UNCG

The Department of
Computer Science

BATALOG

AN INTEGRATED SOLUTION SUITE FOR BAT SCIENTISTS

By Hadi Soufi, Thien Le, Kevin Keomalaythong, David Massey, Rahim Iqbal
Advisors: Dr. Somya Mohanty, Dr. Han Li

try it yourself | source code | contact us

INTRODUCTION

Bat scientists at UNCG record bat calls from all over the state. While bats make hundreds of echolocation calls per night the researchers are interested in so-called abnormal calls, a poorly understood region of bat behavior. Sorting through their data and finding all the abnormal calls can take years of work. Batalog seeks to classify bat calls automatically in a user-friendly web interface using deep learning techniques and custom algorithms.

NOISE REMOVAL

The first step in processing bat calls for classification is to extract them from potentially noisy Zero Crossing files. To do this, we attempt to split a given file into its component calls by simply checking for a threshold distance between points. From there, we perform a slope threshold check, preserving areas of the calls that move in a consistent direction. Then we compare each call to a smoothed version of itself generated using a Savitzky Golay filter and remove noisy areas. Algorithm parameters can be tuned for optimal results.

UTILITIES

Download Zip

Image Mapping

Display Images

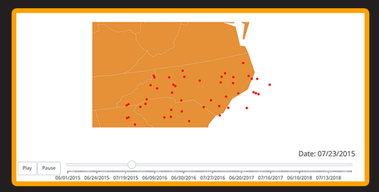
CNN DESIGN

The purpose of the CNN (Convolutional Neural Network) is to classify images as echolocation or abnormal calls. The CNN receives images as PNG images that are of size 200x300. The images are processed through six hidden layers. The output from the CNN is a value that represents the prediction. An output of 0 represents an abnormal call and a 1 represents an echolocation call. The convolutional layers consist of filters that are of size 5 with a stride of 5. The pooling layers consist of filters that are of size 2 with a stride of 2. Training abnormal call images were flipped and mirrored and echolocation call images were rotated 2° to increase the training data and anticipate future calls of unknown bat species. CNN accuracy from testing is 99% for echolocation files and 91% for abnormal files.

RESULTS

After processing ZC data, we serve it to the user in an interactive and intuitive way. The data is split according to classification, and each category is fully interactive. The user can navigate, select, and view every pulse in their library.

If the uploaded ZC data has GPS metadata, we can generate a map of bat call locations crossed by time. This allows researchers to easily visualize trends in abnormal bat call behavior.

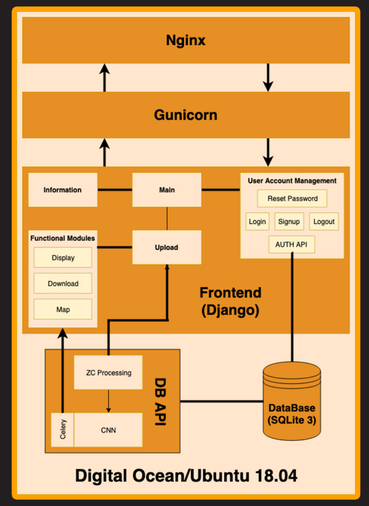


TECHNOLOGIES

DigitalOcean, django, SQLite, Python, Keras, TensorFlow

SYSTEM DESIGN

We used a fairly standard layout for setting up our website; a Django framework for building web pages, a backend for database interactions & a Celery wrapper for asynchronous pulse rendering, Gunicorn as our WSGI middleware, and finally Nginx for our web server. Since none of us had ever built a website before, a lot of this design was built on the fly and it has been modified heavily as we've learned more about our system requirements



CONCLUSION

Contributions to Users:

- Reduce time in the labeling process.

Contribution to Industry:

- Novel algorithms for bat call analysis
- First to provide a user friendly environment
- Open source project