Whale Wave Analyzer

Eliza Homer

Level of Interest: 10/10

Description

Whales, dolphins, porpoises, and other cetaceans are renowned for their social nature, often forming tight-knit groups. To communicate with pod members, they make different clicks, pops, and whistles that create the language of whale song. Whales are known to be gentle giants of the oceans; however, there has been a sudden increase in orca attacks on boats. To gain deeper insights into the world of cetaceans and foster safer coexistence with these majestic mammals, it would be helpful to be able to automatically identify species using only their vocalization. Such an endeavor would aid scientific research as well as offer the possibility of preemptive measures to avoid encounters with pods of orcas known to exhibit aggression. For this project, I propose using machine learning models to predict the different species of cetacean based on a small audio clip.

What would the data look like?

A metadata file would include a list of audio files associated with different species of cetacean.

Audio file	Location	Observation Date	Scientific Name	Common Name
55113002.wav	St. David's Island, Bermuda	30-Apr-1953	Megaptera novaeangliae	Humpback Whale

Using the Mel Frequency Cepstral Coefficient (MFCC) algorithm, we can break our sound file into different numerical features that represent the qualities of the audio file. I have only included six features in this example data instance, though we could easily include many more features for a single audio file.

Feature 1: Loudness	Feature 2: Perceived Pitch	Feature 3: Timbre	Feature 4: Texture	Feature 5: Dynamics	Feature 6: Energy	Class
-273.6928	182.60896	-79.25755	32.81927	32.62913	9.885547	1

Gathering the Data

The data would be gathered from the Watkins Marine Mammal Sound Database[1] website. There are 32 different species with about 15,000 sound cuts[2] of which 1,694 files are the 'best cuts. Each of the clips is about 2 - 15 seconds long, with some clips being longer. They are in the waveform audio file format and are free to download for academic use. We would download each of the sound clips along with the name of the associated species. Then we would use the librosa[3] library, a python package for music and audio analysis, to load and process the files. To extract the features from the audio files, we can use the Mel Frequency Cepstral Coefficient[4] (MFCC) algorithm. It has been used since the 1980s[5] to turn frequency and time characteristics into numerical features. We can choose how many features we would like our dataset to include.

---

[1] whoicf2.whoi.edu

[2] "Watkins Marine Mammal Sound Database, Woods Hole Oceanographic Institution and the New Bedford Whaling Museum."

[3] Mandapaka, Karthik. "Handling Audio Data for Machine Learning." *Medium*, MLearning.ai, 23 Oct. 2021, medium.com/mlearning-ai/handling-audio-data-for-machine-learning-7ba225f183cb.

[4] "Mel Frequency Cepstral Coefficient (MFCC) Tutorial." *Practical Cryptography*, practicalcryptography.com/miscellaneous/machine-learning/guide-mel-frequency-cepstral-coefficients-mfccs/. Accessed 29 Sept. 2023.

[5] Ogola, Willies. "Machine Learning for Audio Classification." *Section*, 29 Sept. 2021, www.section.io/engineering-education/machine-learning-for-audio-classification/. Accessed 29 Sept. 2023.