

**E<sub>4</sub> ENGINEERS  
FOR  
E EXPLORATION**



**San Diego Zoo  
Wildlife Alliance**

# **A Deep Learning Approach to the Automated Segmentation of Bird Vocalizations from Weakly Labeled Crowd-sourced Audio**

**Jacob Glenn Ayers, Sean Perry, Samantha Prestrelski, Tianqi Zhang, Ludwig von Schoenfeldt, Mugen Blue, Gabriel Steinberg, Mathias Tobler, Ian Ingram, Curt Schurgers, Ryan Kastner**

# The Amazon is Under Threat



4 MIN READ

## Amazon inhales more carbon than it emits, NASA finds



(Tracking Amazon Deforestation from Above, NASA. <https://earthobservatory.nasa.gov/images/145988/tracking-amazon-deforestation-from-above>)

(“Amazon inhales more carbon than it emits, NASA finds”, NASA,

<https://science.nasa.gov/science-research/earth-science/carbon-cycle/amazon-inhales-more-carbon-than-it-emits-nasa-finds/>)

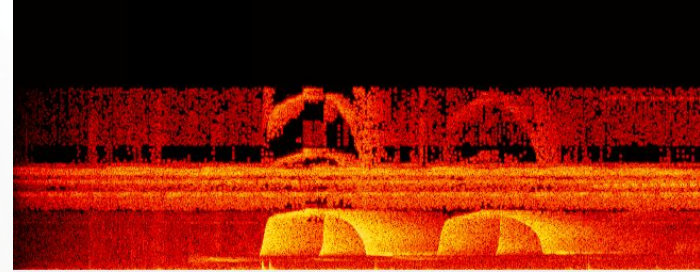
# Measure the Health of the Amazon



Record



Identify



Screaming  
Piha

(Screaming Piha, *Gabriel Bonfa*. <https://macaulaylibrary.org/asset/153969671>.)

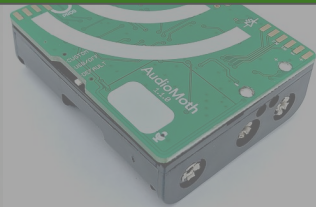
(Screaming Piha, *Lipaigus vociferans*. Audio data obtained from

Campbell, Oscar. XC332895 · Screaming Piha · Lipaugus Vociferans. <https://www.xeno-canto.org/332895>.)

# Measure the Health of the Amazon



**This is Expensive to Do By Hand...  
Lets Automate With AI!**



Screaming  
Piha

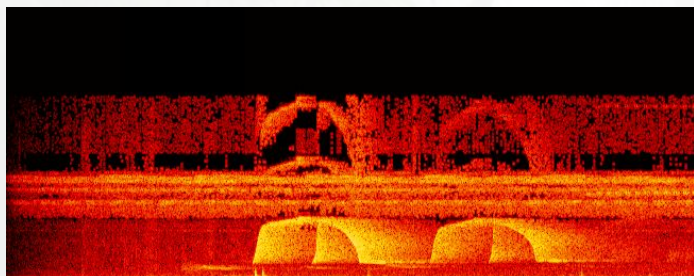
(Screaming Piha, *Gabriel Bonfa*. <https://macaulaylibrary.org/asset/153969671>.)

(Screaming Piha, *Lipaugus vociferans*. Audio data obtained from

Campbell, Oscar. XC332895 · Screaming Piha · Lipaugus Vociferans. <https://www.xeno-canto.org/332895>.)

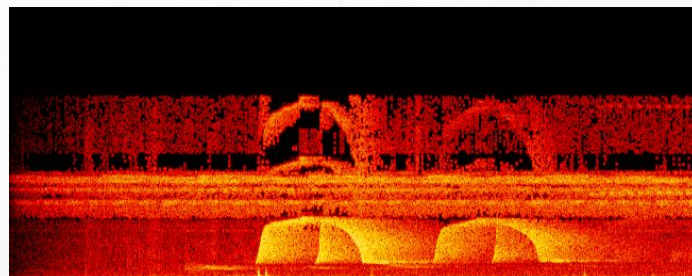
# Weakly Vs Strongly Labeled Problem

**Weakly Labeled**  
Majority of Training Data



Screaming  
Piha

**Strongly Labeled**  
Ideal for Supervised Training



Screaming  
Piha

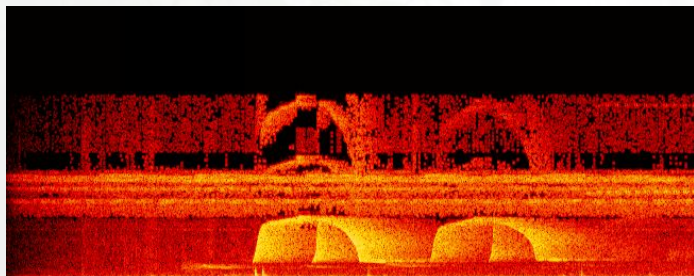
(Screaming Piha, *Lipaugus vociferans*. Audio data obtained from Campbell, Oscar. XC332895 · Screaming Piha · Lipaugus Vociferans. <https://www.xeno-canto.org/332895>.)



**How can we Better Filter Weakly Labeled Data?**

# Weakly Vs Strongly Labeled Data

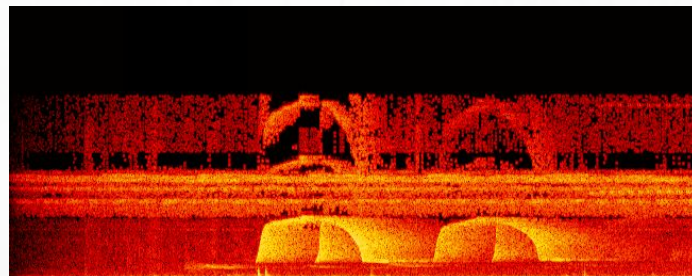
Weakly Labeled



Screaming  
Piha



Strongly Labeled (Bird)



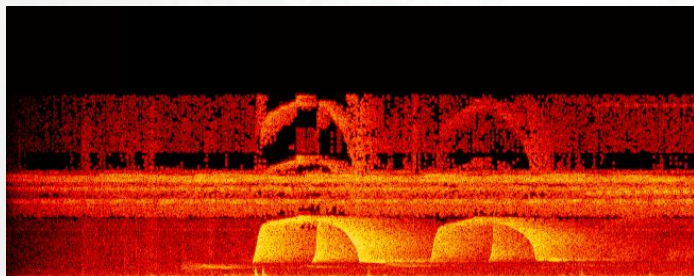
Bird

Bird

(Screaming Piha, *Lipaugus vociferans*. Audio data obtained from Campbell, Oscar. XC332895 · Screaming Piha · Lipaugus Vociferans. <https://www.xeno-canto.org/332895>.)

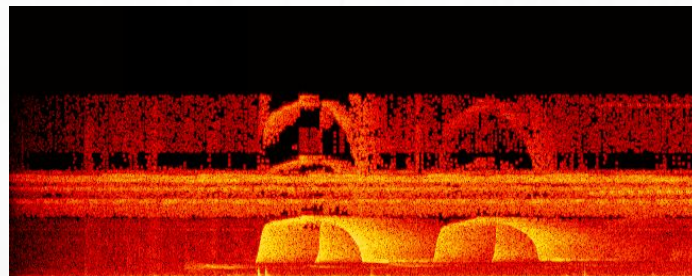
# Weakly Vs Strongly Labeled Data

Weakly Labeled

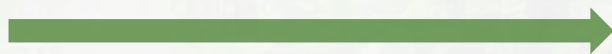


Screaming  
Piha

Strongly Labeled



Screaming  
Piha    Screaming  
Piha





# Weakly Vs Strongly Labeled Data

Weakly Labeled

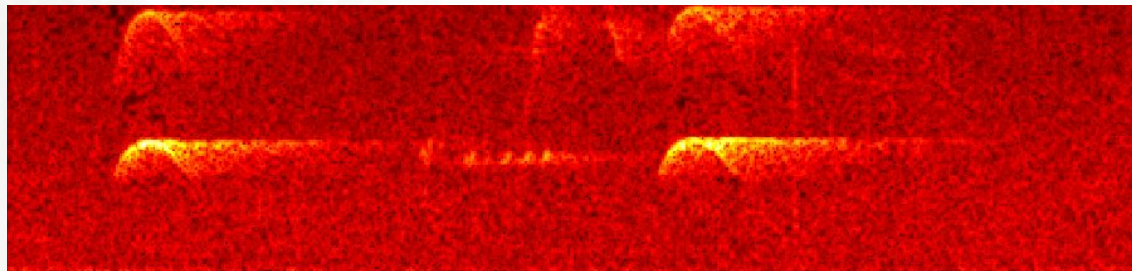
Strongly Labeled

**How is This Currently Done?**

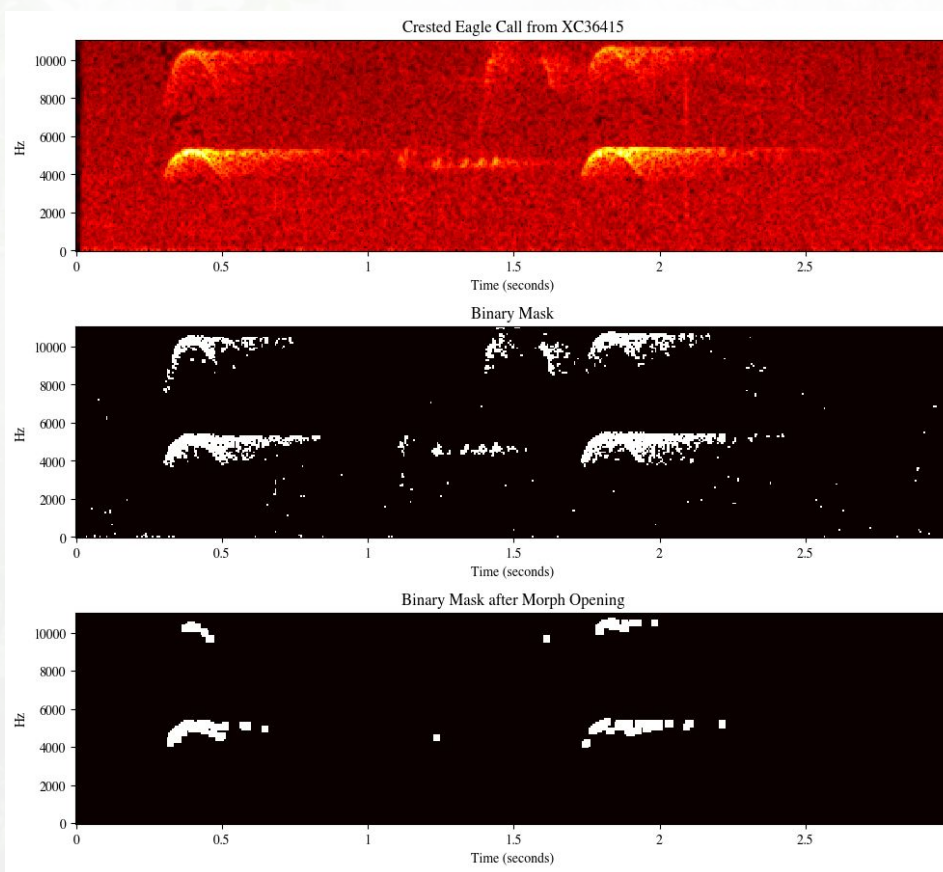
Screaming  
Piha

Screaming  
Piha    Screaming  
Piha

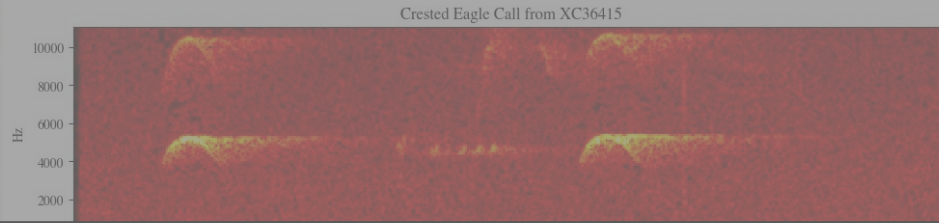
# Current Methods



# Current Methods



# Current Methods



**Can we use AI To Increase the Precision of Generated Strong Labels?**



# Experiment

1) Collect a Dataset of Many Species in XC

# Experiment

- 1) Collect a Dataset of Many Species in XC
- 2) Create 3 second chunks

# Experiment

- 1) Collect a Dataset of Many Species in XC
- 2) Create 3 second chunks
- 3) For each 3 second chunk, use FGBG and preexisting ML pipelines to identify bird not bird

# Experiment

- 1) Collect a Dataset of Many Species in XC
- 2) Create 3 second chunks
- 3) For each 3 second chunk, use FGBG and preexisting ML pipelines to identify bird not bird
- 4) Replace all bird labels with Weak Label



# Experiment

- 1) Collect a Dataset of Many Species in XC
- 2) Create 3 second chunks
- 3) For each 3 second chunk, use FGBG and preexisting ML pipelines to identify bird not bird
- 4) Replace all bird labels with Weak Label
- 5) Compare with human annotators

# Test Two Binary/Bird Models

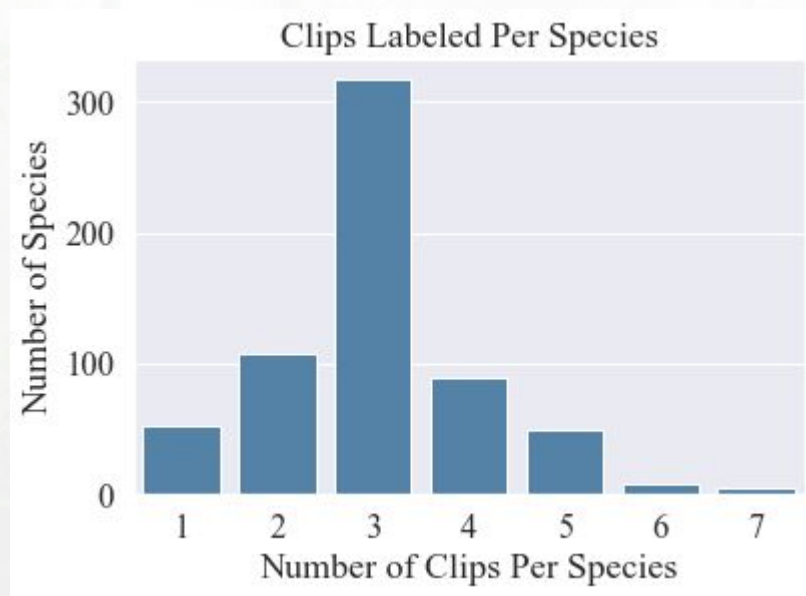
## Microfaune (Preexisting Method)

- Designed by Veronica Morfi and Dan Stowell
- Training Data:
  - Warblr
  - Freefield1010
  - BirdVox
- **RNN**

## Tweetynet

- Designed by Cohen Et Al.
- Retrained by us on:
  - Freefield1010
  - Warblr
- **CNN/RNN**

# Labeling



# Labeling

The image displays six individual audio recording cards from the Xeno-Canto website, arranged in a 2x3 grid. Each card includes a spectrogram, a waveform, a play button, a progress bar, and a stop button. The recordings are as follows:

- XC525583**: Screaming Piha *Lipaugus vociferans*. Mauricio Álvarez-Rebolledo (Colección de Sonidos Ambientales - Instituto Humboldt). PNN Chiribiquete, cunaré- Amú, Colombia. 2001-01-02. Duration: 0:00 to 1:59. IAVH-CSA-06356 To obtain a wav file of... [more »](#). Labels: A | B | C | D | E.
- XC525378**: Screaming Piha *Lipaugus vociferans*. Mauricio Álvarez-Rebolledo (Colección de Sonidos Ambientales - Instituto Humboldt). PNN Chiribiquete, cunaré- Amú, Colombia. 2001-01-02. Duration: 0:00 to 1:31. IAVH-CSA-06614 To obtain a wav file of... [more »](#). Labels: A | B | C | D | E.
- XC425555**: Screaming Piha *Lipaugus vociferans* - call. Joshua Weiss. Tambopata, Madre de Dios, Peru. 2017-07-30. No modification Habitat: dense forest... [more »](#). Labels: A | B | C | D | E.
- XC939878**: Screaming Piha *Lipaugus vociferans* - song. JAYRSON ARAUJO DE OLIVEIRA. Corredeiras do Bem Querer, Roraima, Brazil. 2024-10-03. Duration: 0:00 to 0:25. Labels: A | B | C | D | E. [also]
- XC934752**: Screaming Piha *Lipaugus vociferans*. Yannick Wiegiers. Kourou, Arrondissement of Cayenne, French Guiana. 2023-05-19. Duration: 0:00 to 0:11. Labels: A | B | C | D | E.
- XC901262**: Screaming Piha *Lipaugus vociferans* - song. Jaime Suárez. La Uribe, Meta, PNN Cordillera de los Picachos, Colombia. 2013-08-11. Duration: 0:00 to 0:10. Labels: A | B | C | D | E.

(Screaming Piha, *Xeno-Canto*. <https://xeno-canto.org/species/Lipaugus-vociferans?view=3>)

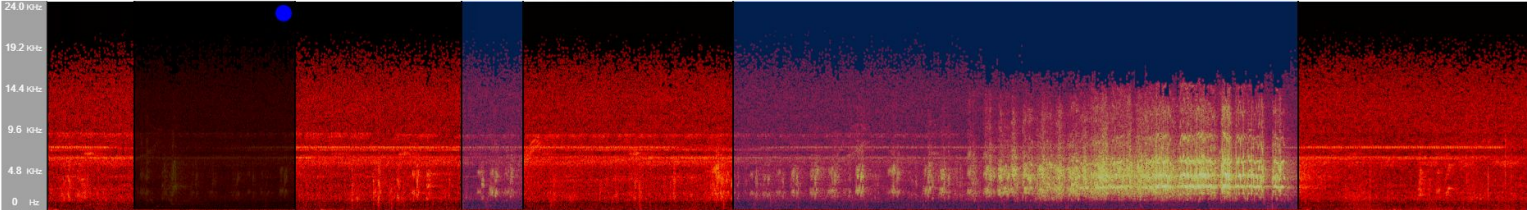
# Labeling

Pyrenote [Dashboard](#) [Admin Panel](#) [Logout](#)

Segment saved ✕

>>

Brotogeris-cyanoptera-129175.wav [Add Regions: On](#)



24.0 kHz  
19.2 kHz  
14.4 kHz  
9.6 kHz  
4.8 kHz  
0 Hz

⏪ ⏩

Species  
Brotogeris cyanoptera Cobalt-winged Parakeet

Delete Save All No Relevant Audio

Feeling Confident In your labels?: No

Previous Next

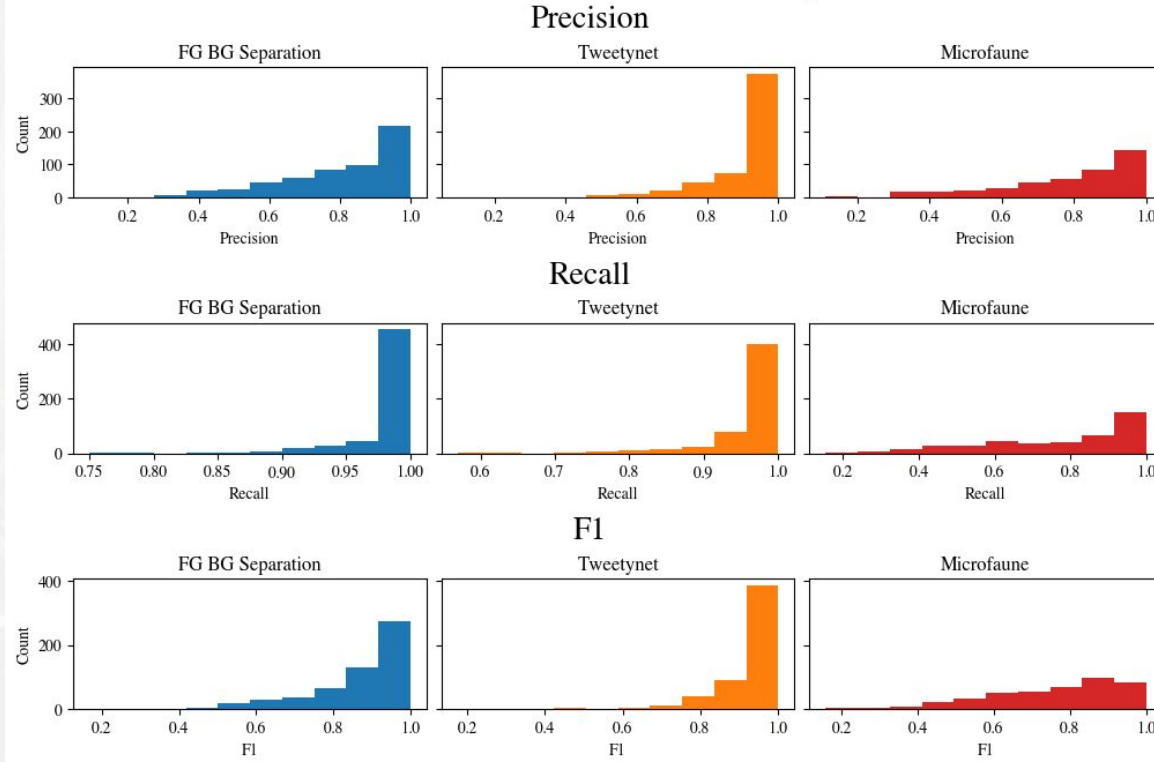
# Results

Technique	Time Ratio	Number of 3 Second Segments	Precision	Recall	F1
FG BG Sep	1.000	<b>21582</b>	0.7797	<b>.9831</b>	0.8697
Microfaune	1.329	13200	0.7767	0.7062	0.7398
TweetyNet	<b>0.853</b>	18365	<b>0.9009</b>	0.9704	<b>0.9344</b>

Table 1: The aggregate metrics for the results of each model.

# Results

## Performance of Models Across Species



## Discussion

Demonstrated a framework for Weakly to Strongly Label pipelines that decreases the amount of false positive labels seen during training



## Discussion

Tweetynet reduces the most false positives in training data

FGBG gets largest recall:  
mirrors hypothesis

# Discussion

## Limitations Future Work

- Downstream Model Classifier experiments based on different Weak to Strongly label pipelines

# Discussion

## Limitations Future Work

- Downstream Model Classifier experiments based on different Weak to Strongly label pipelines
- Transformer based methods

# Discussion

## Limitations Future Work

- Downstream Model Classifier experiments based on different Weak to Strongly label pipelines
- Transformer based methods
- Expert labels

# Discussion

The screenshot shows the GitHub repository page for 'PyHa'. The repository is owned by 'UCSD-E4E' and has 16 stars, 12 watchers, and 11 forks. The main content area features a README with a title 'PyHa' and a description: 'A tool designed to convert audio-based "weak" labels to "strong" moment-to-moment labels. Provides a pipeline to compare automated moment-to-moment labels to human labels. Current proof of concept work being fulfilled on Bird Audio clips using Microfaune predictions.' Below the description is a link to the 'Engineers for Exploration Acoustic Species Identification Team' and the 'San Diego Zoo Wildlife Alliance'. A logo for PyHa is shown, featuring a brown bird with its beak open, set against a green circular background with the letters 'PYHA' above it. The README also includes a section for 'Contents' with links to 'Installation and Setup', 'Functions', and 'Examples', and a section for 'Installation and Setup' with four numbered steps. The right sidebar shows repository statistics: 15 releases (latest is v0.2.1 from last month), 0 packages, 20 contributors (6 shown), 50 deployments (github-pages from last month), and a language bar showing 96.5% for Jupyter Notebook and 3.5% for Python 3.5.

README License

## PyHa

A tool designed to convert audio-based "weak" labels to "strong" moment-to-moment labels. Provides a pipeline to compare automated moment-to-moment labels to human labels. Current proof of concept work being fulfilled on Bird Audio clips using Microfaune predictions.

This package is being developed and maintained by the [Engineers for Exploration Acoustic Species Identification Team](#) in collaboration with the [San Diego Zoo Wildlife Alliance](#).

PyHa = Python + Piha (referring to a bird species of our interest known as the screaming-piha)

### Contents

- [Installation and Setup](#)
- [Functions](#)
- [Examples](#)

### Installation and Setup

- Navigate to a desired folder and clone the repository onto your local machine. `git clone https://github.com/UCSD-E4E/PyHa.git`
- If you wish to reduce the size of the repository on your local machine you can alternatively use `git clone https://github.com/UCSD-E4E/PyHa.git --depth 1` which will only install the most up-to-date version of the repo without its history.
- Install Python 3.8, Python 3.9, or Python 3.10
- Create a `venv` by running `python3.x -m venv .venv` where `python3.x` is the appropriate python.
- Activate the `venv` with the following commands:

Custom properties

- 16 stars
- 12 watching
- 11 forks
- Report repository

### Releases 15

v0.2.1 Latest  
last month

+ 14 releases

### Packages

No packages published  
[Publish your first package](#)

### Contributors 20

+ 6 contributors

### Deployments 50

github-pages last month

+ 49 deployments

### Languages

Jupyter Notebook 96.5% Python 3.5%

**E<sub>4</sub> ENGINEERS  
FOR  
E EXPLORATION**



**San Diego Zoo  
Wildlife Alliance**

# **A Deep Learning Approach to the Automated Segmentation of Bird Vocalizations from Weakly Labeled Crowd-sourced Audio**

**Jacob Glenn Ayers, Sean Perry, Samantha Prestrelski, Tianqi Zhang, Ludwig von Schoenfeldt, Mugen Blue, Gabriel Steinberg, Mathias Tobler, Ian Ingram, Curt Schurgers, Ryan Kastner**