CS486C – Senior Capstone Design in Computer Science
Project Description

Project Title: Machine learning classification of acoustic data components	
Sponsor Information:	**Colin Quinn**, PhD student NAU SICCS: Global Earth Observation and Dynamics of Ecosystems Lab cq73@nau.edu Patrick Burns, Research Associate SICCS: Global Earth Observation and Dynamics of Ecosystems Lab pb463@nau.edu

Background

Natural and anthropogenic environmental changes affect a range of species around the world and as more species are being affected, proper management and observation are required to understand their response. Traditional field methods require trained observers to determine species presence/absence and are thus expensive and logistically challenging to employ at large scales. The recent evolution of passive acoustic monitoring (PAM), the practice of deploying acoustic recorders to record the in situ variability of sound (AKA the local “soundscape”), has allowed for ecologists and land managers to have a long-lasting record of the audible spectrum of a landscape. Using PAM coupled with expert listener verifications and machine learning, we can efficiently identify many species’ calls in large acoustic datasets. This method is especially practical for vocal species like birds, which are sometimes difficult to detect visually.

Using sound to monitor biodiversity across landscapes is a fairly recent development (Farina et al. 2011). The availability of inexpensive, autonomous acoustic recorders that detect sounds with sufficient quality has made possible the scaling of research from single locations and organisms to full animal communities across landscapes (Sueur et al. 2014; Krause and Farina 2016). Acoustic recorders facilitate the survey of large areas for long periods of time, and in doing so generate terabytes of recording data.

New developments in our capacity to rapidly process sound data with high performance computing systems and machine-learning algorithms now permit the detection of species at a site with reasonable accuracy (Stowell et al. 2018; Priyadarshan et al. 2018). The NASA-funded project **Soundscapes to Landscapes** uses inexpensive AudioMoth devices for recording soundscapes across Sonoma County, California, including natural forests, urban areas, and agricultural fields. These devices are programmed to record 1 minute every 10 minutes for 3 to 5 days at each site and have thus far resulted in over 240 days of stored acoustic data.

The Problem

Since recorders can be located in dense urban areas or near roads, there is often a range of anthropogenic noise in the recordings. Similarly, recorders are deployed in the breeding season for birds, which runs from March to June, and this spring weather period can often bring wind and rain that negatively impacts recordings. Depending on the habitat setting, effects from background noise can be highly influential on the quality of sound recordings, potentially impacting our ability to determine whether or not a species is present.

There are, of course, various ways to “clean up” a soundscape recording to filter out unwanted sounds, including the use of AI algorithms to recognize and remove unwanted sounds. The problem is that these tools are fairly arduous to use, particularly for non-experts in technology. Using a purely functional basic command line interface, each sound file must be loaded, appropriate parameters must be set, and the filtering algorithms applied…and the process potentially repeated with different algorithms or parameters if the result was not satisfactory. When multiplied by the sheer volume of soundscape data that a site produces and the potentially large number of sites in a study, this "data scrubbing" preparatory process can easily dominate the entire analysis workflow.

Envisioned Solution

What is needed is a streamlined software tool that will allow researchers to (1) accurately identify and remove non-biologic audio components, and (2) better understand how various sources of noise in soundscape recordings diminish our ability to detect individual bird species and quantify avian diversity. We will call this product the Soundscape Noise Analysis Workbench (SNAW). In particular, SNAW will be implemented as a user-friendly web application running on a cloud-based server, and with direct access to high-performance computing power, e.g., via a backend compute instance also hosted on the cloud site. This will allow researchers anywhere, anytime, to upload their soundscapes for noise analysis, then snapshot and/or download results. Key functions to support include the following. Tier 1 features represent the minimum viable product to be useful at all; Tier 2 features are more complete solutions that could actually be usable by non-geeks; Tier 3 represents stretch goals.

1. Web application capable of running on all modern browsers that provides a clean, highly-usable analysis interface for researchers to upload a novel audio file and receive aforementioned products. Solution includes the specialized cloud-based architecture developed to host the complete GUI+HPC solution.
2. Ability to load soundscape files in various ways, i.e., individual file selection using a file browser from a local machine, and ways to batch load files to the cloud-based analysis server running SNAW.
3. Ability to return results of analysis in a range of useful fashions, including a summary of the components of the sound file, acoustic indices, and export of a sound file with the noise masked out.
4. A machine learning module that can classify (decide) whether or not a given soundscape file contains confounding (non-biologic) acoustic components: human (anthrophony) and/or abiotic natural (geophony). The classifier should be flexible, with the functionality to return either absolute presence/absence or a percentage of time that a particular sound component is present.
   a. * Tier 1 Result: a tabular summary of the percentage of each sound component for each audio file.
   b. ** Tier 2 Result: a visualization (e.g., figure 1) showing masked out regions identified as having anthrophony, geophony, or biophony present. This can also be output as a .wav file with regions masked out or a table with coordinates for the user to mask.

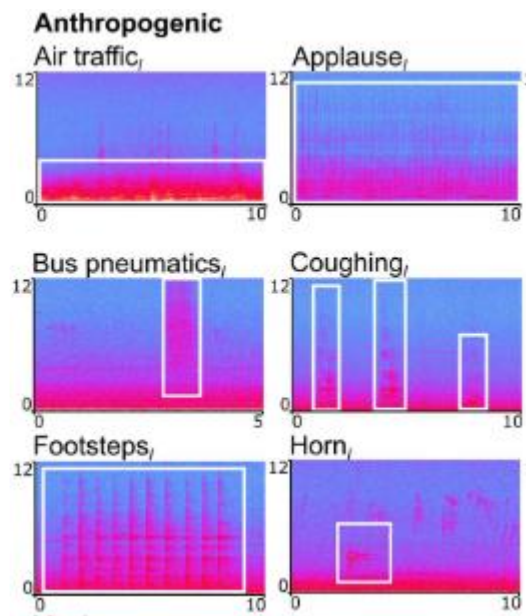

Fig 1. *Examples of human (anthropogenic) sources of sound in 2D spectrograms where the x-axis = time, y-axis = frequency, and z-axis = amplitude (adapted from Fairbrass et al. 2017).*

5. A proper assessment of the accuracy and error associated with the final machine learning algorithm.
   a. ** Tier 1 Result*: For absolute presence/absence, classification accuracy should be assessed using a confusion matrix, and the sensitivity, specificity, and overall accuracy metrics. For percentage

predictions, accuracy should be assessed using R_2 and root mean squared error (RMSE). Each type of model should use at least 50 files (or sound regions) for model validation.

6. Able to produce acoustic diversity metrics to quantify different dimensions of variability in audio files. These metrics condense the broad variability of sounds found in some soundscape files into a single summary metric. A number of R packages and python libraries already exist for performing these calculations on audio files.
   a. * *Tier 1 Result*: Acoustic diversity metrics calculated for each audio file before and after accounting for confounding acoustic components.
   b. *** *Tier 3 Result*: Base the machine learning model on a Raspberry Pi with a hard drive and wireless transmitter (to be provided by Project leads) to use in the field for near real time results.
7. Tier 3: a standalone remote version of the SNAW product. Make the product portable, i.e., allowing it to be loaded onto a laptop or Raspberry PI for transportation to a field site; fresh files could be loaded from recorders with analysis results returned in a near real-time. A quick instance analysis like this could be quite helpful in rapidly pruning away sites that are noisy and/or uninteresting.

Required Deliverables

- The software products outlined about, deployed and demonstrate on a platform of client's choice.
- A strong as-built report detailing the design and implementation of the product in a complete, clear and professional manner. This document should provide a strong basis for future development of the product.
- Complete professionally-documented codebase, delivered both as a repository in GitHub, BitBucket, or some other version control repository; and as a physical archive on a USB drive.

Knowledge, skills, expertise, equipment required for this project:

- A basic understanding of machine learning model building and implementation
- Expertise with appropriate coding languages to implement machine learning module(s)
- Interfaceable web-based app construction and design
- Client will provide access to and expertise on any machine learning tools or algorithms required by the team, as well as the techniques used to apply the tools to clean up soundscape datasets.
- No equipment required other than a development platform and open source software/tools
    - Note: access to Amazon Web Service to both host model and web-app is possible

Broader Application

The creation of a flexible and accurate model for identifying confounding audio components, as well as an associated web-based app, would be a notable advance for NASA Soundscapes to Landscapes and the field of sound ecology. It will allow for researchers to 'clean' acoustic data without having to create models from scratch and may be used by individuals who otherwise are not able to develop such a product themselves. Additionally, we plan to host the web-app on the Soundscapes to Landscapes website (https://soundscapes2landscapes.org/) for public access to process audio files for related research projects and engage the citizen scientist community. If the stretch goal, loading the model on a mobile platform, is successful, this will allow for a novel, near real-time processing and distilling of large amounts of acoustic data; a process that currently requires extensive post-processing effort.

References

Fairbrass, A. J., Rennert, P., Williams, C., Titheridge H., and Jones K. E, (2017). Biases of acoustic indices measuring biodiversity in urban areas. Ecological indicators, 83, pp. 169-177.

Farina, A., Pieretti, N. and Piccioli, L., 2011. The soundscape methodology for long-term bird monitoring: A Mediterranean Europe case-study. Ecological Informatics, 6(6), pp.354-363.
Krause, B. and Farina, A., 2016. Using ecoacoustic methods to survey the impacts of climate change on biodiversity. Biological Conservation, 195, pp.245-254.
Priyadarshani, N., Marsland S., and Castro, I, (2018). Automated birdsong recognition in complx acoustic environments: a review. Journal of Avian Biology, 49(5), pp.1-27.
Stowell, D., Wood, M.D., Pamuła, H., Stylianou, Y. and Glotin, H., 2019. Automatic acoustic detection of birds through deep learning: the first Bird Audio Detection challenge. Methods in Ecology and Evolution, 10(3), pp.368-380.
Sueur, J., Farina, A., Gasc, A., Pieretti, N. and Pavoine, S., 2014. Acoustic indices for biodiversity assessment and landscape investigation. Acta Acustica united with Acustica, 100(4), pp.772-781.