Identification of Avian Chirps and Songs of Forest Bird Species on Kauai using Machine Learning

Team Members: Ivy Wang (wangivy@stanford.edu)

Motivation

The goal of this project is to aid the identification of endangered avian species using machine learning. Even after 70 years of human point counts, basic demographics facts such as the population size of endangered avian species on Kauai is still unknown. Inaccessibility and cost considerations such as needing to fly into remote habitats make accurate numbers from human observations uneconomical. Without accurate numbers, conservationists do not know if conservation efforts are effective. Remote sensing and remote listening stations provide a fundamentally more economical approach to getting this data. However, there are no techniques developed to efficiently distinguish avian sounds.

For this project, we will take a small first step which is to separate out chirps and calls from field recordings against the background noise of the forest. Most existing research has been focused on bird songs, but since many of these critically endangered species are cryptic, these methods of programs have limited applicability. Isolating the chirps and calls from the vast amount of field recordings is an essential first step for efficiently collecting enough data to train more complex classification problems.

Method

We will be using field recordings from the Kauai Forest Bird Recovery Project database to train our machine learning methods. We will focus on supervised learning algorithms such as neural networks and logistic regression to identify which time segments in a field recording contain avian chirps and calls.

The general method involves segmenting each clip into short timeframes, and applying FFT spectrum analysis on each timeframe. We will then train a neural network to classify which timeframes contain chirps and calls and which do not. Logistic regression can also be used by first identifying useful features of each sound segment such as maximum frequency, mean amplitude, and finding the most useful features to use for classifying timeframes.

Intended Experiments

The effectiveness of the algorithms will be measured by the percentage of timeframes accurately classified in our test data. We will use this metric to compare the usefulness of various features chosen for logistic regression, as well as the parameters chosen for an FFT spectrum analyzer. We may also examine how our algorithms are affected by applying different audio pre-processing methods such as a band pass filter.