ENGINEERS FOR EXPLORATION

San Diego Zoo Wildlife Alliance

Challenges in Applying Audio Classification Models to Datasets Containing Crucial Biodiversity Information

Jacob Ayers

Project Lead - Engineers for Exploration Automated Acoustic Species Identification

Education

- B.S. in Electrical Engineering at UC San Diego (2022)
- M.S. in Electrical Engineering Data Science/Machine Learning at UC San Diego (2023)

Research Interests

- Climate Change, Deep Learning, Machine Learning, Data Science, Ecology, Natural Language Processing, Statistics, Ecoacoustics, Digital Signal Processing, Ornithology, Indicator Species, and Sound-processing

Yaman Jandali

Technical Contributor - Engineers for Exploration Automated Acoustic Species Identification

Education

- ❖ B.S. in Computer Science at UC San Diego (2020)
- ❖ M.S. in Electrical Engineering Data Science/Machine Learning at UC San Diego (2022)

Research Interests

- ❖ Machine Learning for Social Good, Identification and Generation of Audio, AI Misuse Prevention, Adversarial Machine Learning

Motivation to Study Biodiversity Loss

Impact of 2019–2020 mega-fires on Australian fauna habitat

Michelle Ward, Ayesha I. T. Tulloch, James Q. Radford, Brooke A. Williams, April E. Reside, Stewart L. Macdonald, Helen J. Mayfield, Martine Maron, Hugh P. Possingham, Samantha J. Vine, James L. O'Connor, Emily J. Massingham, Aaron C. Greenville, John C. Z. Woinarski, Stephen T. Garnett, Mark Lintermans, Ben C. Scheele, Josie Carwardine, Dale G. Nimmo, David B. Lindenmayer, Robert M. Kooyman, Jeremy S. Simmonds, Laura J. Sonter & James E. M. Watson

Nature Ecology & Evolution **4**, 1321–1326 (2020) | Cite this article

Climate Change, Deforestation, and the Fate of the Amazon

Yadvinder Malhi[1,*], J. Timmons Roberts[1,2], Richard A. Betts[3], Timothy J. Killeen[4], Wenhong Li[5], Carlos A. Nobre[6]
+ See all authors and affiliations

Science 11 Jan 2008:
Vol. 319, Issue 5860, pp. 169-172
DOI: 10.1126/science.1146961

IPBES-IPCC Co-Sponsored Workshop: Spotlighting Interactions of the Science of Biodiversity and Climate Change

Biodiversity Monitoring

Animal Tracks

Feeding Sites

Capture

Machine Learning in Biodiversity Monitoring

Camera Trap Arrays

- ❖ Natural next step of biodiversity monitoring due to advancements in the field of image classification that has been driven by deep learning.
- ❖ Relatively easy for laymen to annotate training and test sets of animal images.
  - ➢ Leads to citizen science projects
- ❖ Hardware is often sophisticated with motion sensing technology as well as night vision.
  - ➢ More points of failure in the hardware
- ❖ Limited to studying larger, oftentimes mammalian species.
- ❖ Clear lines of sight must be established which can necessitate clearing out natural obstacles in staging area.

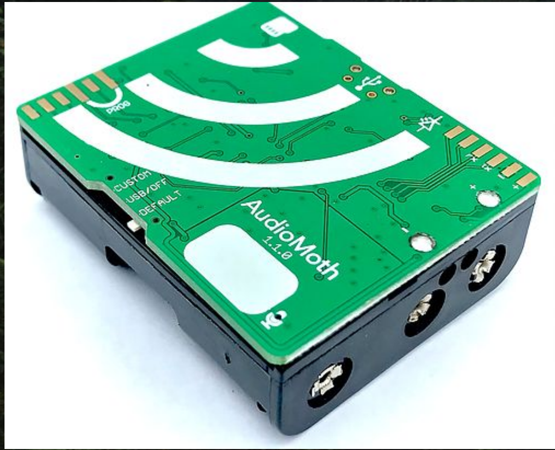
Passive Acoustic Monitoring

Audio Arrays

- Another natural next step in the field of biodiversity monitoring due to advancements in the field of natural language processing
  - Also leverages techniques from image processing with spectrograms
- Collect audio from species that are too inconvenient for camera trap capture
  - Many of which are key indicator species
- Low-cost open-source audio recorders that is minimally invasive
- Challenging for laymen to annotate audio recordings
- Not a wide abundance of labeled audio to reliably train models and classify most species of interest.
- Challenging to label audio clips

Summer 2019 Peruvian Amazon Audiomoth Deployment

- 35 Audiomoths deployed
- Deployed in Madre de Dios, Peru
- Deployed by San Diego Wildlife Alliance Researchers
- Deployed nearby roads on land managed by an FSC certified logging company
- Set to record for one minute every ten minutes
- 3.9 terabytes/1500 hours of audio data collected



Resources Used

Microfaune

- Hybrid Recurrent Neural Network - Convolutional Neural Network model
  - Model derived from Veronica Morfi and Dan Stowell
- Github Link: https://github.com/microfaune/microfaune

DCASE 2018 - Training/Validation Dataset

- “freefield1010” - 7690 field recordings of bird presence/absence from around the world
- “warblr10k” - 8000 bird presence/absence smartphone audio recordings from around the United Kingdom
- 80/20 split of these datasets from training and validation respectively
- Dataset Link: http://machine-listening.eecs.qmul.ac.uk/bird-audio-detection-challenge/

Resources Used

Xeno-canto/Google Audioset Ontology - Testing Dataset

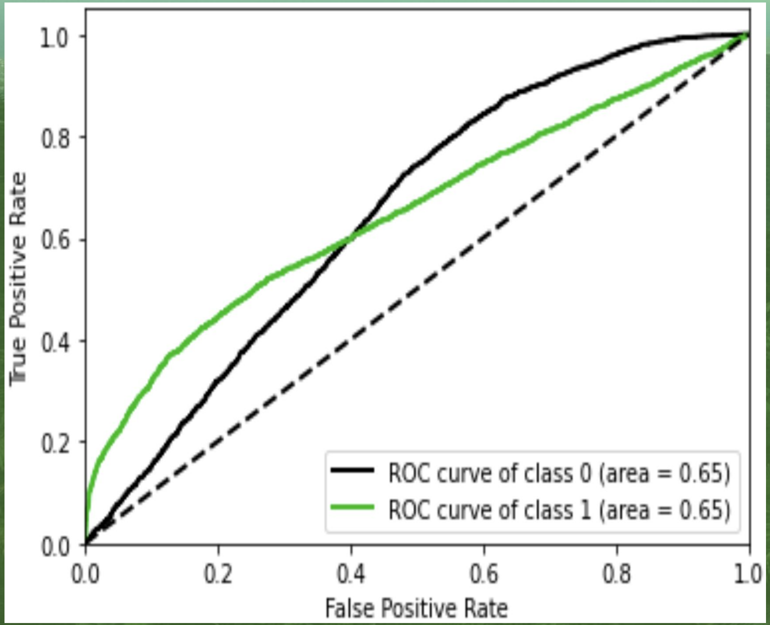
- 4774 bird-present audio clips of Madre de Dios species scraped from xeno-canto
  - 2-3 random clips from ~ 1000 bird species
  - 50 random clips from 50 particular species of interest
- 4774 Google Audioset Ontology audio clips from classes unlikely to contain bird calls.

Audiomoth Data - Testing Dataset

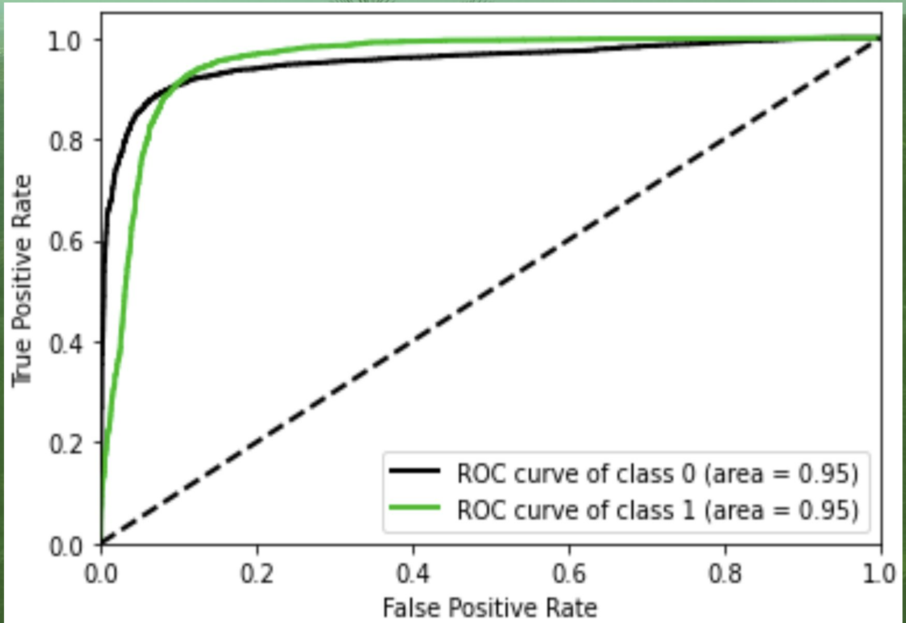
- Stratified Random Sample of Madre de Dios field recordings
  - One minute from every hour of the day from each Audiomoth device
- Stratified clips from 16 Audiomoths were taken and broken down into 3 second segments a labeled for bird presence/absence
  - 7120 3 second clips in total

Results after Training on Original Data

Audiomoth ROC Curve



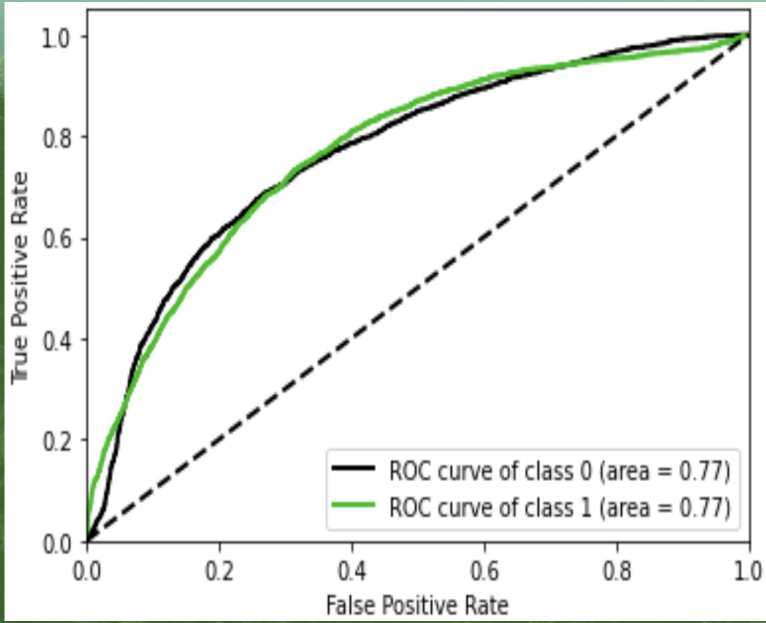
Xeno-canto ROC Curve



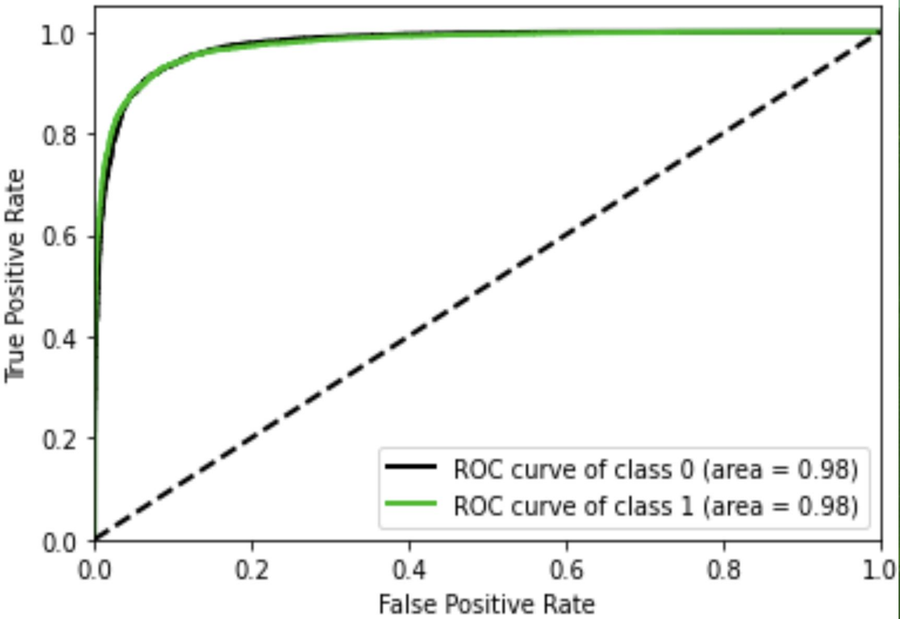
Results after Training on Augmented Data

Speed Augmentations: 0.9, 1.1; Gaussian Noise Augmentations: 0.005, 0.1

Audiomoth ROC Curve

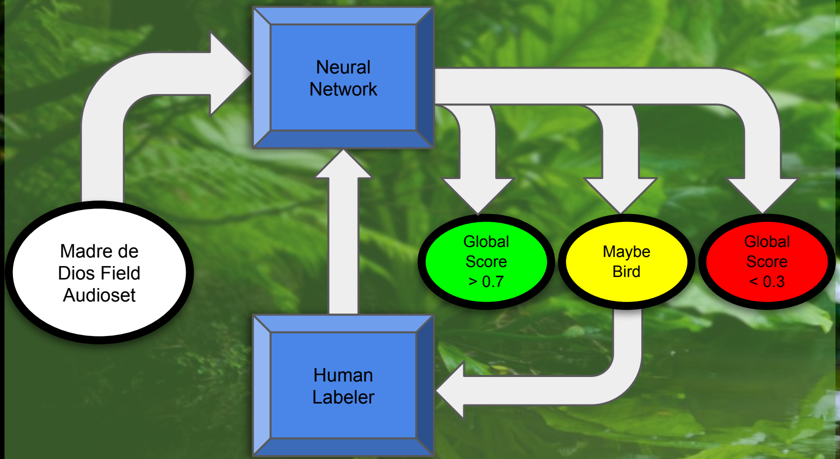


Xeno-canto ROC Curve



Future Plans to Tackle Field Recordings

Active Learning



Referee Labeling Process

Audio	Labeler 1	Labeler 2	Referee
Clip 1	Bird	Bird	N/A
Clip 2	Bird Absent	Bird Absent	N/A
Clip 3	Bird	Bird Absent	Bird *Final label for clip 3

New Data Augmentation Techniques

- Pink Noise
- Tempo Modulation
- Pitch Modulation
- Random Filtering
- Salt and Pepper

Transfer Learning

- Derive lower layers from a larger neural network and train final layers on dataset of interest.
- Has the potential of skewing very powerful networks towards an ecosystem of interest
- Reduces dataset size and training time.

Further Information

- Paper Github Repository: https://github.com/UCSD-E4E/AID_ICML_2021
- Engineers for Exploration: http://e4e.ucsd.edu/
- San Diego Zoo Wildlife Alliance: https://science.sandiegozoo.org/population-sustainability

E4E ENGINEERS FOR EXPLORATION

San Diego Zoo Wildlife Alliance

Contact Information

- Jacob Ayers:
  Email: jgayers@ucsd.edu
  LinkedIn: https://www.linkedin.com/in/jgayers/
- Yaman Jandali:
  Email: yamanjandali@gmail.com
  Website: https://www.yamanjandali.com