency (Hz)