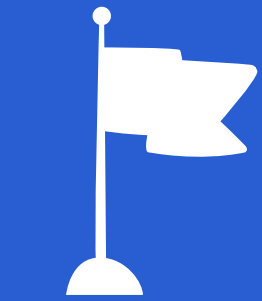


MUSIC GENRE CLASSIFICATION

Projecte de Xarxes Neuronals
i Aprenentatge Profund

Arnau Ortiz
Martí Mussons
Eloi Aranda

Índex



Starting
Point

Introducció
Arquitectura

Audio processing
Embedding
Ensambling methods



Problemes i
Millores
proposades



Resultats
i anàlisi

Precision
Recall
F1 Score
Accuracy

Conclusions
Possibles millores



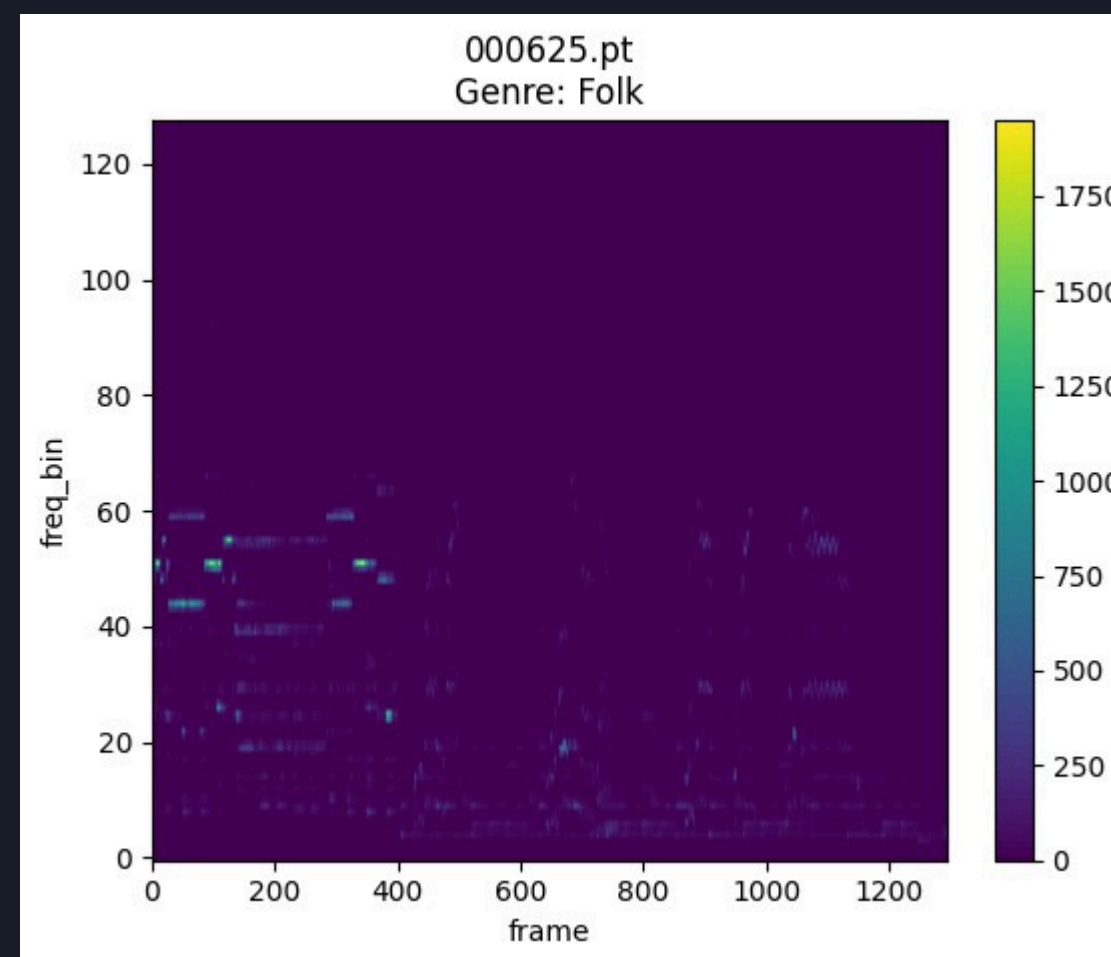
Anàlisi

Starting Point

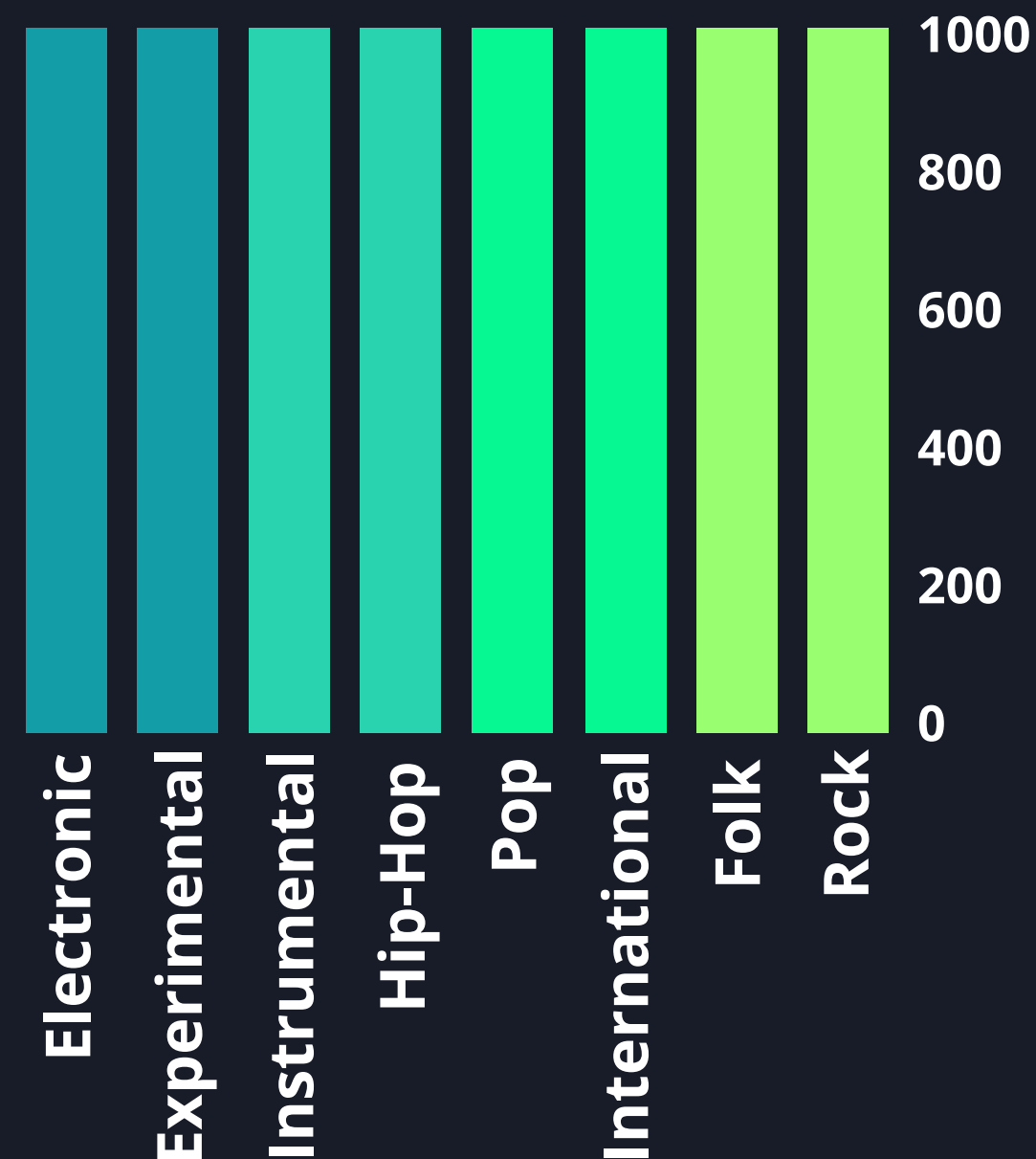
Introducció



Learning Rate = 0.001
DataAugmentation = False
Optimizer = Adam
Scheduler = ReduceLROnPlateau
Epochs = 30

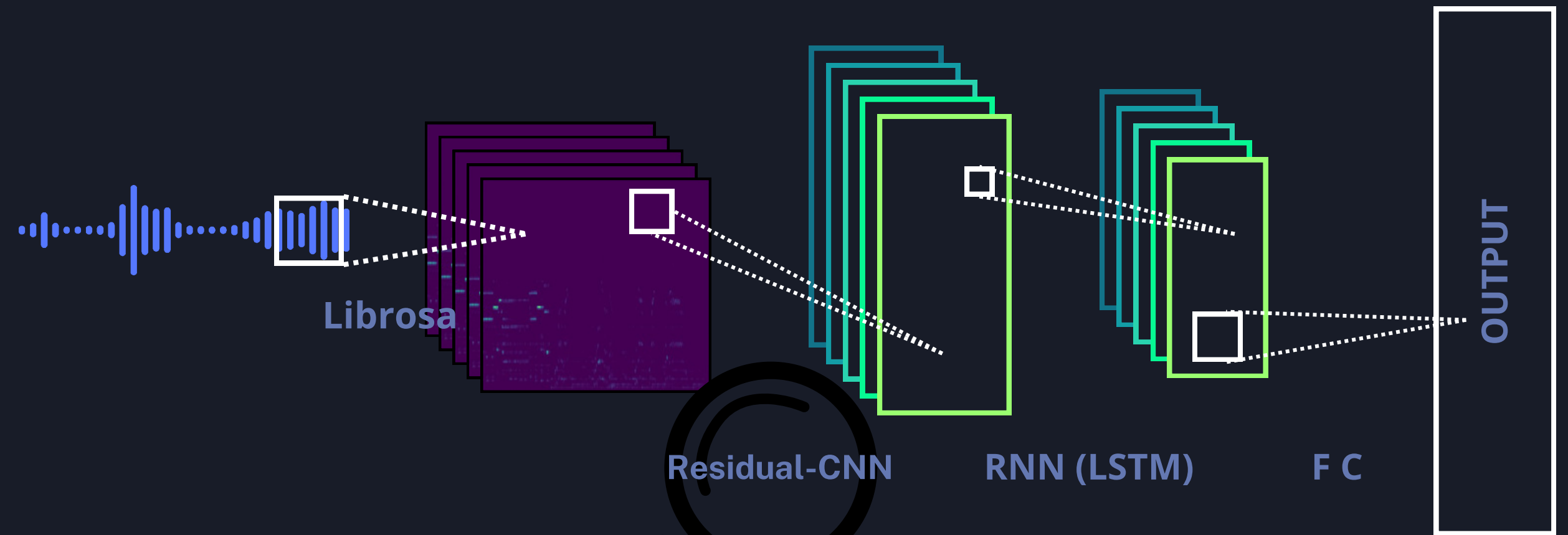


Spectrograma Mel



Starting Point

Arquitectura

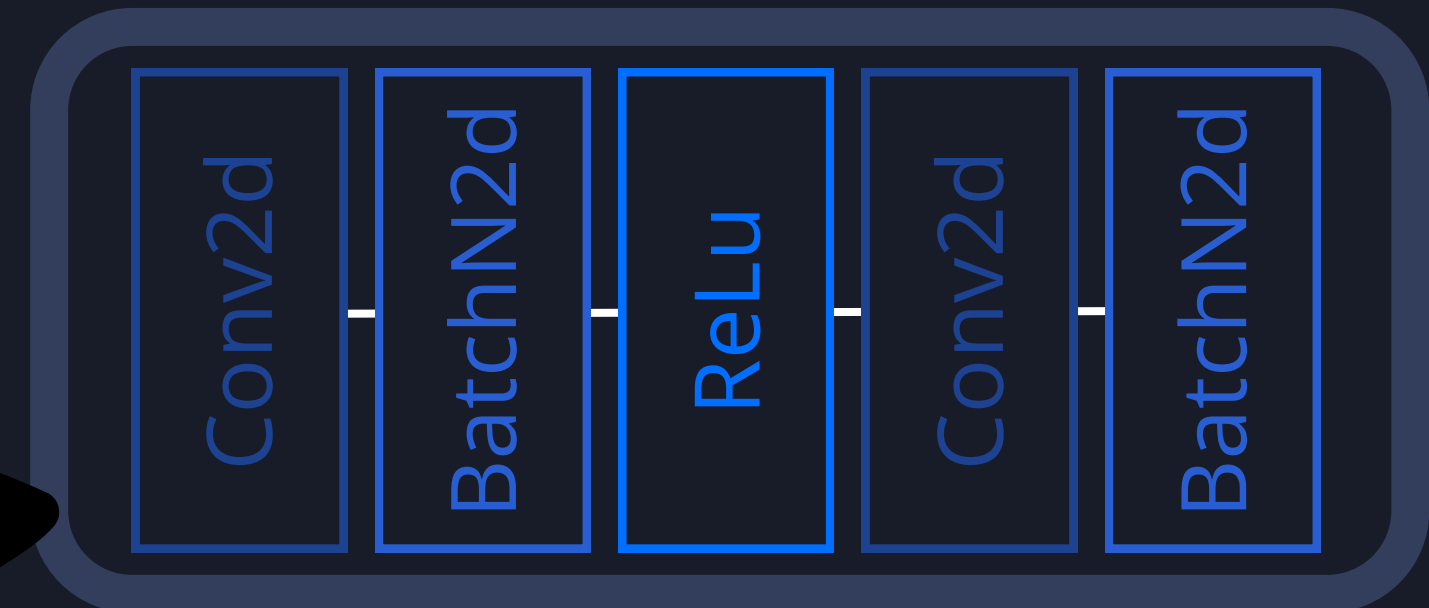


Residual Convolutional Block



Accuracy = 55% | Test Loss = 1.1

ResBlock2d



Propostes de millora



**Audio
processing**

Embeddings

**Ensamble
methods**

Audio processing Librosa

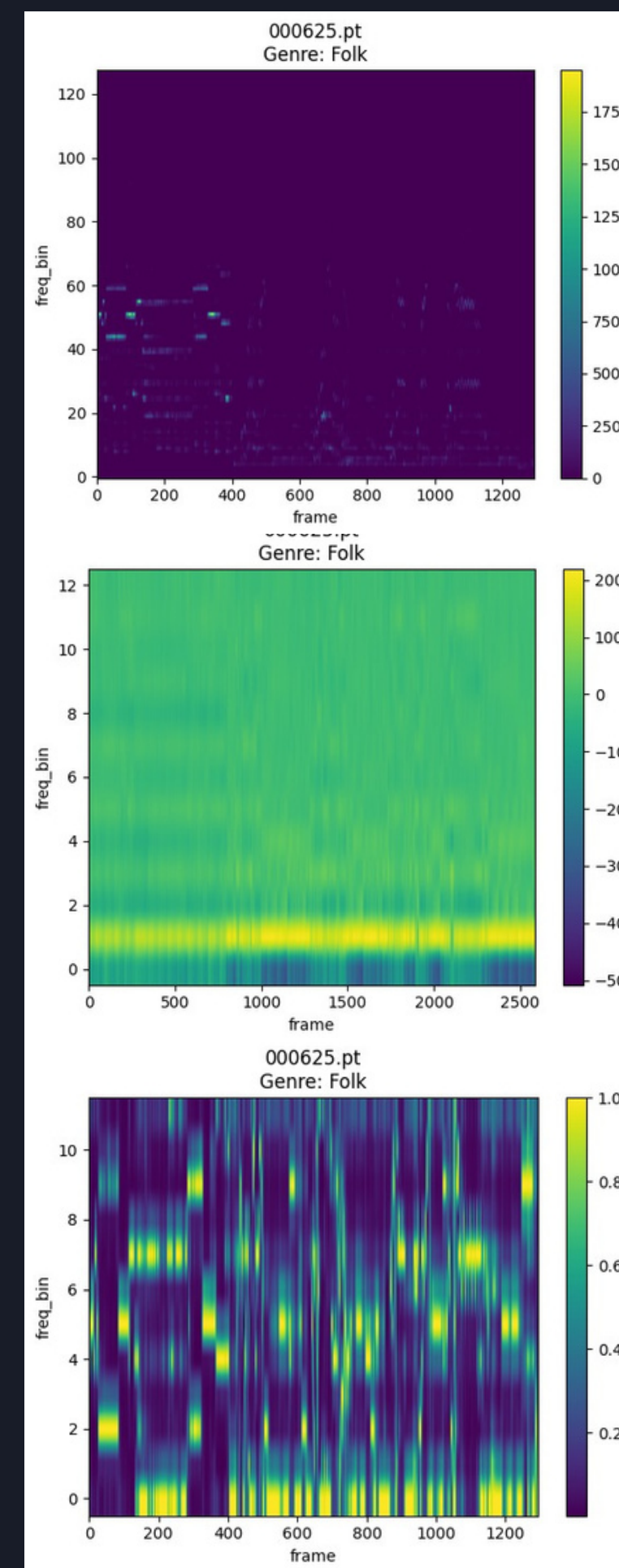


Track de 30s

= Spectrograms

+ MFCCS

+ Chroma



Audio Processing

MFCCs

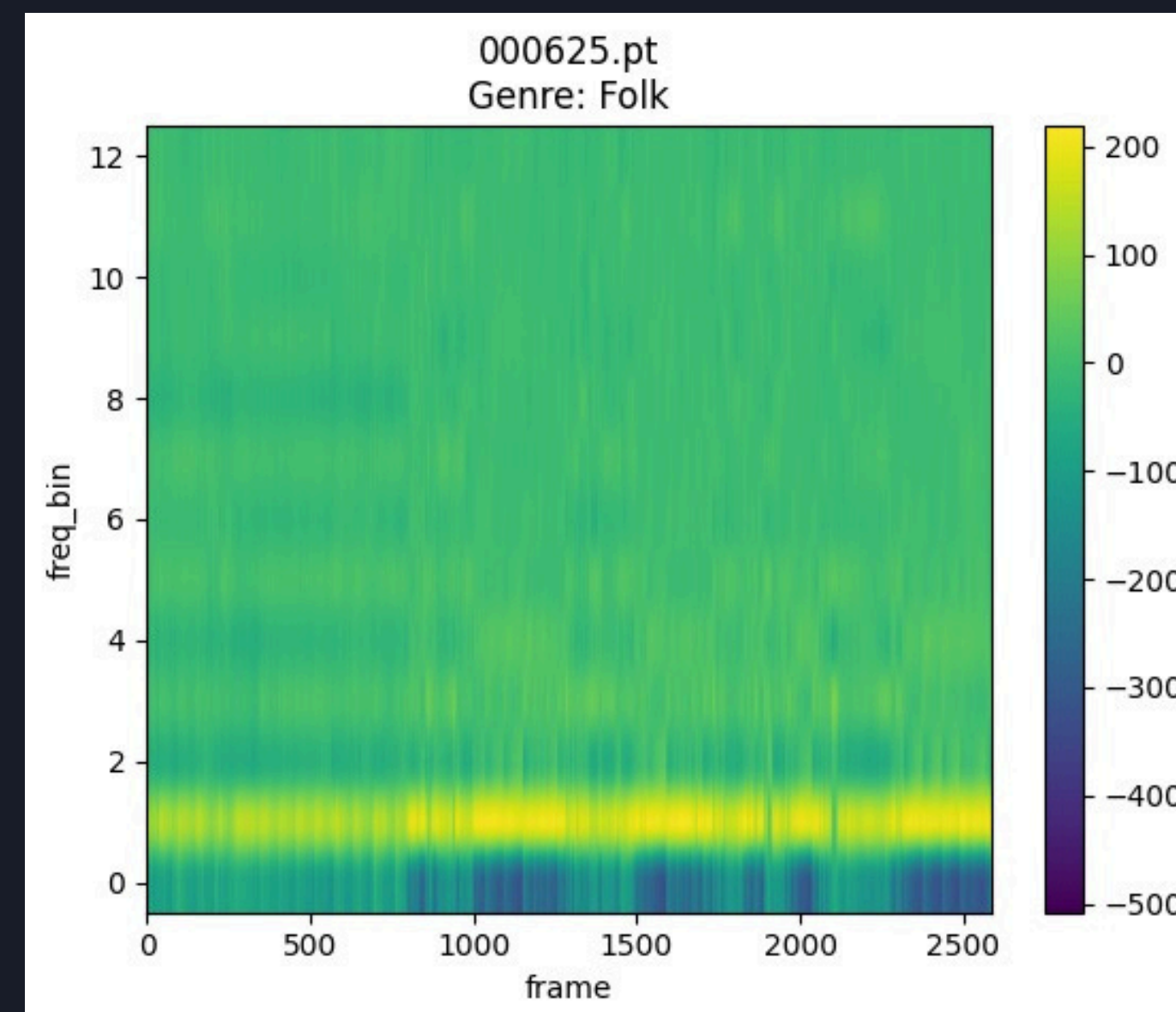
Càlcul



MFCC

Propietats

- Robust
- Compacte
- Melodia
- Timbre

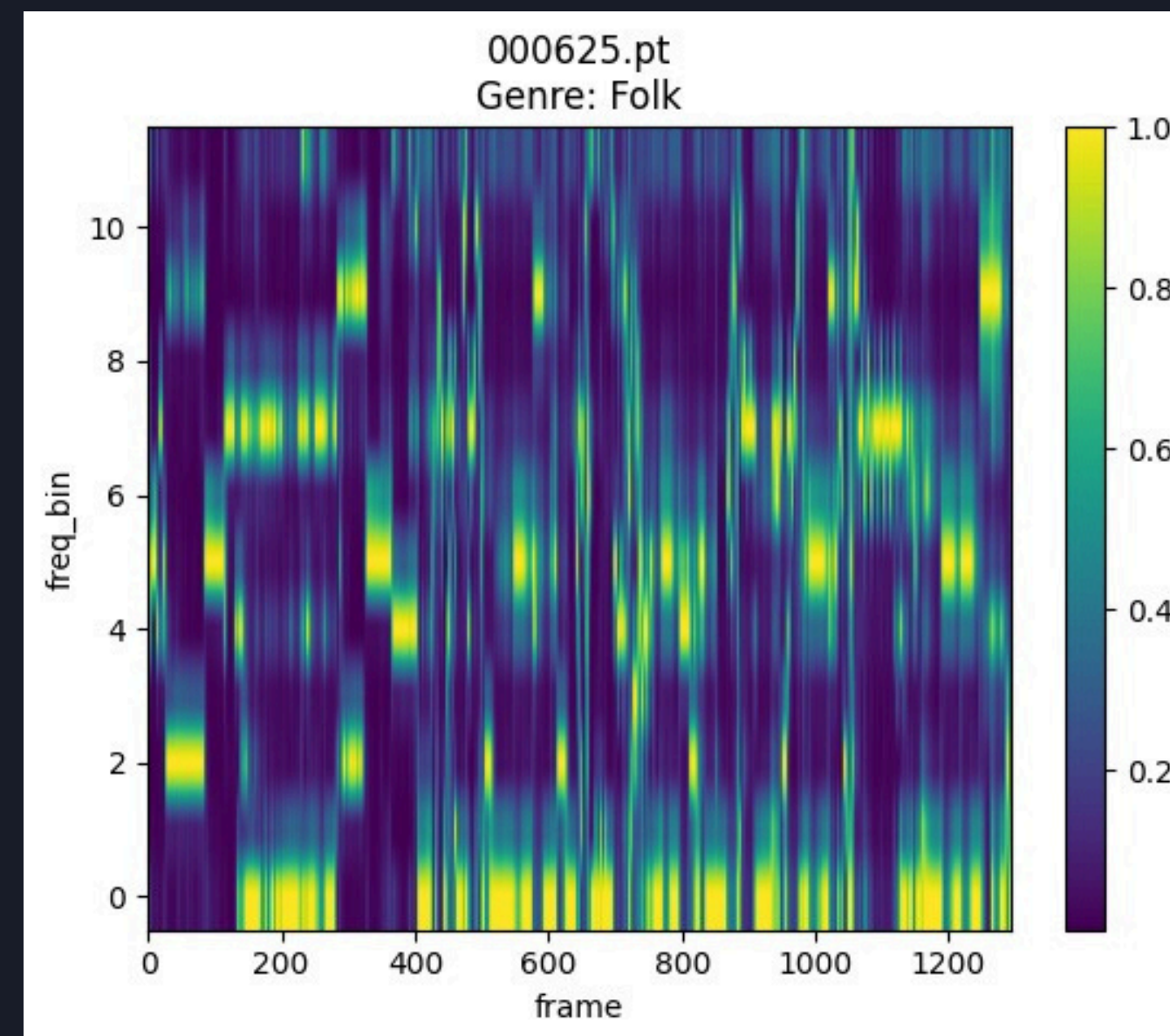


Audio Processing Chroma

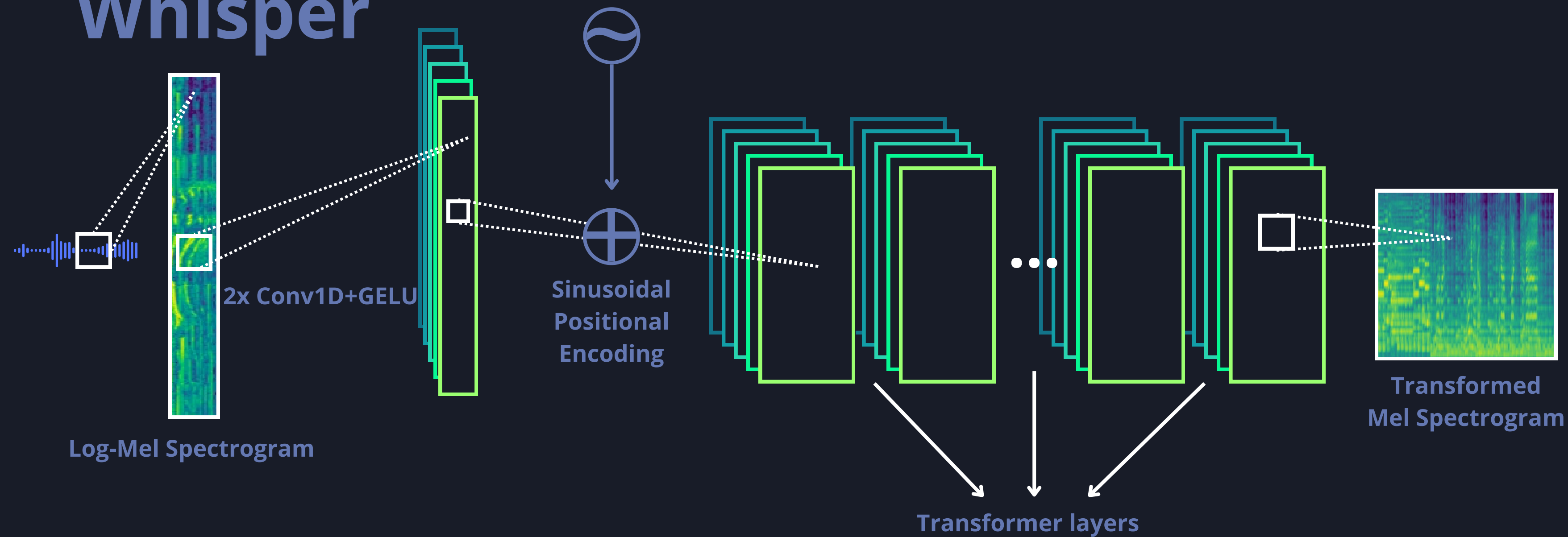


Propietats

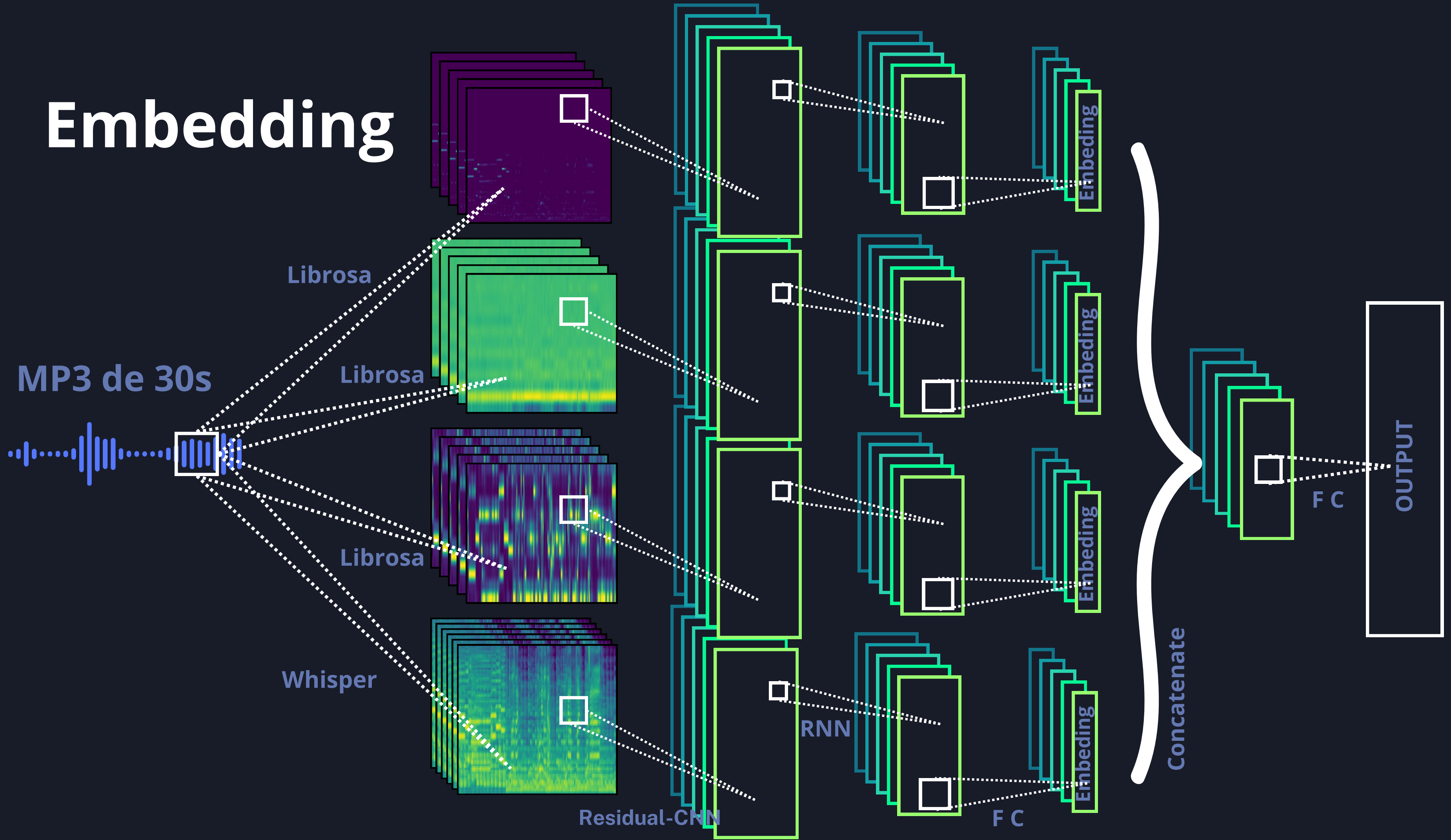
- Notes
- Armonia
- Tonalitat musical



Audio processing Whisper

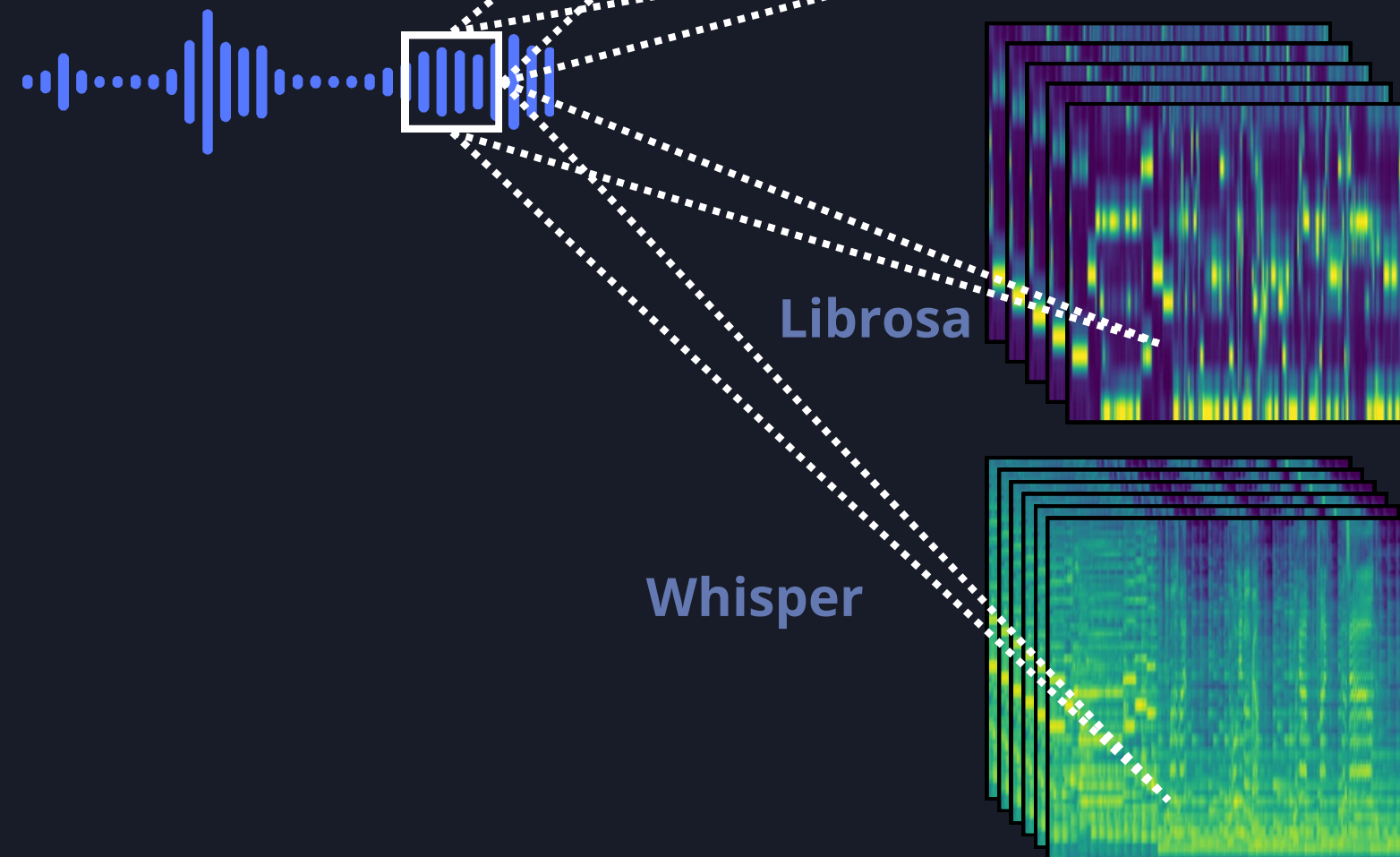


Embedding



Ensembling methods

MP3 de 30s



RNN del Starting point

RNN del Starting point

RNN del Starting point

RNN del Starting point

SPECTROGRAMS
PREDICT

FMCCS
PREDICT

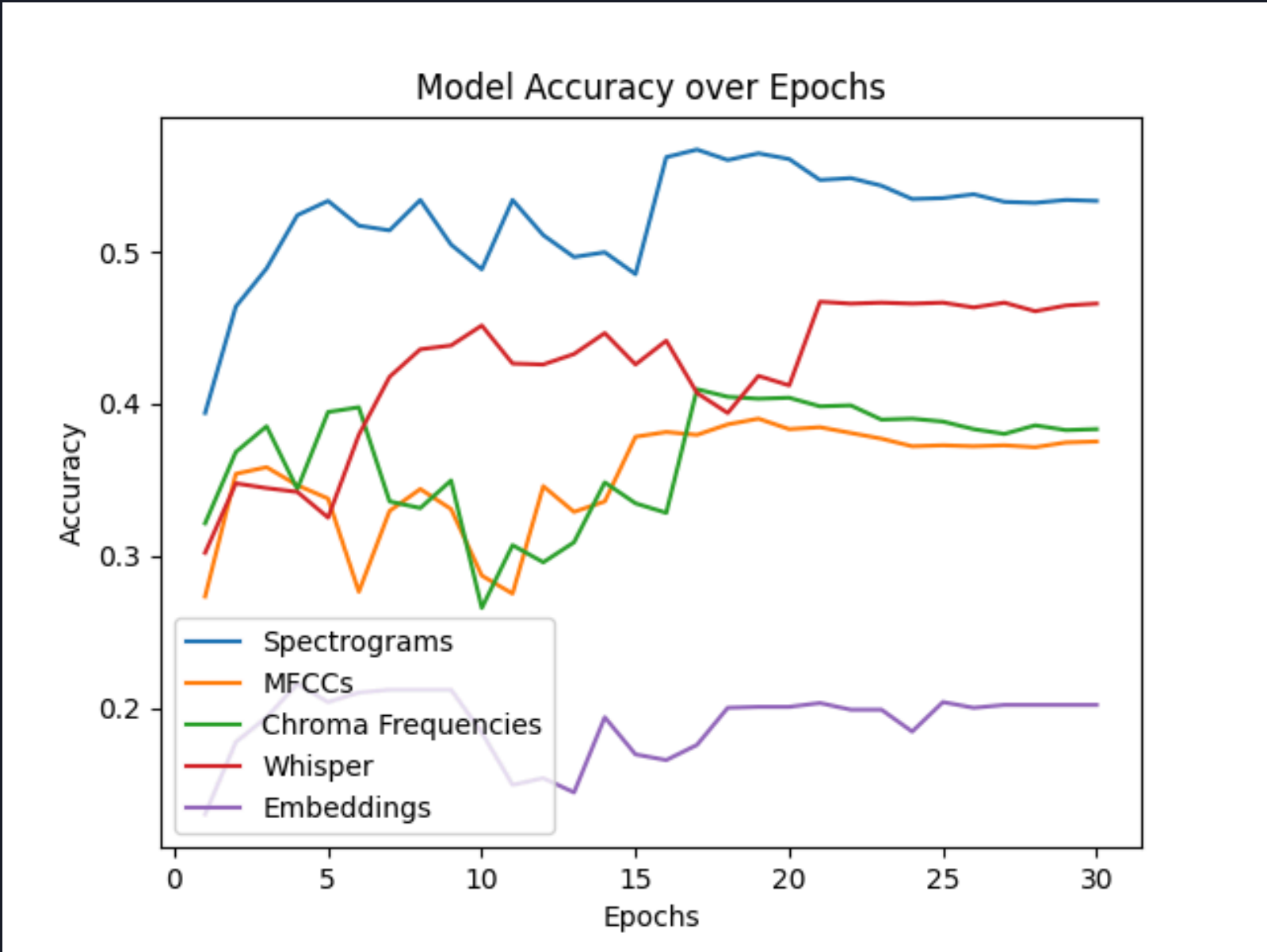
CHROMA
PREDICT

WHISPER
PREDICT

Majority Vote
Highest Probability
Highest Mean Probability

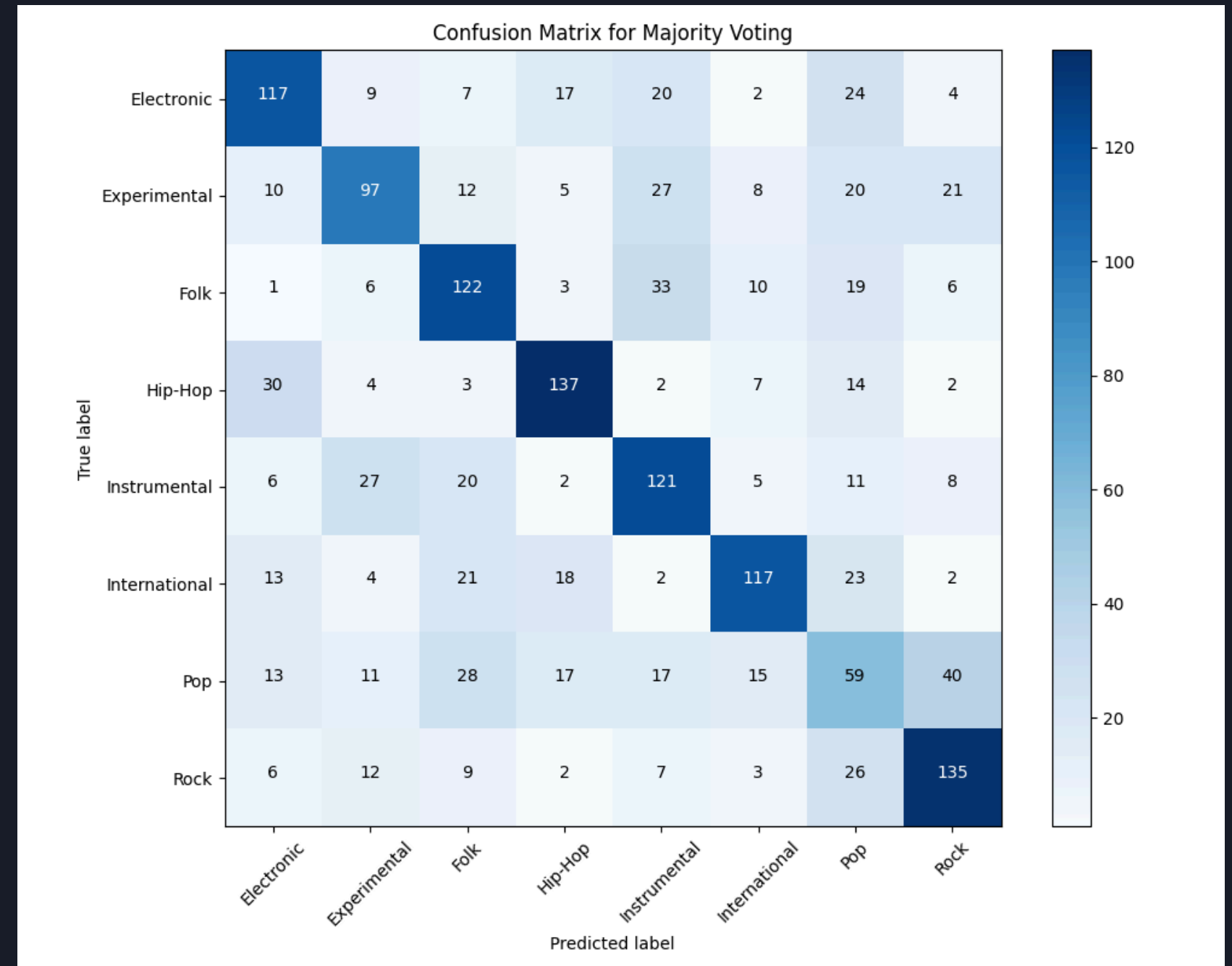
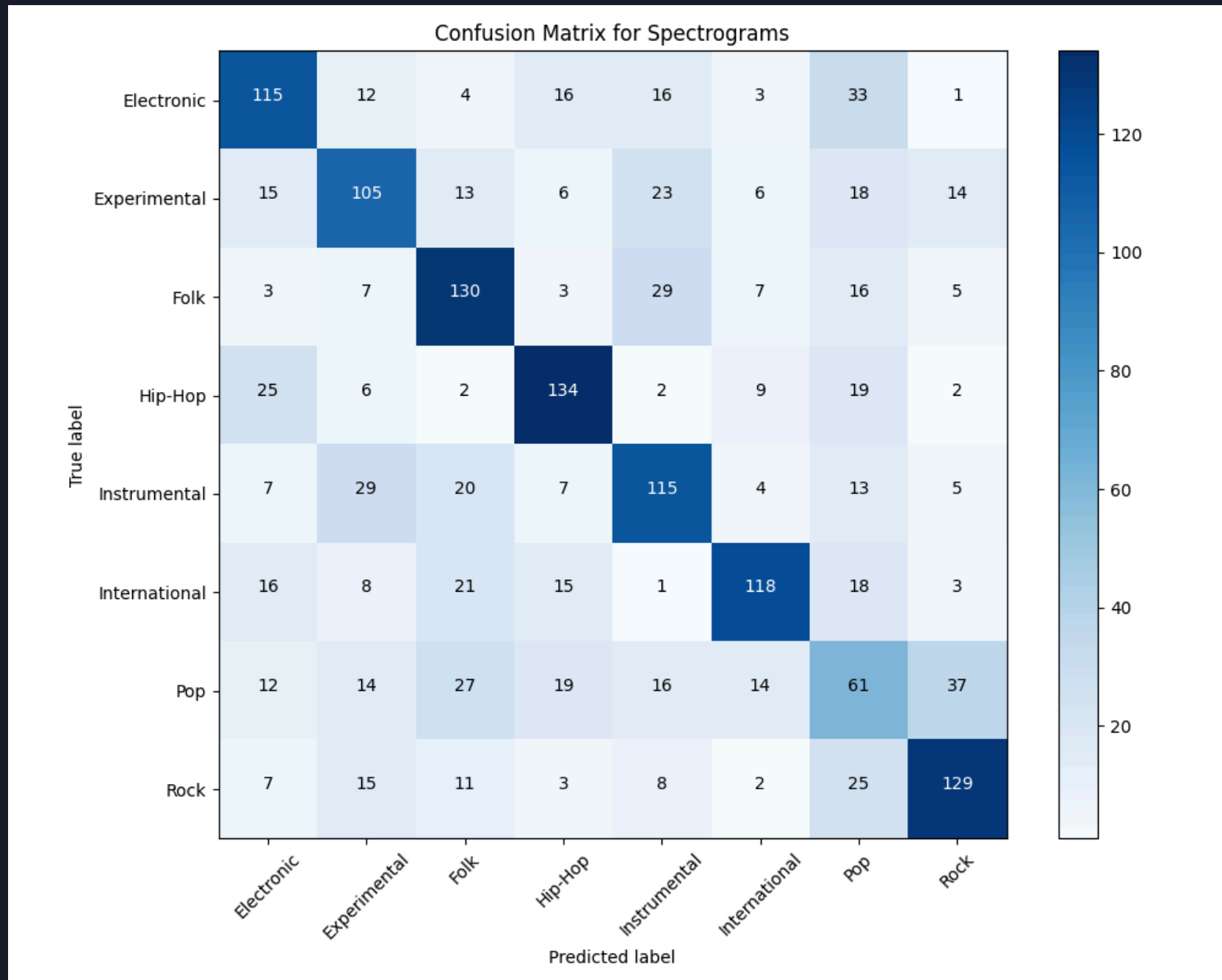
FINAL
PREDICT

Resultats



	Precision	Recall	F1 Score	Accuracy
2. Spectrograms	55	55	55	55
MFCCs	40	39	39	39
Chroma	42	41	41	41
Whisper	46	47	46	47
Embeddings	2	13	4	20
1. Majority Voting	57	57	57	57
Higher Mean Probability	49	49	49	49
3. Confidence	53	54	53	57

Resultats



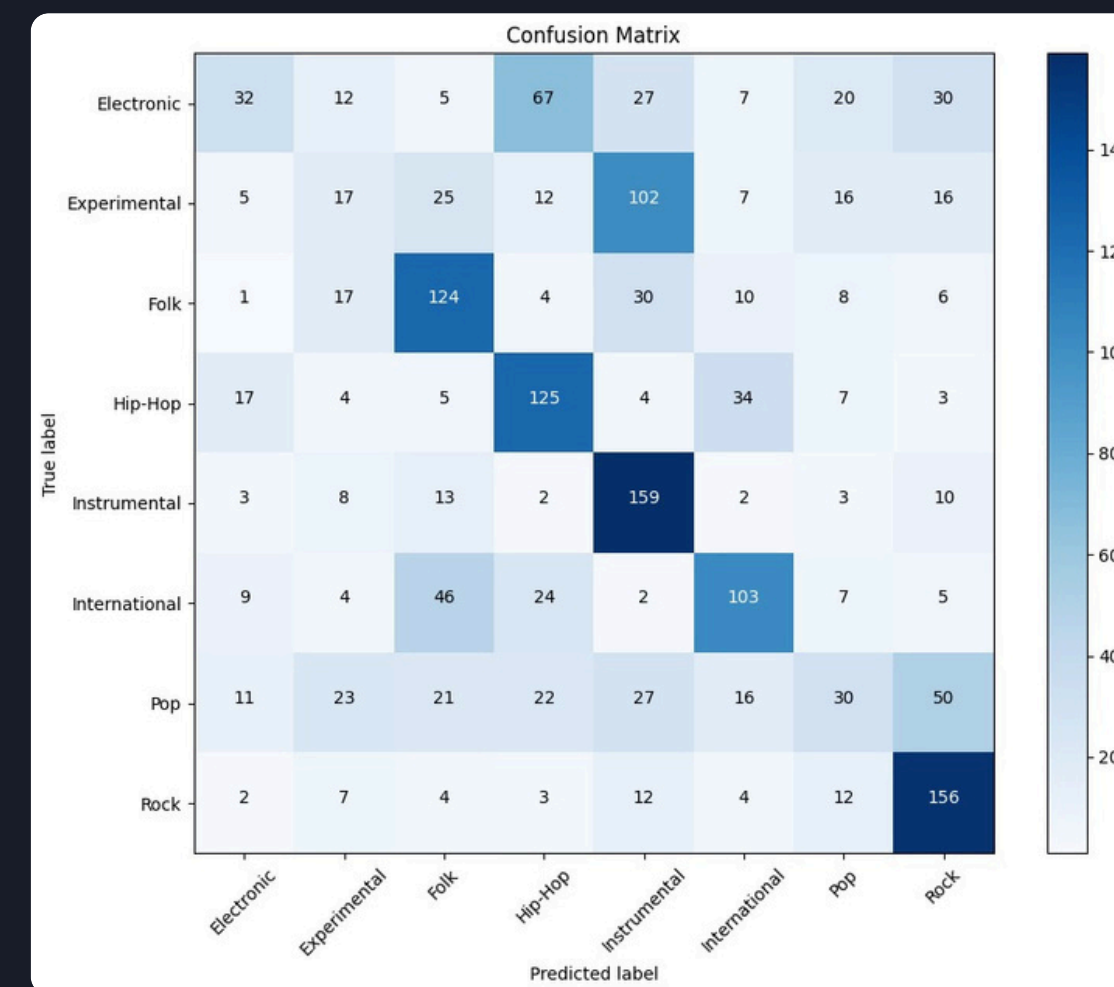
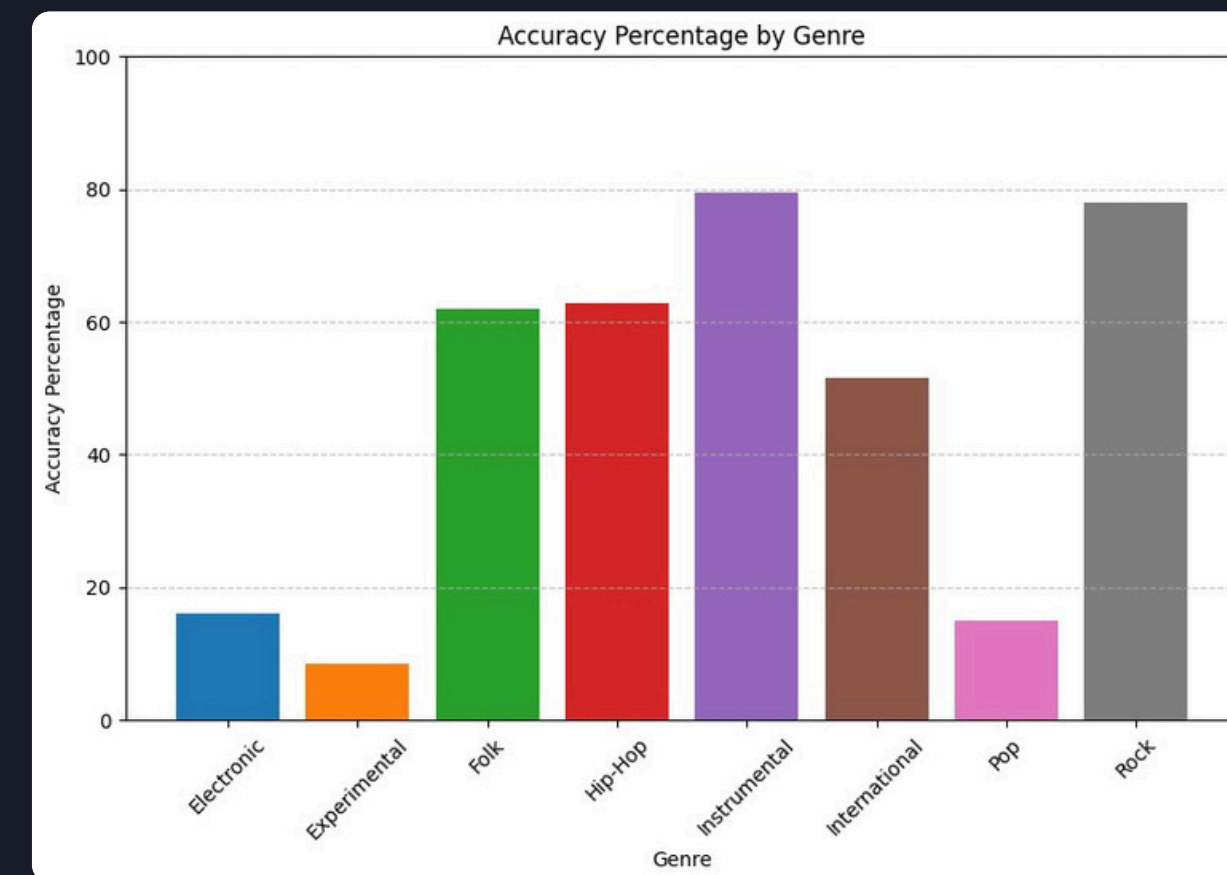
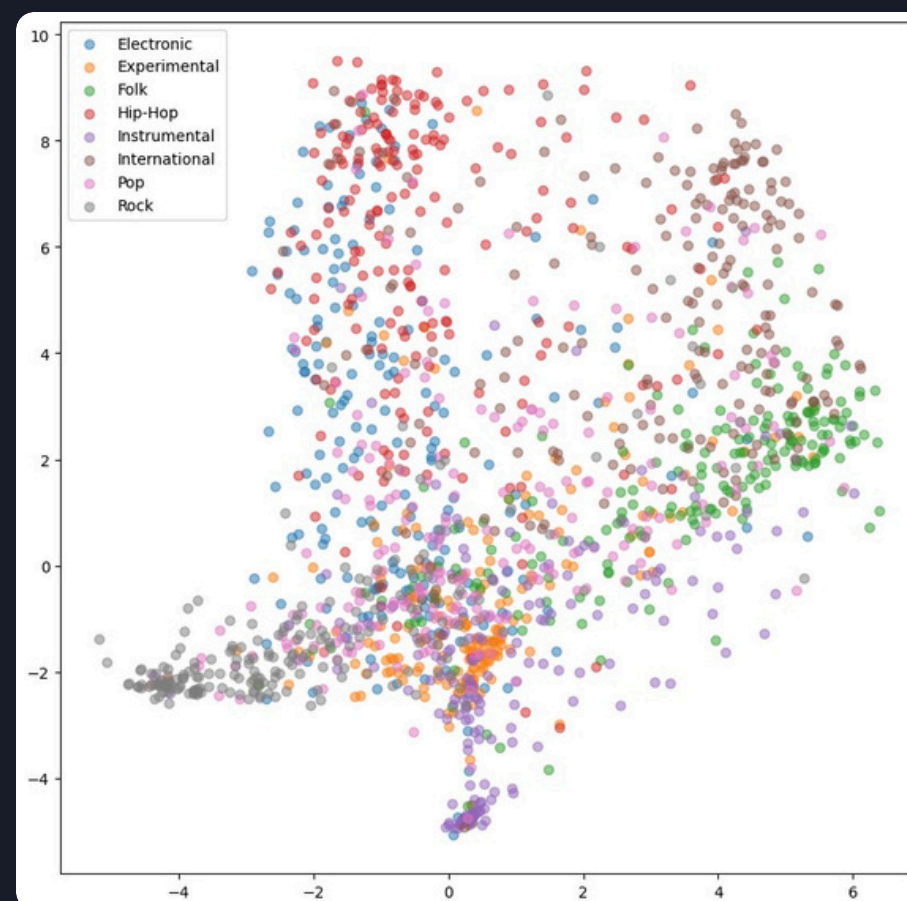
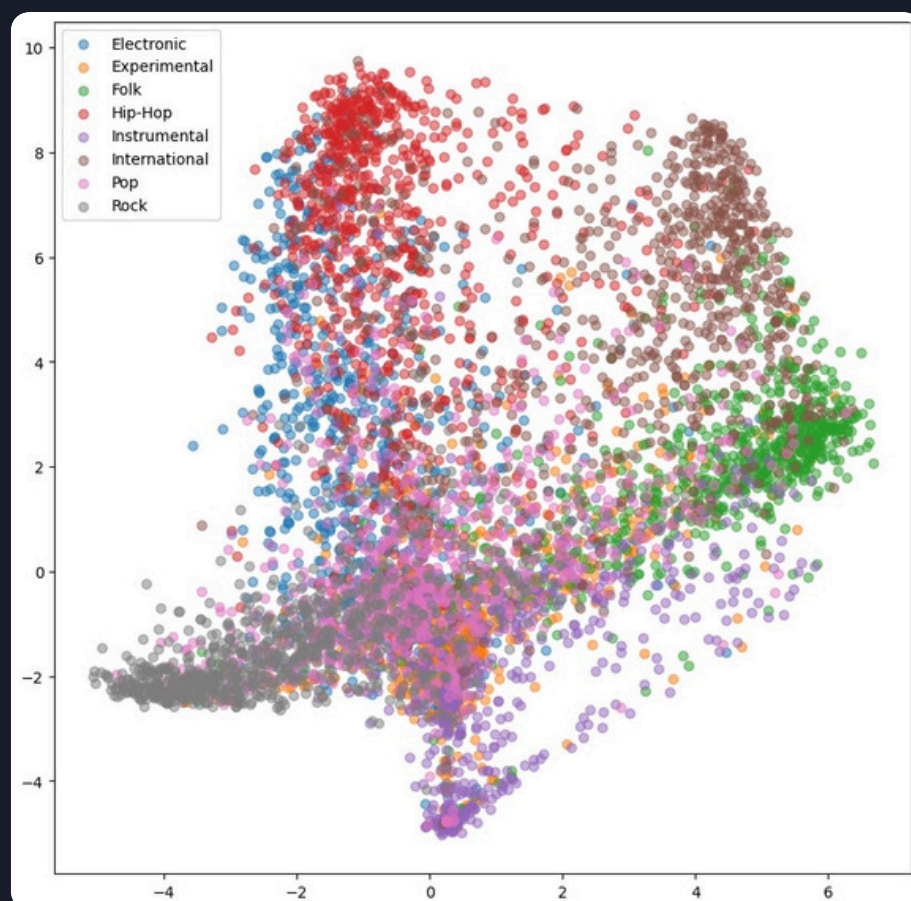
Conclusions



Majority Voting

Train

Test



**Possibles
milliores**

**Attention
Mechanisms**

**Millor data (data
augmentation?)**

**Més data
processing**

