Thar bird call Analysis

By -

Saksham Jain
B21EE059

Under
Dr Mitali Mukerji
Dr Manasi Mukherjee

Objective

- To analyse whether the calls of the same species is same are different in desert region as compared to non desert region.
- We are analysing for Indian Region only so the desert region is Thar and rest of the region is non desert region .
- The differences in calls can be used to detect the onset of desertification in a certain region and act as an measure of environment health

Data preparation

- We are using xeno canto dataset which is a large and diverse collection of bird vocalizations from around the world.
- The dataset has columns of latitude and longitude which is used to determine the region .
- For extracting features from audio and denoising librosa library is used.
- Spectrograms are used as they capture both temporal and spatial information of sound signals.

Comparing Spectrograms

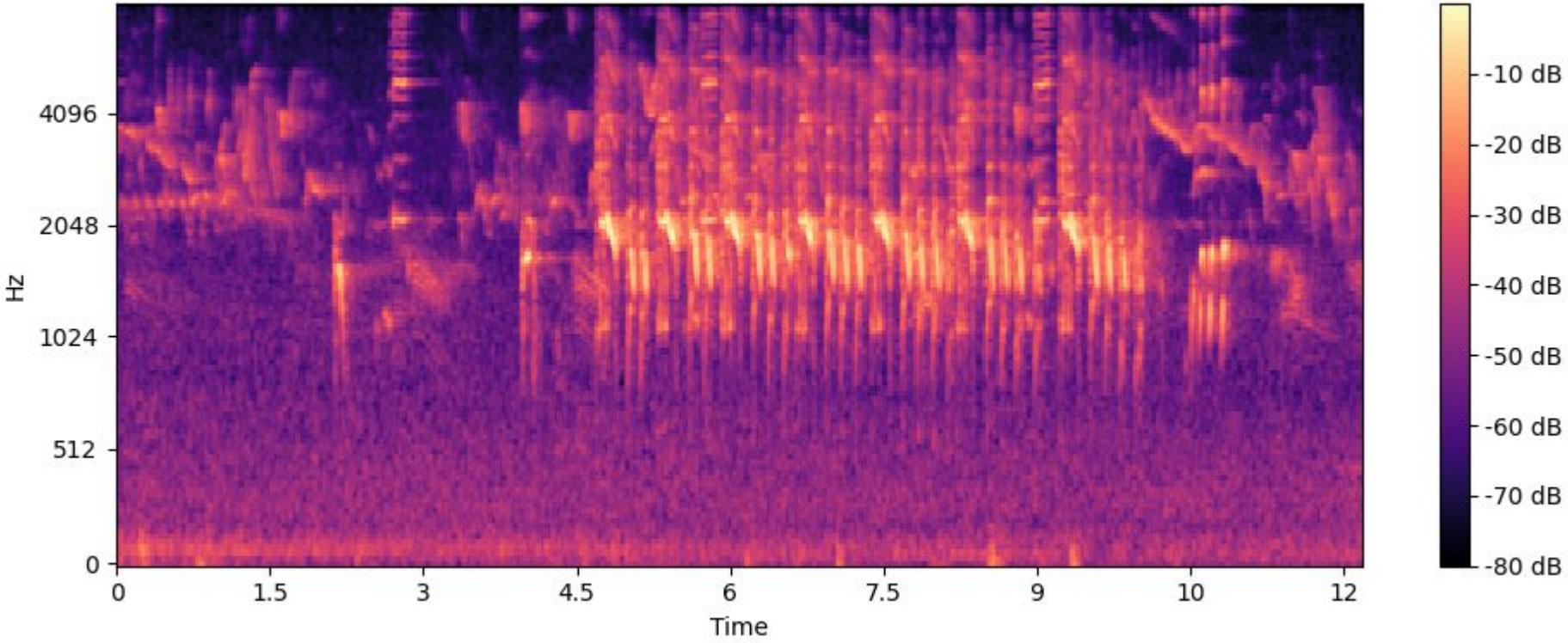
- We needed to compare spectrograms of audio recordings, but we faced a challenge due to the different durations of the recordings, which made direct comparison impossible
- To address this issue, we tried using zero-padding, a technique that involves adding zeros to the spectrograms to equalize their dimensions.
- We then used cosine similarity as our similarity metric to compare the calls. However, we observed that the results were the same for both within-region and across-region comparisons
- One possible explanation for this outcome could be that zero-padding diluted the similarity measure, leading to less discriminative results

- We attempted to extract only the energy-frequency data from the spectrogram and focused on comparing a specific frequency band centered around the peaks.

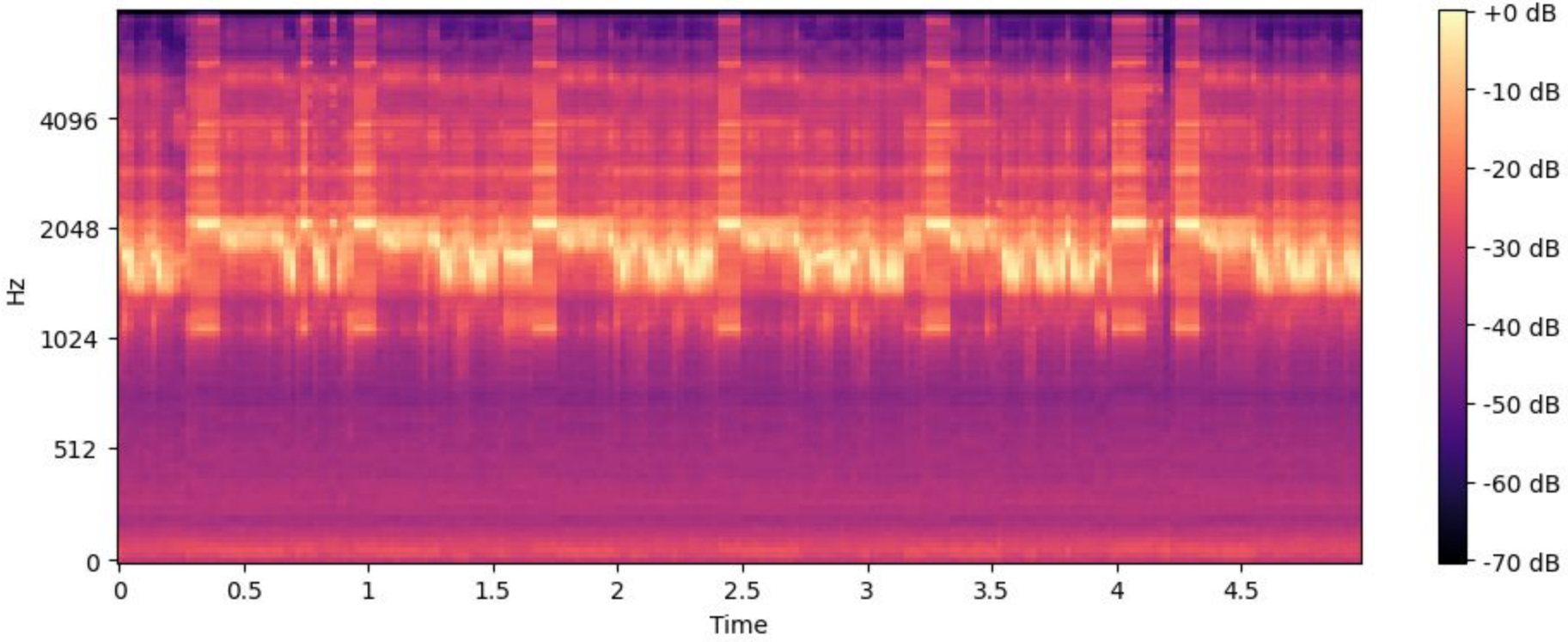
- However, in this approach, the similarity scores varied significantly even for comparisons within the same region. This suggests that this approach is not fit for our purpose.

- To fix the issue with the shape of the spectrogram , Now we are extracting 5 seconds of maximum energy and using that for comparison purposes .

Original Spectrogram



After denoising and extracting 5 seconds of maximum energy

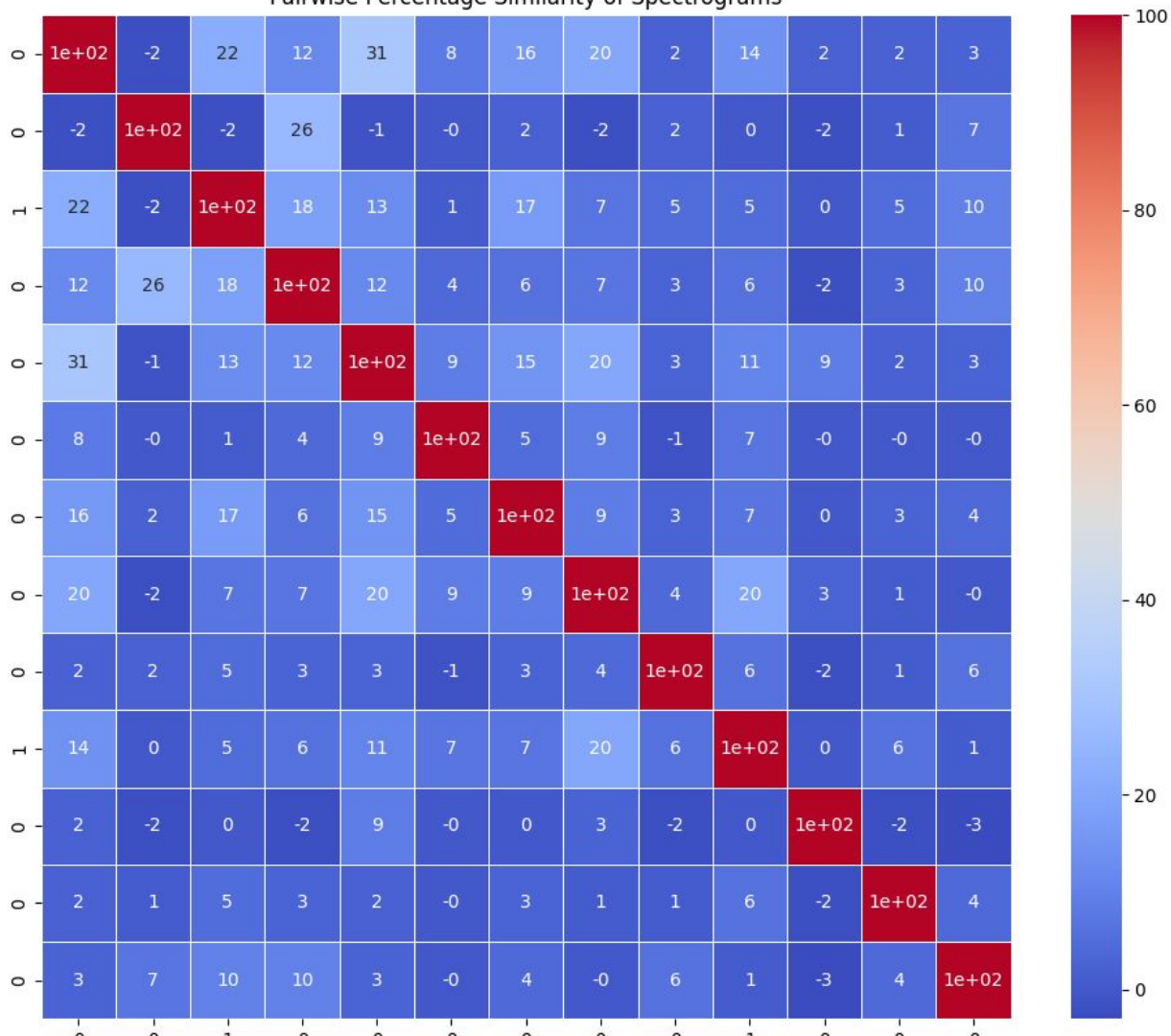


So our denoising function and extraction of 5second of maximum energy works

Pearson Correlation

Pearson correlation is a statistical measure that quantifies the linear relationship between two variables. It can be used to assess the similarity between two spectrograms by calculating the correlation between corresponding elements in the spectrograms. The Pearson correlation coefficient ranges from -1 to 1, where 1 indicates a perfect positive linear relationship, -1 indicates a perfect negative linear relationship, and 0 indicates no linear relationship.

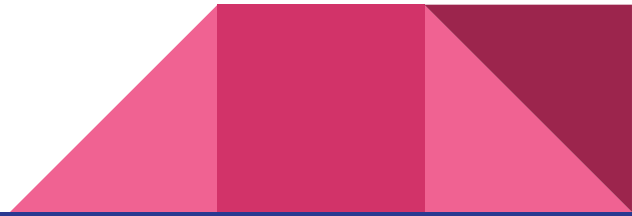
We find that pearson correlation is not able to differentiate between desert and non desert sounds.

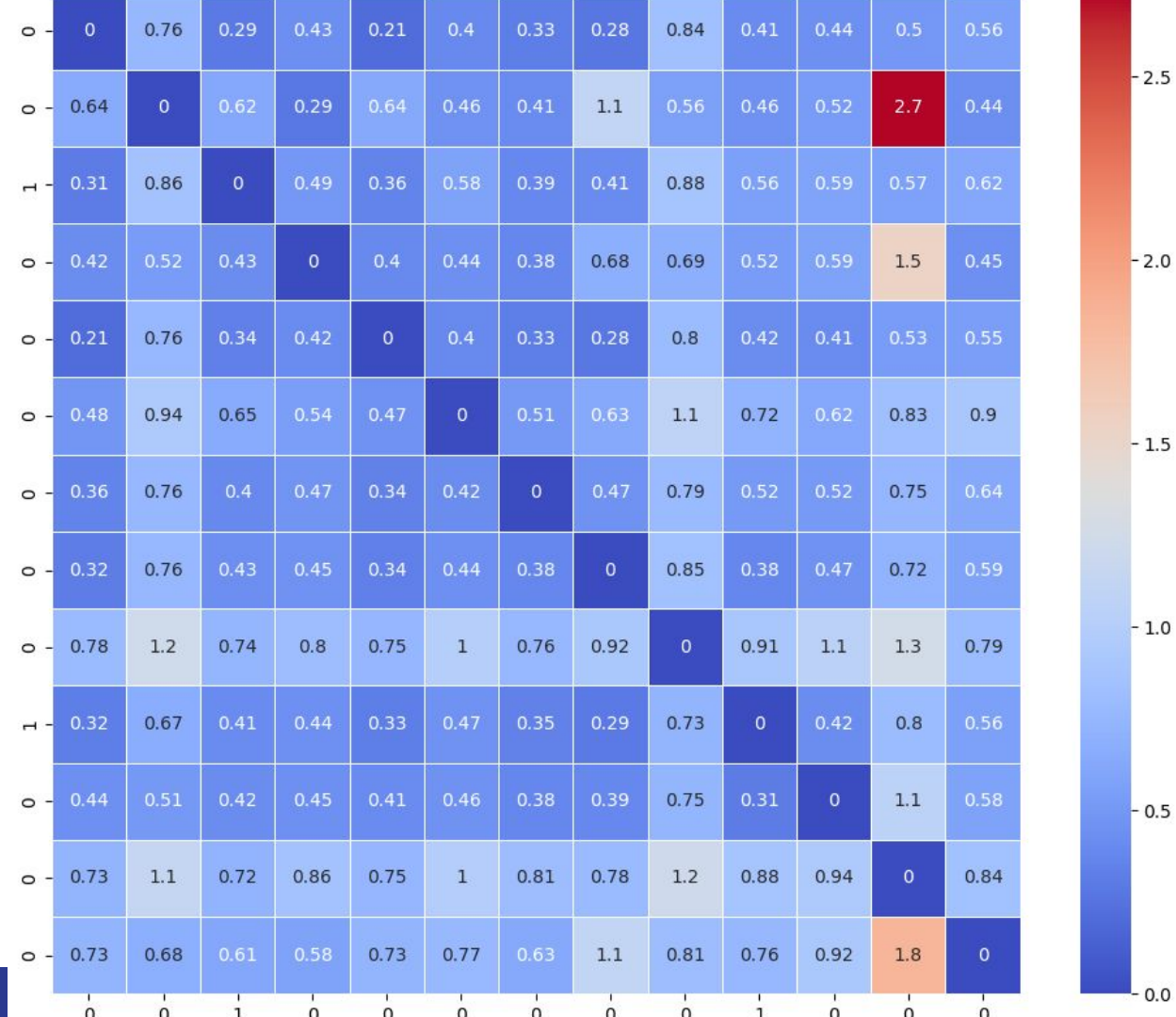

Pairwise Percentage Similarity of Spectrograms
0 1e+02 -2 22 12 31 8 16 20 2 14 2 2 3
0 -2 1e+02 -2 26 -1 -0 2 -2 2 0 -2 1 7
1 22 -2 1e+02 18 13 1 17 7 5 5 0 5 10
0 12 26 18 1e+02 12 4 6 7 3 6 -2 3 10
0 31 -1 13 12 1e+02 9 15 20 3 11 9 2 3
0 8 -0 1 4 9 1e+02 5 9 -1 7 -0 -0 -0
0 16 2 17 6 15 5 1e+02 9 3 7 0 3 4
0 20 -2 7 7 20 9 9 1e+02 4 20 3 1 -0
0 2 2 5 3 3 -1 3 4 1e+02 6 -2 1 6
1 14 0 5 6 11 7 7 20 6 1e+02 0 6 1
0 2 -2 0 -2 9 -0 0 3 -2 0 1e+02 -2 -3
0 2 1 5 3 2 -0 3 1 1 6 -2 1e+02 4
0 3 7 10 10 3 -0 4 -0 6 1 -3 4 1e+02
100
80
60
40
20
0

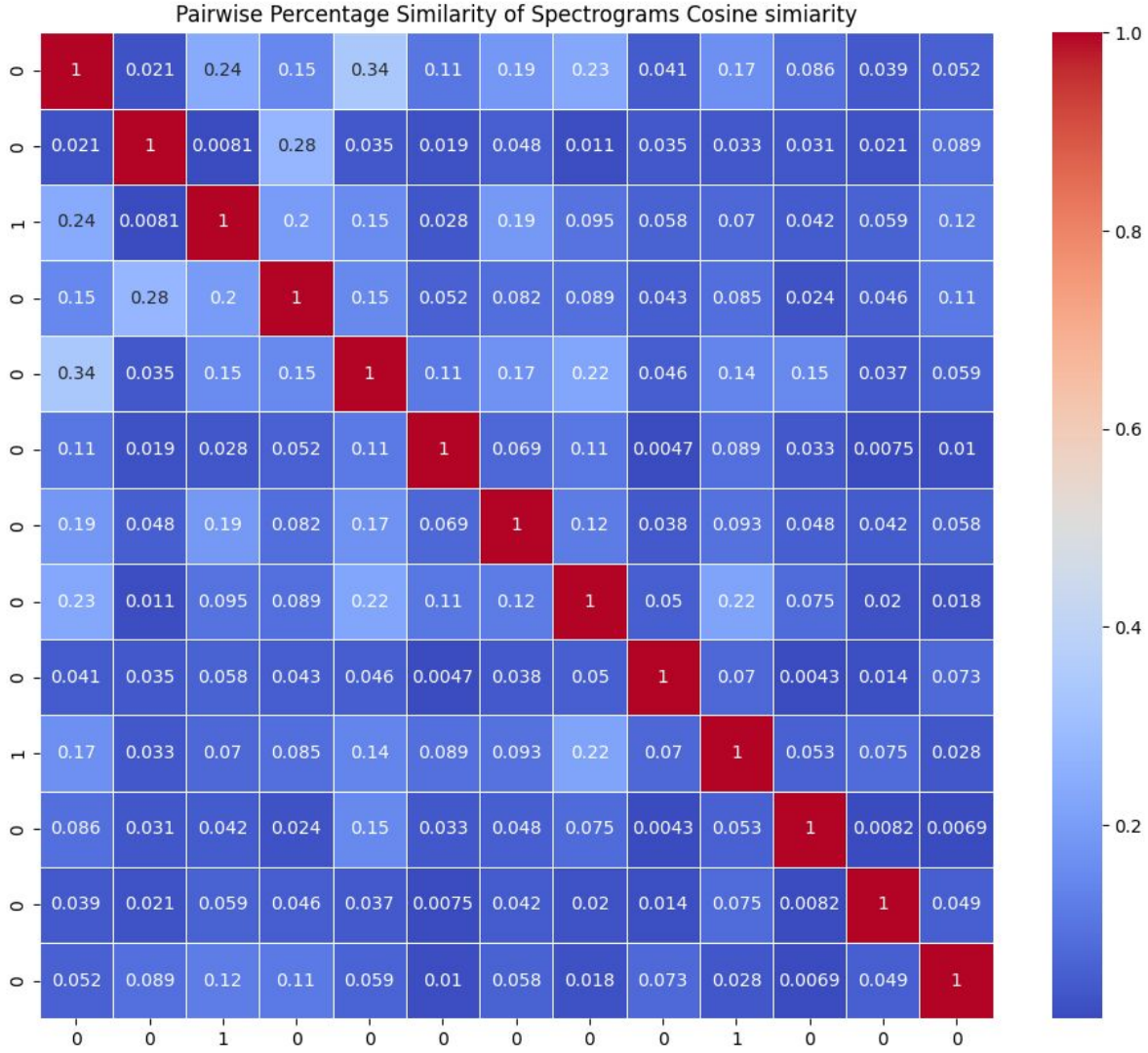
KL divergence

KL divergence is a non-symmetric metric that measures the relative entropy or difference in information represented by two distributions.

Even KL divergence is not able to differentiate between desert and non desert sound as is evident from the results .



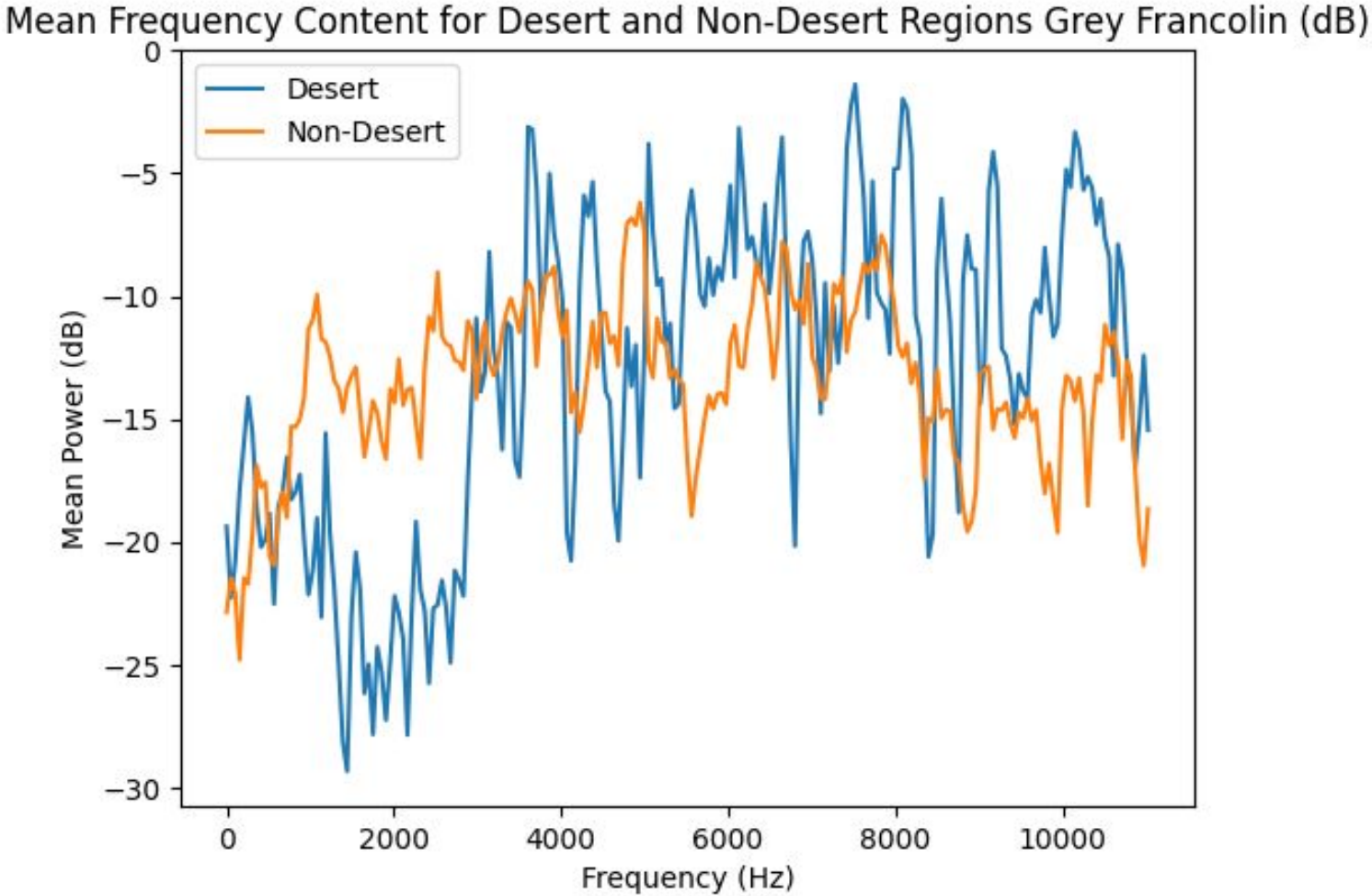

0 0.76 0.29 0.43 0.21 0.4 0.33 0.28 0.84 0.41 0.44 0.5 0.56
0.64 0 0.62 0.29 0.64 0.46 0.41 1.1 0.56 0.46 0.52 2.7 0.44
0.31 0.86 0 0.49 0.36 0.58 0.39 0.41 0.88 0.56 0.59 0.57 0.62
0.42 0.52 0.43 0 0.4 0.44 0.38 0.68 0.69 0.52 0.59 1.5 0.45
0.21 0.76 0.34 0.42 0 0.4 0.33 0.28 0.8 0.42 0.41 0.53 0.55
0.48 0.94 0.65 0.54 0.47 0 0.51 0.63 1.1 0.72 0.62 0.83 0.9
0.36 0.76 0.4 0.47 0.34 0.42 0 0.47 0.79 0.52 0.52 0.75 0.64
0.32 0.76 0.43 0.45 0.34 0.44 0.38 0 0.85 0.38 0.47 0.72 0.59
0.78 1.2 0.74 0.8 0.75 1 0.76 0.92 0 0.91 1.1 1.3 0.79
0.32 0.67 0.41 0.44 0.33 0.47 0.35 0.29 0.73 0 0.42 0.8 0.56
0.44 0.51 0.42 0.45 0.41 0.46 0.38 0.39 0.75 0.31 0 1.1 0.58
0.73 1.1 0.72 0.86 0.75 1 0.81 0.78 1.2 0.88 0.94 0 0.84
0.73 0.68 0.61 0.58 0.73 0.77 0.63 1.1 0.81 0.76 0.92 1.8 0
2.5
2.0
1.5
1.0
0.5
0.0


Pairwise Percentage Similarity of Spectrograms Cosine simiarity
0 1 0.021 0.24 0.15 0.34 0.11 0.19 0.23 0.041 0.17 0.086 0.039 0.052
0 0.021 1 0.0081 0.28 0.035 0.019 0.048 0.011 0.035 0.033 0.031 0.021 0.089
1 0.24 0.0081 1 0.2 0.15 0.028 0.19 0.095 0.058 0.07 0.042 0.059 0.12
0 0.15 0.28 0.2 1 0.15 0.052 0.082 0.089 0.043 0.085 0.024 0.046 0.11
0 0.34 0.035 0.15 0.15 1 0.11 0.17 0.22 0.046 0.14 0.15 0.037 0.059
0 0.11 0.019 0.028 0.052 0.11 1 0.069 0.11 0.0047 0.089 0.033 0.0075 0.01
0 0.19 0.048 0.19 0.082 0.17 0.069 1 0.12 0.038 0.093 0.048 0.042 0.058
0 0.23 0.011 0.095 0.089 0.22 0.11 0.12 1 0.05 0.22 0.075 0.02 0.018
0 0.041 0.035 0.058 0.043 0.046 0.0047 0.038 0.05 1 0.07 0.0043 0.014 0.073
1 0.17 0.033 0.07 0.085 0.14 0.089 0.093 0.22 0.07 1 0.053 0.075 0.028
0 0.086 0.031 0.042 0.024 0.15 0.033 0.048 0.075 0.0043 0.053 1 0.0082 0.0069
0 0.039 0.021 0.059 0.046 0.037 0.0075 0.042 0.02 0.014 0.075 0.0082 1 0.049
0 0.052 0.089 0.12 0.11 0.059 0.01 0.058 0.018 0.073 0.028 0.0069 0.049 1
0 0 1 0 0 0 0 0 0 1 0 0 0
1.0
0.8
0.6
0.4
0.2

T - tests and Mean frequency content

The t-test is used in this code to statistically compare the mean frequency content of vocalizations between two groups: desert regions and non-desert regions. Specifically, the code performs a two-sample t-test, which is a hypothesis test used to determine whether there is a significant difference between the means of two independent groups.

We calculate the mean frequency content of species' vocalizations to assess the distribution of energy across different frequency bands. This is achieved by computing the power spectral density of the mel spectrograms and averaging the power values within each frequency band.


Mean Frequency Content for Desert and Non-Desert Regions Grey Francolin (dB)
Desert
Non-Desert
Mean Power (dB)
0
−5
−10
−15
−20
−25
−30
0
2000
4000
6000
8000
10000
Frequency (Hz)

Mean Frequency Content for Desert and Non-Desert Regions Purple Sunbird (dB)