Week 2

Goals

- [ ] GOAL 1: Proceed with pre-processing pipeline on the newly annotated 8000+ Ecuador images.
    - [ ] Produce visualizations on the new distribution at animal class, order and species level.
    - [ ] Create new datasets: 1) combining Brazil and Ecuador images 2) only Ecuador images

Finishing this goal will provide my team with cleaned larger datasets to re-train existing classification and detection models. This will also allow us to identify species/orders that requires more sample images to be extracted from video annotations.

- [ ] GOAL 2: Repeat model building process for object detection with new image datasets
    - [ ] Re-train models at binary, class, order levels
    - [ ] Produce evaluation metrics to compare with previous detection models

Finishing this goal will move my team closer to the updated detection models with all of the image data currently available to us. This will also allow us to quantify the prediction performance at current level of class

Expected Working Timeline

Name	Assign	Date	Expected Hours Spent	Status
Client Meeting		@Jan 25, 2021	1	Completed
Goal 1	j Jiayue Xu	@Jan 25, 2021 → Jan 29, 2021	6	In Progress
Goal 2	j Jiayue Xu	@Jan 30, 2021 → Jan 31, 2021	4	Not Started

Name	Assign	Date	Expected Hours Spent	Status
Internal Meeting		@Jan 31, 2021	1	Not Started

Report

Actual Working Timeline

Name	Assign	Date	Actual Hours Spent	Status
Client Meeting	Team	@Jan 25, 2021	1	Completed
Goal 1	Individual		2	Blocked
Goal 2	Individual			Blocked
Urgent Internal Meeting	Team	@Jan 27, 2021	1	Completed
Urgent Client Meeting	Team	@Jan 28, 2021	1	Completed
Revised Goal 1	Individual	@Jan 28, 2021 → Jan 29, 2021	4	Completed
Internal Meeting	Team	@Feb 1, 2021	1	Not Started

⛔ GOAL 1: Proceed with pre-processing pipeline on the newly annotated 8000+ Ecuador images.

While working on this goal, I discovered that 973 of the 8000+ new Ecuador images, which were initially classified into animal species by the client, but are missing bounding box annotations. Upon clarification with Ryan and Caitlin, who are client's data managers, we were informed that these are images that student volunteers fail to clear identify the animal species. This discrepancy surface a critical decision to be made: **should this images be labeled as "Ghost" images**

(does not contain animal species)? Through a urgent client meeting with Ryan, we came to the final decision that these images will be reviewed to classify into below 3 categories:

- If the animal species in the image is identifiable by human annotator, without any external information from other images, it will be labeled as the animal species and generate json file with bounding box annotation
- If it can be identified that there is an animal, but not clear which animal species it belongs to, the image will be labeled as "Unknown" and generate json file with bounding box annotation
- If animal cannot be clear identified from the background, the image will be labeled as "Ghost" and generate son file with no bounding box annotation

We also established that it is necessary to also review the old images that were marked as "Blur", and apply the same decision rule as above. There were 543 old images from Brazil and Ecuador that need to be reviewed for this reason.

✅ REVISED GOAL 1: Create a folder with copies of the 973 + 543 images to be reviewed, so that Ryan and Caitlin can proceed with the review process more efficiently
Commit ID: 7dd8bfdf

⛔ GOAL 2: Repeat model building process for object detection with new image datasets

Due to the incompletion of Goal 1, Goal 2 was not able to proceed and put on hold until the review process of the images were completed.