Icelandic Marine Audio Classification

Ella Linder
*eklinder22@earlham.edu*

Earlham
C O L L E G E

Intro

Earlham's Icelandic Field Studies (IFS) collect multiple types of data from Skalanes, Iceland every summer. One of those types is audio samples taken in the fjord next to Skalanes. This helps determine how humans may influence animal traffic there. Currently, the method of collection is attaching various Go Pros scattered across the water for 1–2 hours, a person then parses through this footage manually. There are two problems with this current setup: there is limited data, and it is laborious to parse. Because of this, I decided to make a project solving this issue. If IFS collects audio through a sonogram, they will be able to use my program to automatically parse through and organize marine life sounds.
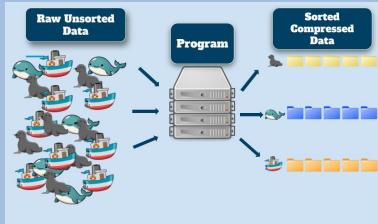
Design

This project uses autoencoders in order to distinguish whale audio from miscellaneous audio. It mainly uses python and bash in the server itself, with no graphical UI. The spectrograms use Sox in order to convert WAV files to graphical spectrograms.

Methods
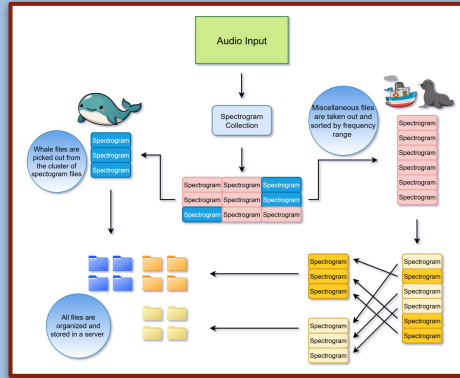
The program begins by taking in audio input and converting the files into spectrograms, in order for the image classification model to distinguish whale sounds from other sources. Those files are then stored in a folder on the Earlham server, while the files that have miscellaneous sounds are sorted by highest frequency, then stored separately on the server.

Results

I hypothesized the frequency sorting program to operate at around 80% accuracy, and the whale classifier to run at around 90% accuracy, but the classifier was not able to run successfully, and the frequency program demonstrated that it would only be able to sort outliers, as most frequency ranges of the selected sources were very similar. HOWEVER, with a larger sample size, it would be more effective. The whale, for instance, had a very large sample size and was able to sort it correctly with 74% accuracy, and an added 11% with it narrowing down to two options. The rest performed around 10% accuracy, however.

Future Work

Graphical interface, classification to sounds other than whales, perfect the frequency sorting. Creating a graphical interface would be a great addition to the project, allowing ease of use for IFS and increasing speed and accessibility. Adding a larger sample size for the miscellaneous sources would greatly improve the outcome.