<u>Sounds of Nature: Integrative Monitoring of Biodiversity Through Acoustic Landscapes and Animal Sound Recognition</u>

Amar Meddahi
ENSEEIHT

Abstract
This project aims to develop methods for analyzing massive acoustic data in natural environments using supervised and unsupervised deep learning techniques. We successfully recognized acoustic landscapes and estimated environmental indicators, including forest cover, altitude, and shape, using machine learning algorithms. Additionally, we developed an automatic recognition system for over 60 bird species, achieving a satisfactory accuracy of 75%. The project has potential applications in ecological research and conservation efforts and provided valuable learning experiences for us.

Introduction
Recent developments in passive autonomous acoustic recorder technology have led to a paradigm shift in the study of biodiversity. These devices, capable of recording soundscapes in natural environments for extended periods, offer a wealth of data for ecologists to better understand animal communication and biodiversity decline. However, the massive amount of data generated by these devices poses a challenge to ecologists. The aim of this project is to develop a set of methods for analyzing massive acoustic data in natural environments, focusing on supervised and unsupervised deep learning techniques for characterizing acoustic landscapes and automatically recognizing species or groups of animal species.

Acoustic Landscapes
This part of the project aims to develop techniques for automatically recognizing acoustic landscapes by benchmarking various machine learning methods, as well as accurately estimating biodiversity indicators.
Acoustic landscapes are characterized by two labels: "blue" or "red" and "estive" or "edge." "Blue" refers to forests that have descended in altitude since the 1950s, while "red" refers to forests that have moved up in altitude due to climate change. The labels "estive" and "edge" refer to areas that are at least 200m above the tree line and the forest edge, respectively.
Various machine learning algorithms were rigorously benchmarked using cross-validation to identify the most effective approach for automated recognition of acoustic landscapes. The performance of each algorithm was evaluated based on the specific audio representation chosen. The results showed that the Artificial Neural Network (ANN) model achieved the highest accuracy of 92% when trained on acoustic indices. When trained on the latent representation, the ANN model still performed well with an accuracy of 89%.
In addition to the primary objective of recognizing acoustic landscapes, this project aimed to estimate a range of environmental indicators using the same sound data. These included measures of dominant species, lithological diversity, proportion of quaternary deposits and crystalline rocks, average altitude and slope, orientation, precipitation and temperature evolution, as well as distances to roads, trails, ski runs, buildings, and a national nature reserve.

Additionally, we sought to quantify changes in forest cover, altitude, and shape. To estimate these parameters, we utilized two different representations of the sound data: acoustic indices and a latent representation generated using a pre-trained neural nets encoder. We employed a variety of machine learning algorithms, including Linear, Ridge, Lasso, ElasticNet, Decision tree, Random forest, and ANN, to accurately estimate these parameters. Our results demonstrated that the random forest model trained on the index representation achieved the best performance. For example, we achieved an average altitude estimate with an accuracy of approximately 50m, an orientation estimate with a precision of around 2°, and a distance estimation to the nearest trail with an accuracy of approximately 2m. This successful estimation of environmental indicators from acoustic data opens up new possibilities for environmental monitoring and conservation efforts.

Bird sounds

The second part of the project focuses on the use of a bank of animal songs, calls, and stridulations from soundscapes, annotated in time and frequency, to develop an automatic recognition system for over 60 different species that produce sounds. We represented the audio sequences in a spectrogram, a grid-like image representation that allows us to apply convolutional neural networks (CNNs) for classification. We trained and compared different models to achieve the best performance. Our most successful model achieved an accuracy of 75%, which is a very satisfactory result considering the large number of species involved. This system has potential applications in ecological research, conservation, and monitoring of bird populations.

Feedback Reflexion

This project has been a valuable learning experience for me, allowing me to gain expertise in acoustic landscapes and machine learning algorithms for sound data. Our team's efforts have yielded promising results, and I am excited about the potential applications of these techniques in ecological research and conservation. Overall, I feel proud of our achievements and grateful for the skills and knowledge gained through this project.

Conclusion

The project developed methods for analyzing massive acoustic data in natural environments using supervised and unsupervised deep learning techniques. We successfully recognized acoustic landscapes and estimated environmental indicators, including forest cover, altitude, and shape, using machine learning algorithms. Additionally, we developed an automatic recognition system for over 60 bird species, achieving a satisfactory accuracy of 75%. The project provided valuable learning experiences and has potential applications in ecological research and conservation efforts.

Acknowledgement

We thank Axel Carlier for his expertise in deep learning and data science, and Maxime Cauchoix, David Funosas, Elodie Massol, Arnaud Elger, and Luc Barabro for their guidance in understanding the project context. Our gratitude also extends to everyone else who has contributed to this project.