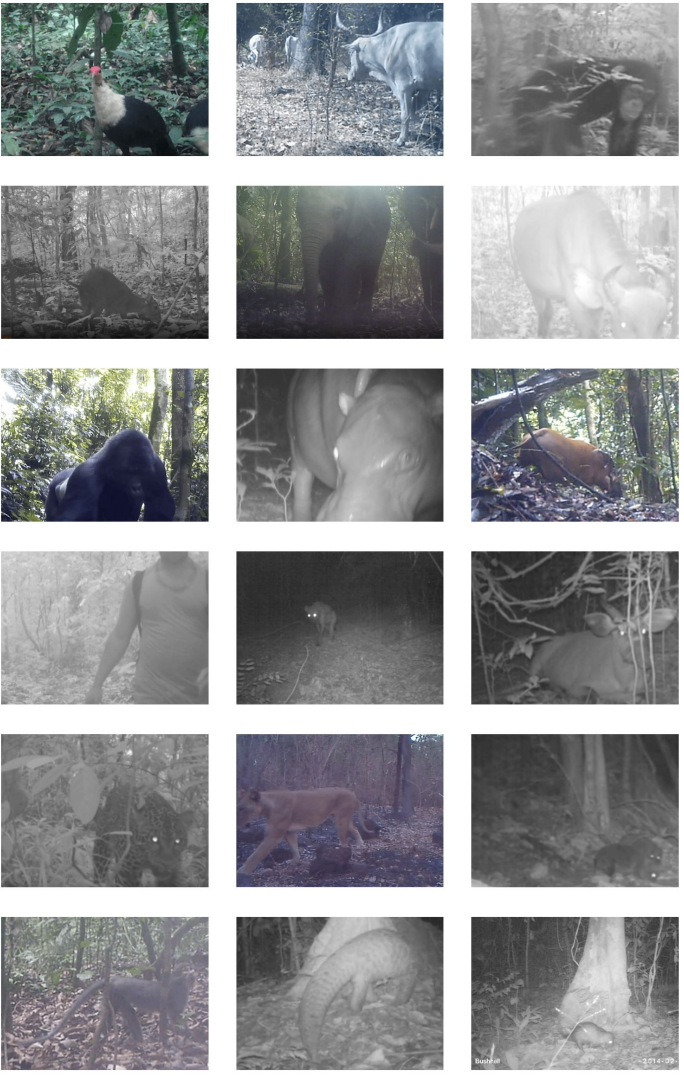
Computer vision for wildlife conservation

An introduction to Zamba

Emily Dorne
Lead Data Scientist

Agenda

- Motivation on working with videos
- Computer vision tools for camera trap videos
    - Zamba: python package
    - Zamba Cloud: web application
- Challenges and future directions



Researchers, conservationists, and park managers are using **camera traps to monitor wildlife.**

The problem is that these camera traps generate an **enormous amount of footage** that needs to be reviewed by human experts.

False triggers – caused by wind, rain, changes in light, etc. – are common, meaning **many videos do not contain an animal** at all.

How can we use machine learning to get to the videos we care about?

Working with videos instead of images

Why does zamba focus on videos

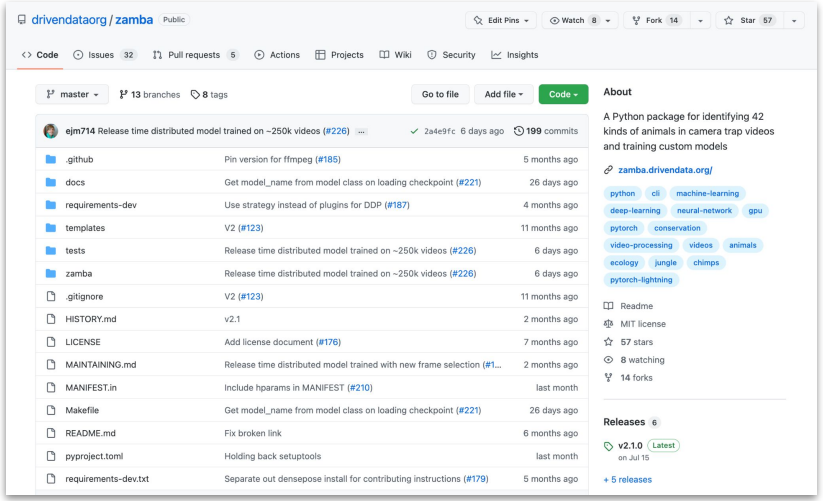
- Few tools out there for working with camera trap videos, which is a harder technical problem than working with images
- Increasing use of camera trap videos in the field because they contain so much more information
    - Animal behavior
    - By-catch
    - Audio
    - Multiple views of the animal
        - Sex
        - Size
        - Age
        - Individual identification
            - Capture re-capture

Why working with videos is hard

- Cannot treat it as an image problem
    - 32 frames per second * 60 seconds = 1,920 frames in a one minute video
    - Frame selection: how do we find the relevant parts of the video quickly and accurately?
- Video data is large
    - Greater compute resources needed for training and running models
    - Hurdles for uploading data for cloud-based applications

Our work

zamba



An open-source python package for identifying 42 kinds of animals in camera trap videos and training custom models.

Zamba Cloud

A web application where you can use the zamba models – without writing any code – by just uploading videos or pointing to where they are stored.

zamba

For those who know python

Zamba capabilities

Species classification

Predict using models trained on ~250,000 videos

- African forest (32 species)
- European (11 species)
- Blank detection

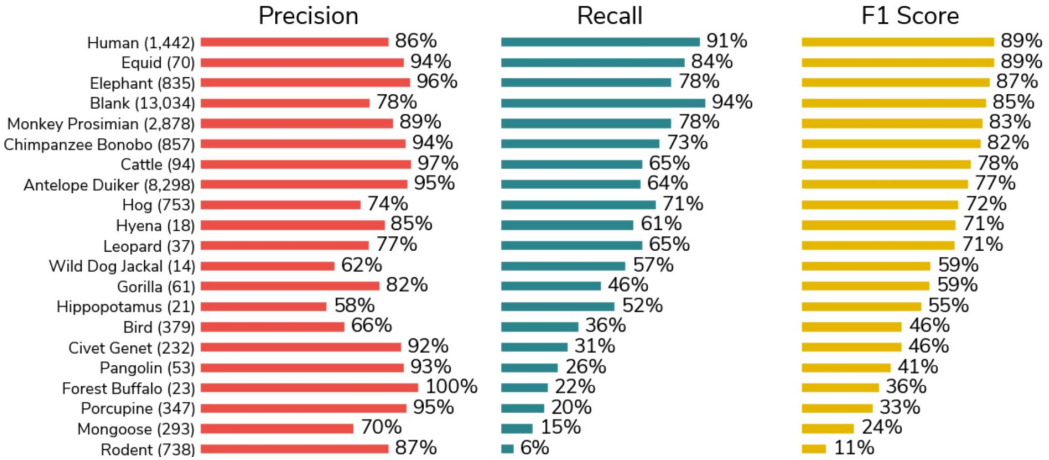
```
zamba predict --data-dir my_videos
```

Train your own models

- Finetune on new species and/or new geographies

```
zamba train --labels panthers.csv
```

Share custom models with the community



African forest model performance on holdout set

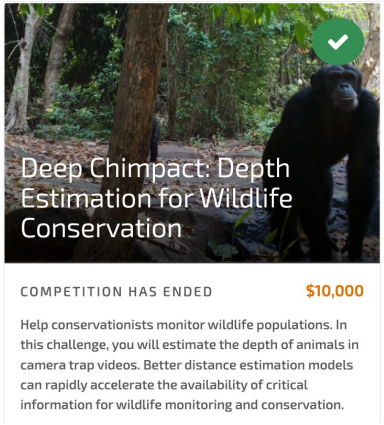
Zamba capabilities

Depth estimation

Estimated release: October 2022

Distance prediction at 1 frame per second for bushbucks, chimpanzees, duikers, elephants, leopards, monkeys

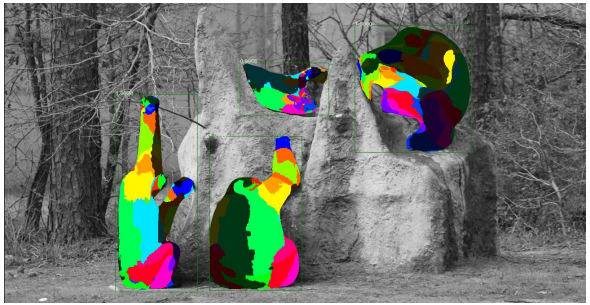
Uses the winning model from the depth estimation machine learning competition



Segmentation

DensePose is a model published by Facebook AI Research that can be used to get **segmentations** for animals that appear in videos.

The model provides mapping of the segmentation output to specific anatomy for chimpanzees.



Methods

Frame selection

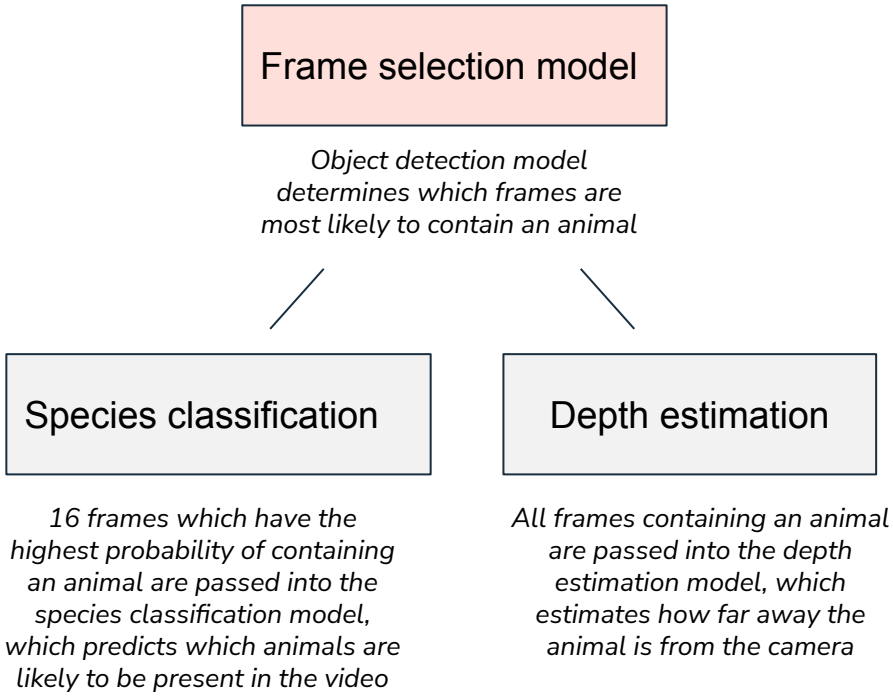
Frame selection is critical to working with videos

Species classification

- Run frame selection on every frame at 4 fps
- Select top 16 frames

Depth estimation

- Run frame selection on every frame at 1 fps
- Select all frames with an animal



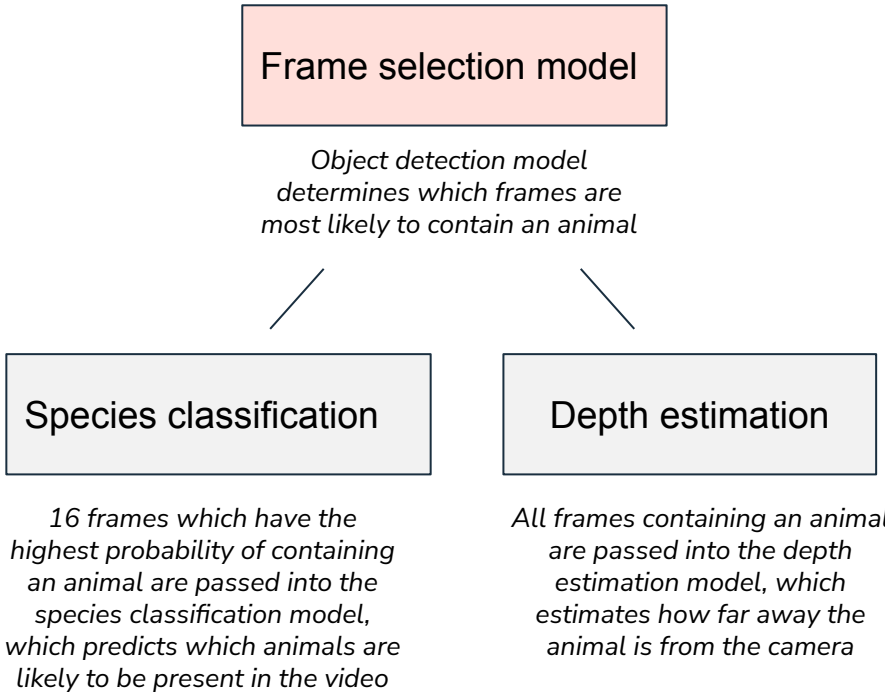
Methods

Frame selection

Use a “student-teacher” approach for training

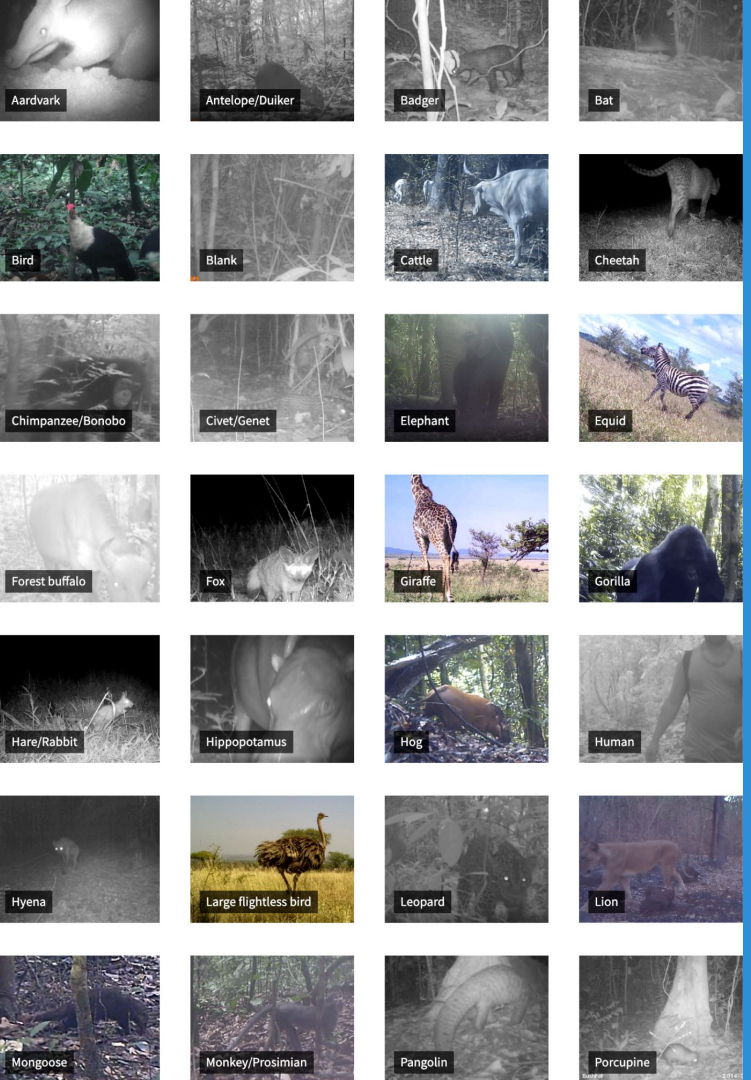
- Teacher labels come from Megadetector (v4) object detection predictions at the frame level

- Student model is a YOLOX nano architecture



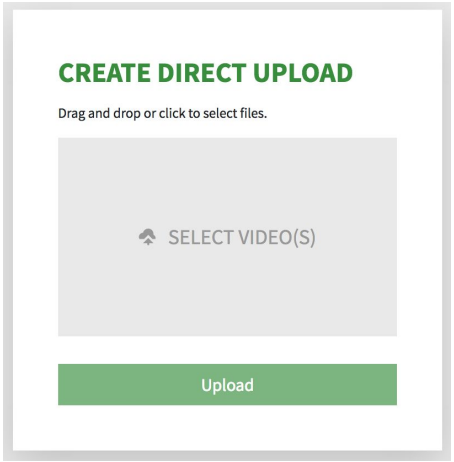
Zamba Cloud

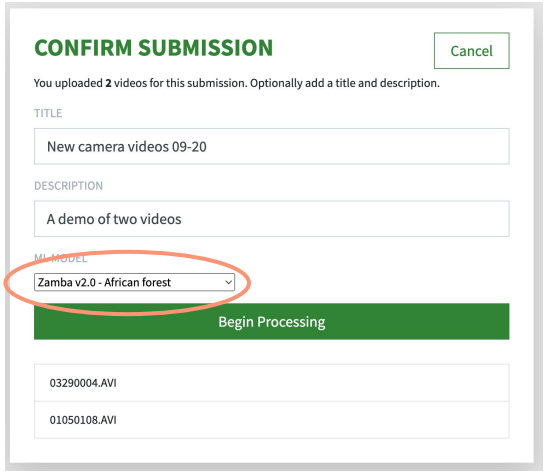
For those who want to point and click



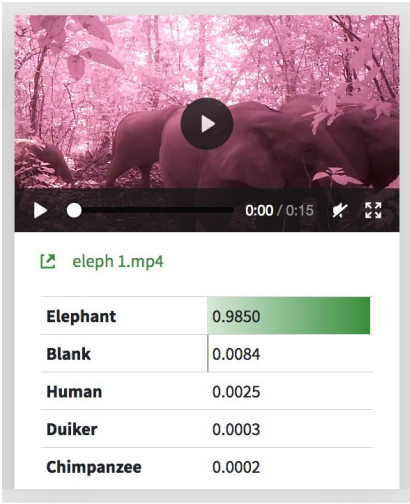
Zamba Cloud is an application that automatically identifies species groups in videos, making it much easier, cheaper, and faster to take advantage of camera trap footage.

Species classification





Model predictions run in the cloud



Upload videos directly or through an FTP server

Choose your model

- African forest
- European
- Blank nonblank

Review your predictions

- Predictions are outputted to a downloadable csv

Species classification

The output format is easy to use for analysis with Excel or statistical software.

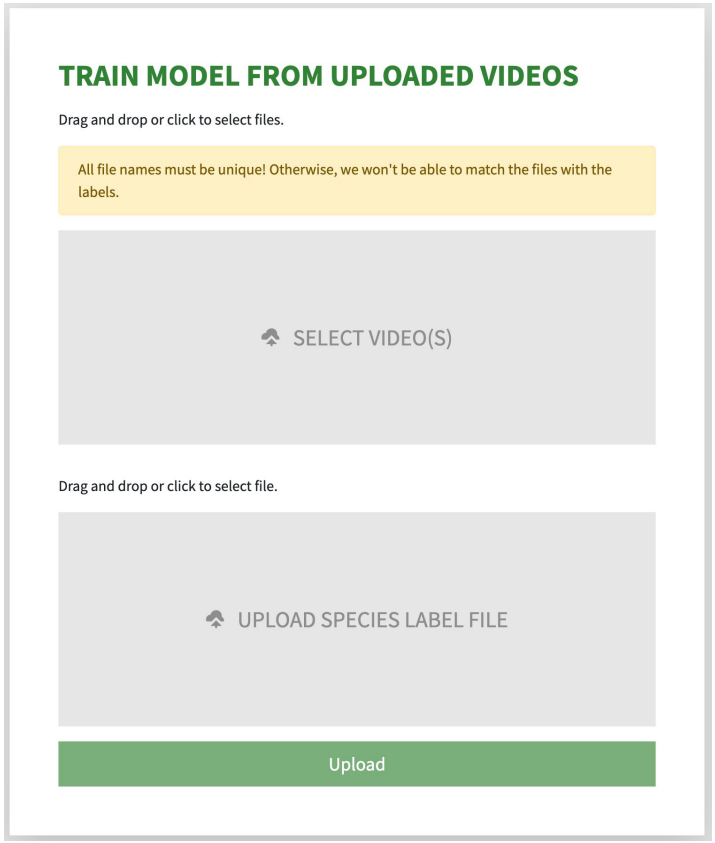
Researchers and conservationists can easily identify videos that are worth reviewing.
Plus, the format is open, so the data will always be reusable.

original_filename	top_1_label	top_1_probability	top_2_label	top_2_probability	...	BIRD	BLANK	CATTLE	...	corrected_label
12180011.AVI	BLANK	0.6446	OTHER_PRIMATE	0.1418		0.0021	0.6446	0.0004		
12060005.AVI	HUMAN	0.9460	BLANK	0.0571		0.0001	0.0571	0.0000		

Species classification

Users can also **generate a custom model** based on their own labeled videos, enabling Zamba Cloud to be adapted for any species in any ecosystem.

Fine-tuned models can then be shared between users and contributed to the broader community.



Looking ahead

Challenges

Technical

Rare or small species

Occlusion

Difficult cases

- Far from camera
- Bottom of screen
- Washed out videos
- Only present for a couple frames

Generalization to new locations

Accessibility

Uploads do not support low-bandwidth connections

Training models (even with a point and click interface) still requires some knowledge of ML

Hard to provide ex-ante guidance for custom model training

Sustainability

Cloud compute requires a consistent source of funding

Development covers a large surface area (model training, package, and web application) making progress slow

Use cases are overlapping but distinct

Future directions

- Combine the depth estimation with species classification to produce abundance estimates
- Develop a robust user base that trains and shares custom models
- Make Zamba Cloud more accessible for users with low-bandwidth connections
- Develop a sustainability model for development and cloud compute costs

If you'd like to partner with us on this work, reach out!

Thank you

Learn more: zamba.drivendata.org

Reach out: emily@drivendata.org

DRIVENDATA