Bird Call Feature Extraction Analysis Using Signal Processing

João Soares
João Vieira
Tiago Beaulieu

Table of Contents

INTRODUCTION
Why Identify Bird Species Through Sound?

Non-Invasive Monitoring

Allows researchers to monitor bird populations without disturbing their natural habitats.

Biodiversity Assessment

Helps in tracking species diversity and abundance in various ecosystems.

Large-Scale Automation

Scales up research efforts with AI-powered tools for identifying multiple species simultaneously

INTRODUCTION

The chosen species were decided based on distinct callings

King Penguin

Red-tailed Hawk

Greater Prairie Chicken

Magpie Goose

DATA PREPROCESSING

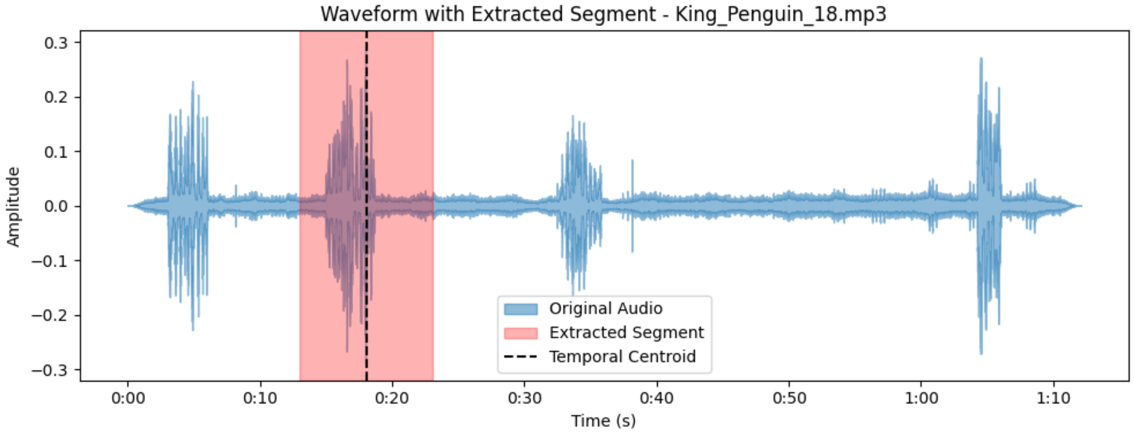
Several Preprocessing mechanisms were applied to the Data

1. **Application of Noise Reduction**

   Noise Reduction Using an External Library

2. **Segmenting Sounds: Temporal Centroid**

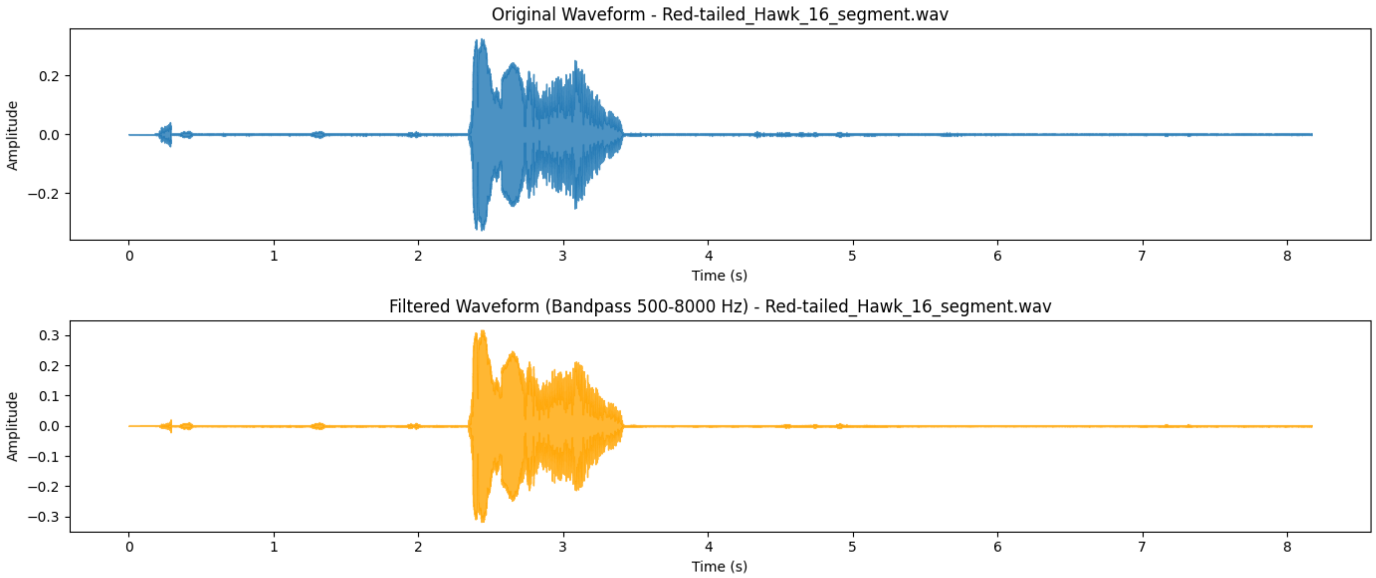
   Segmenting audios into 10 second segments centered about the Temporal Centroid



DATA PREPROCESSING
Application of Band Pass Filter

3. Application of a Band Pass Filter

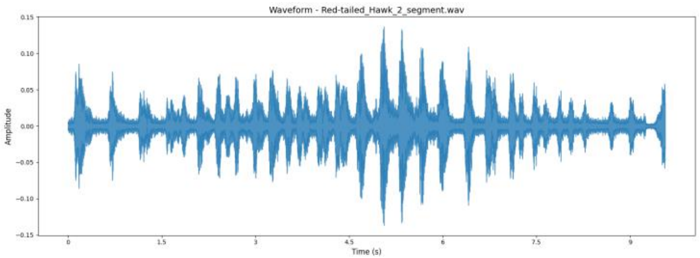
Application of band pass filter between 500 Hz and 8000Hz (frequencies where bird calling tends to happen)



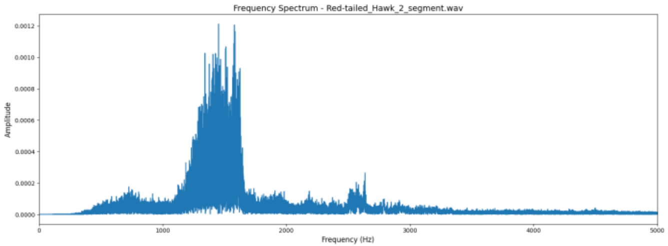
FEATURE EXTRACTION

Features were extracted from the Time and Frequency Domains

1. Time Domain
    1. **Zero Crossing Rate (ZCR)** - Detecting noisy, percussive, or sharp sounds.
    2. **Root Mean Square Energy (RMS Energy)** - Measuring loudness or signal presence.



2. Frequency Domain
    1. **Spectral Centroid** - Identifying brightness or pitch characteristics.
    2. **Spectral Bandwith** - Distinguishing tonal vs. complex sounds.
    3. **Spectral Rolloff** - Finding energy cutoff points in the spectrum.
    4. **Spectral Contrast** - Analyzing dynamic variations in calls.



These features were extracted using the Librosa Python Library

FEATURE EXTRACTION

The MFCCs is designed to mimic human auditory perception

3. Time and Frequency Domain

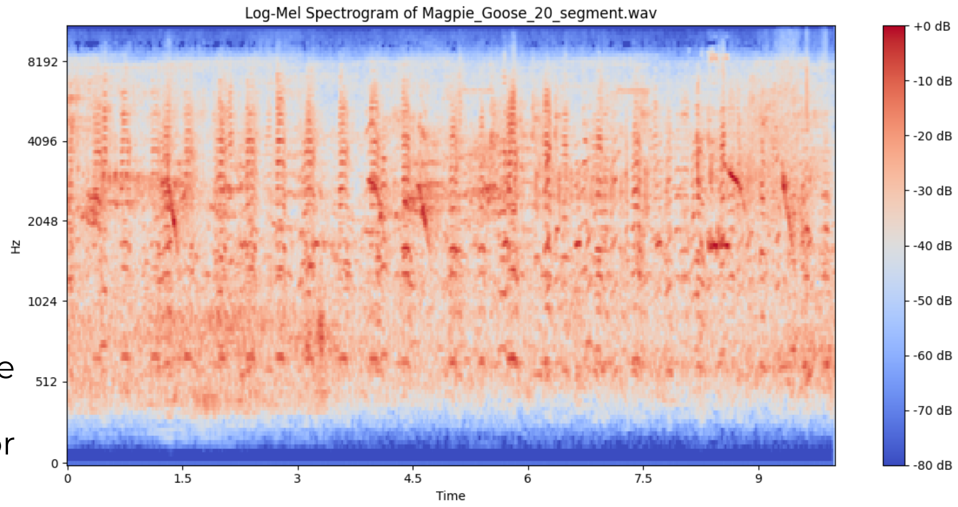
MFCCs are a **compact representation** of the spectral information, designed to **mimic human auditory perception** by mapping frequencies to the Mel scale and compressing the data. It captures characteristics like tone, pitch and timbre.

How does MFCCs work?

1. Break waveform into small chunks of time
2. Focuses on low frequencies more closely (using Mel Scale)
3. Reduces noise and redundancy

What do MFCCs Values Represent?

1. MFCC1: Overall loudness or energy of the sound
2. MFCCs 2-5: broad characteristics like bass vs treble balance
3. MFCCs 6-13: captures finer details like sharpness or variations of tone



FEATURE EXTRACTION

Custom Feature Extraction – Basic "Lima" Spectrogram

3. What are Lima Spectrograms?

Lima Spectrograms are **visual representations of sound signals**, showing how their **frequency content changes over time**. These are particularly useful for analyzing animal calls and other natural sounds. These manually implemented spectrograms are basic.

Key Methodology

1. Extracts features in the frequency domain using Short-Time Fourier Transform (STFT).
2. Divides the frequency spectrum into customizable bins for detailed analysis.
3. Allows for time-frequency relationships to be observed clearly.

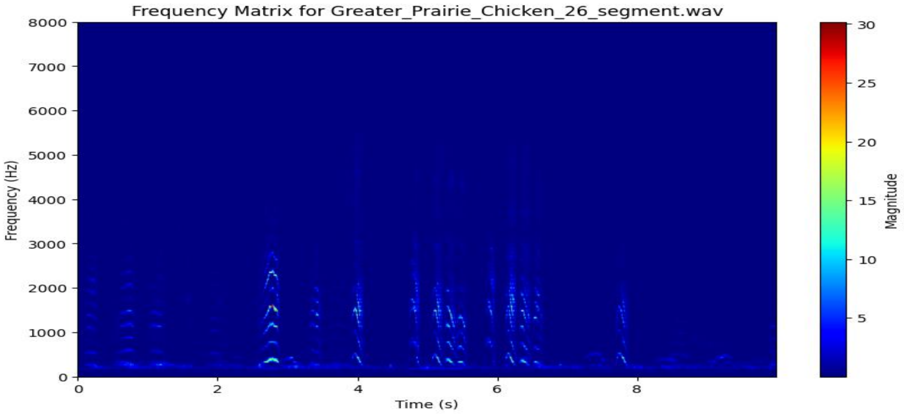
Parameters used:

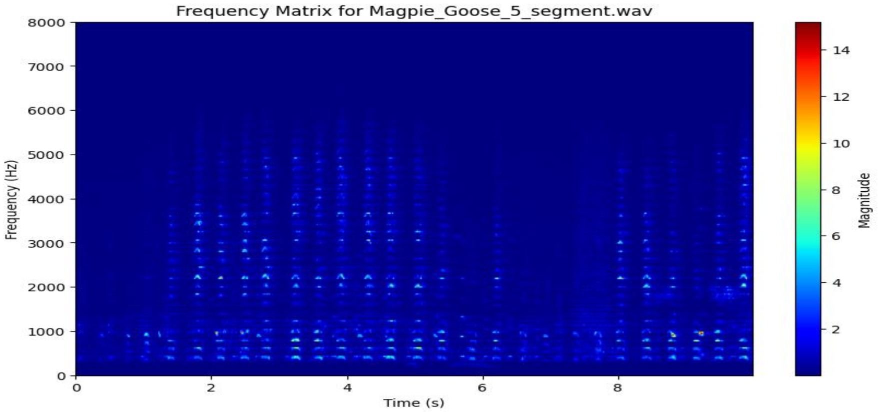
- Maximum Frequency: 8 kHz
- Number of frequency bins: 250
- Hop size: 512 samples
- Window size: 30 milliseconds

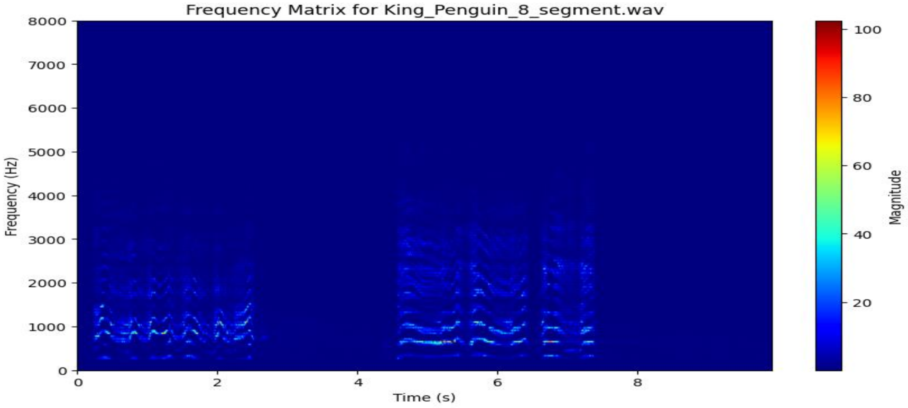
Applications: These basic spectrograms can help differentiate between species-specific vocalizations, identify noise patterns, and analyze the structure of calls in various contexts.

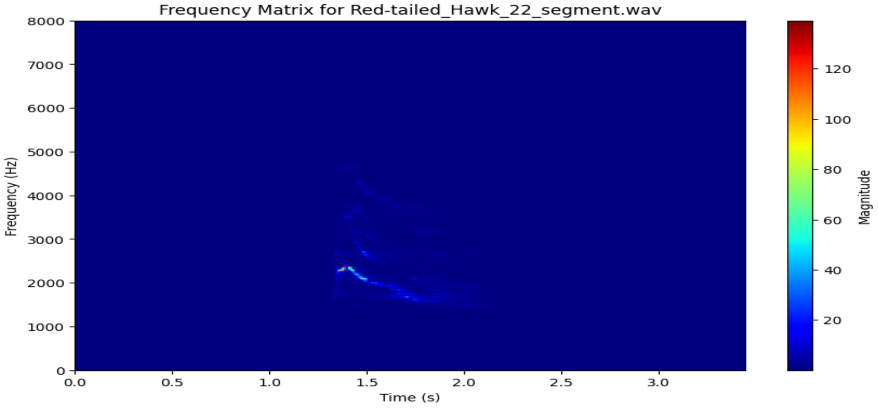
FEATURE EXTRACTION

Example Basic Spectrograms



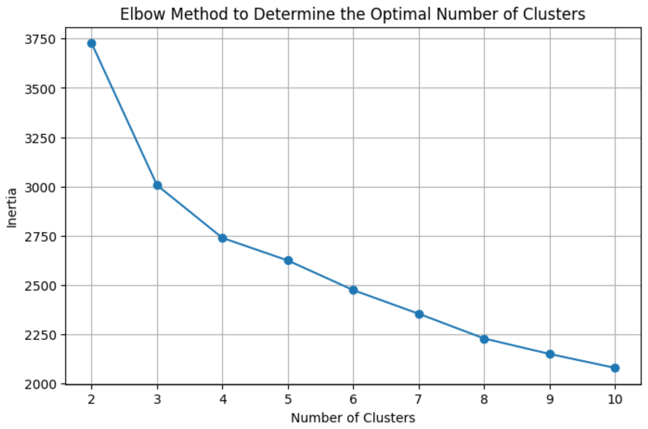






K-MEANS

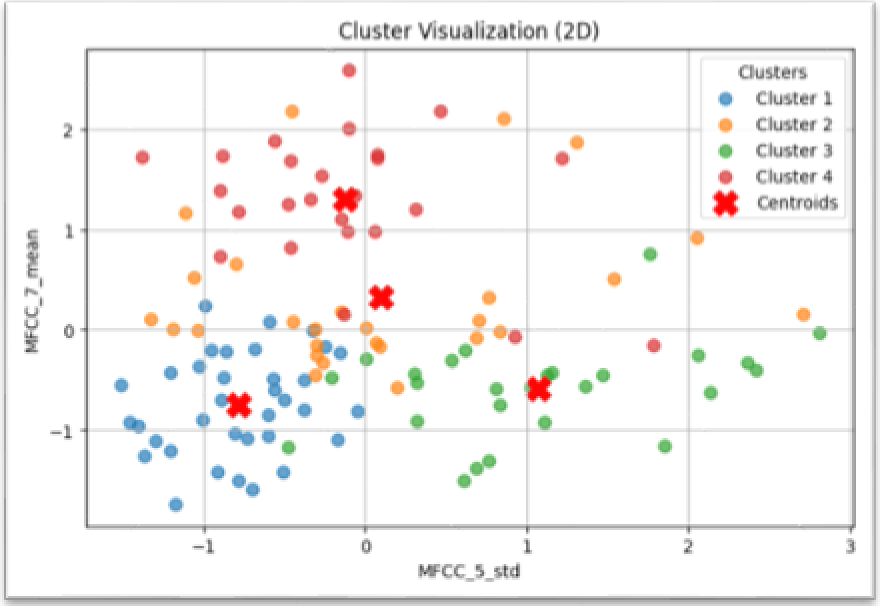
The application of K-means shows us good results for 4 clusters

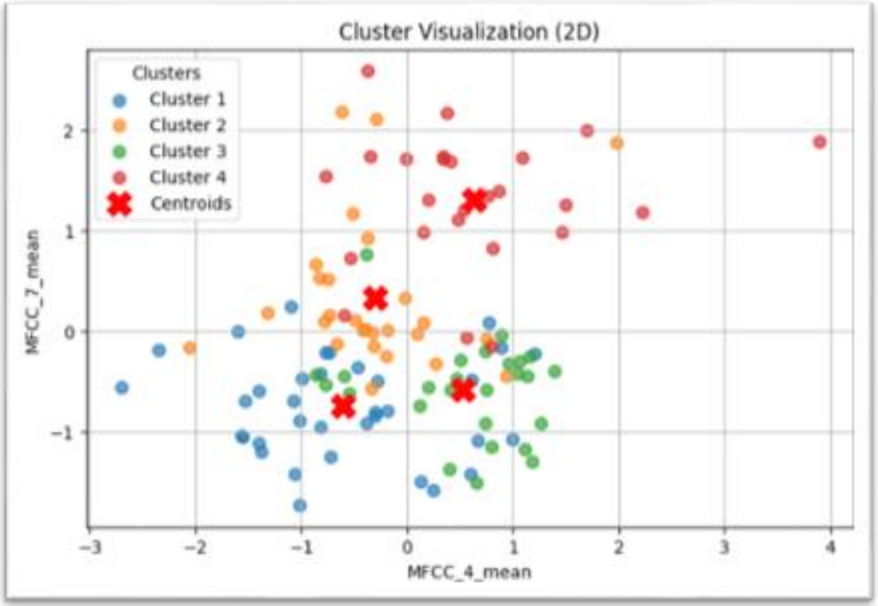


Bird\Cluster	1	2	3	4
Magpie Goose	**22**	3	0	2
King Penguin	6	**20**	4	0
Greater Pairie Chicken	4	2	**22**	0
Red-tailed Hawk	3	2	0	**23**
Accuracy	62.86%	74.07%	84.62%	92.00%

K-MEANS

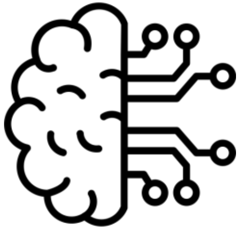
High MFCCs and Spectral Rollof can help to identify clusters in a 2D plot





CONCLUSION

Defining the correct features to use in Machine Learning has a great impact in the capabilities of the model



1. Integration of Supervised Learning

2. Dynamic Feature Expansion

3. Ecological Insights

Thank You