

Whale ReID: Identifying whales by their flukes



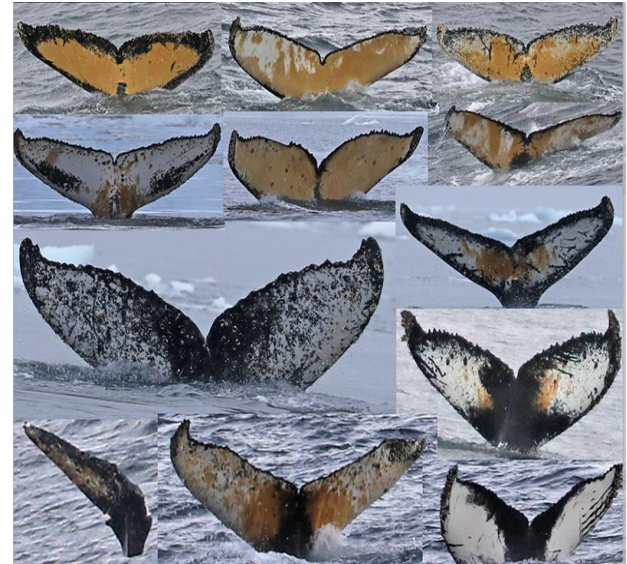
Group: 30

Name: Atharva Phatak

ID : 1141133

Introduction : Humpback whales and their decreasing population

- After centuries of intense whaling, recovering whale populations still have a hard time adapting to warming oceans and struggle to compete every day with the industrial fishing industry for food.
- To aid whale conservation efforts, scientists use photo surveillance systems to monitor ocean activity.
- They use the shape of whales' tails and unique markings found in footage to identify what species of whale they're analyzing and meticulously log whale pod dynamics and movements



Objective: Automatic Identification of Whale by their flukes

- For the past 40 years, most of this work has been done manually by individual scientists, leaving a huge trove of data untapped and underutilized.
- Recent progress in the field of computer vision has given us AI tools that can help us ID these whales !
- Computer vision algorithms have achieved SOTA performance on tasks like face recognition, image classification and object recognition.

Similar approaches can be applied to recognize whales by their caudal fins or flukes !



Face Recognition

- Facial recognition is the task of making a positive identification of a face in a photo or video image against a pre-existing database of faces.

Similar approach can be applied to recognize whales by their caudal fins or flukes !

Face Recognition

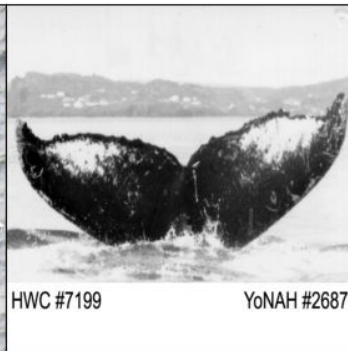
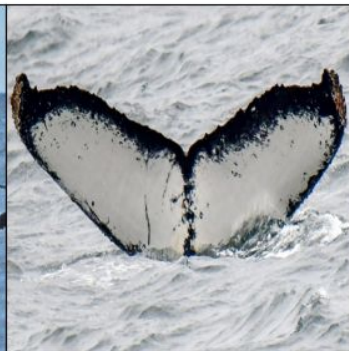
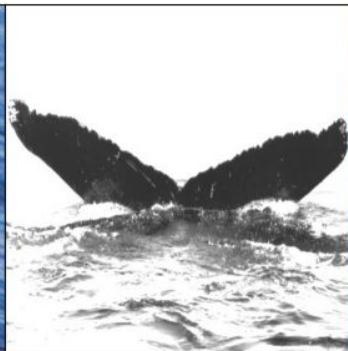


Whale Recognition



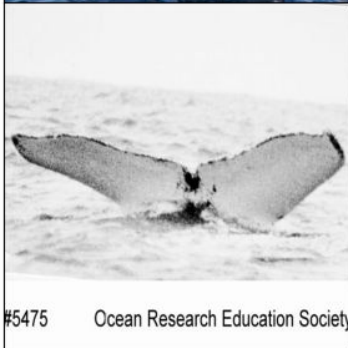
Exploratory Data Analysis and Data Pipelines

- Data Size ~ 33k images with varied resolution
 - Train Size ~ 25k images
 - Test Size ~ 8k images
- **5005 classes**
- The problems with the data
 - **Noisy images/Distorted Images**
 - **The dataset is long tailed i.e there are classes which consist of very less amount of samples.**
 - **The images are both RGB and Gray**



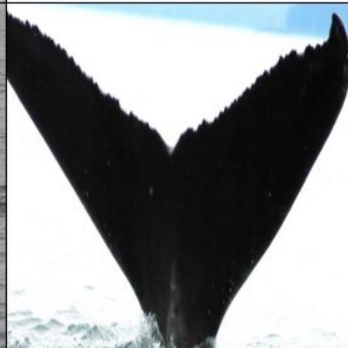
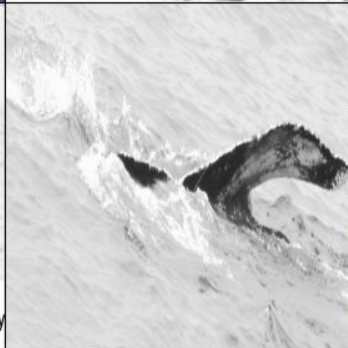
HWC #7199

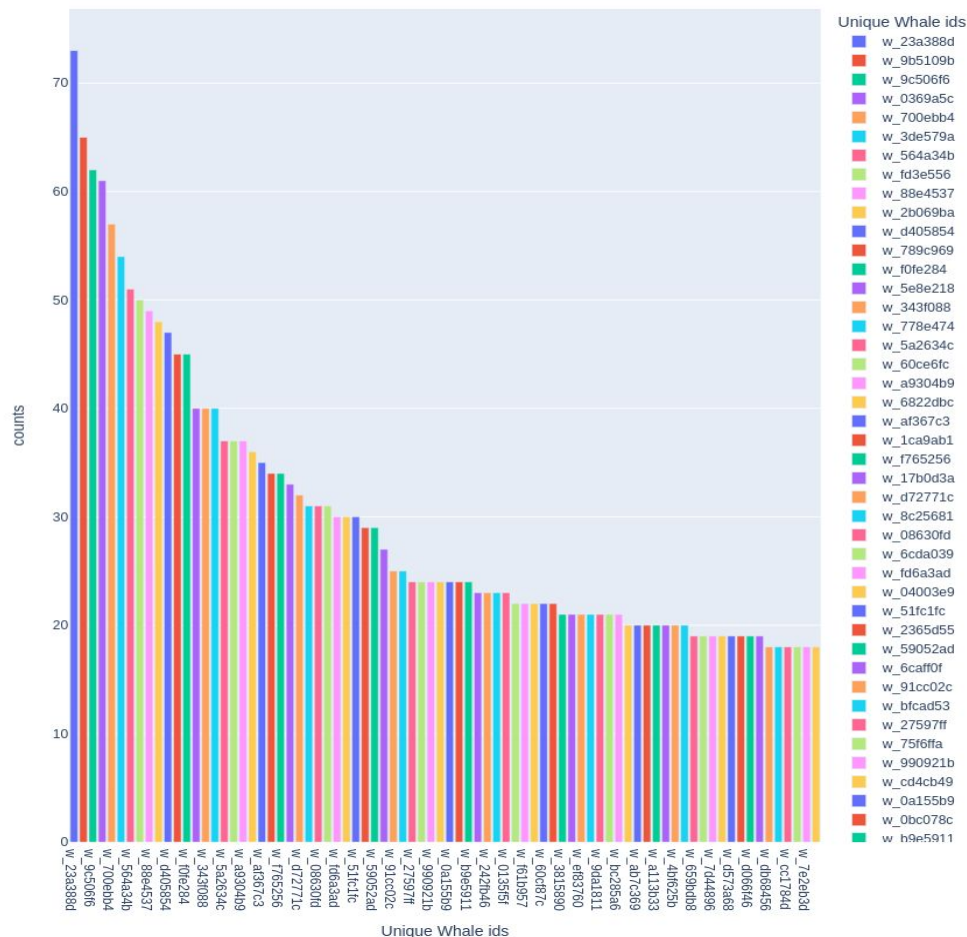
YoNAH #2687



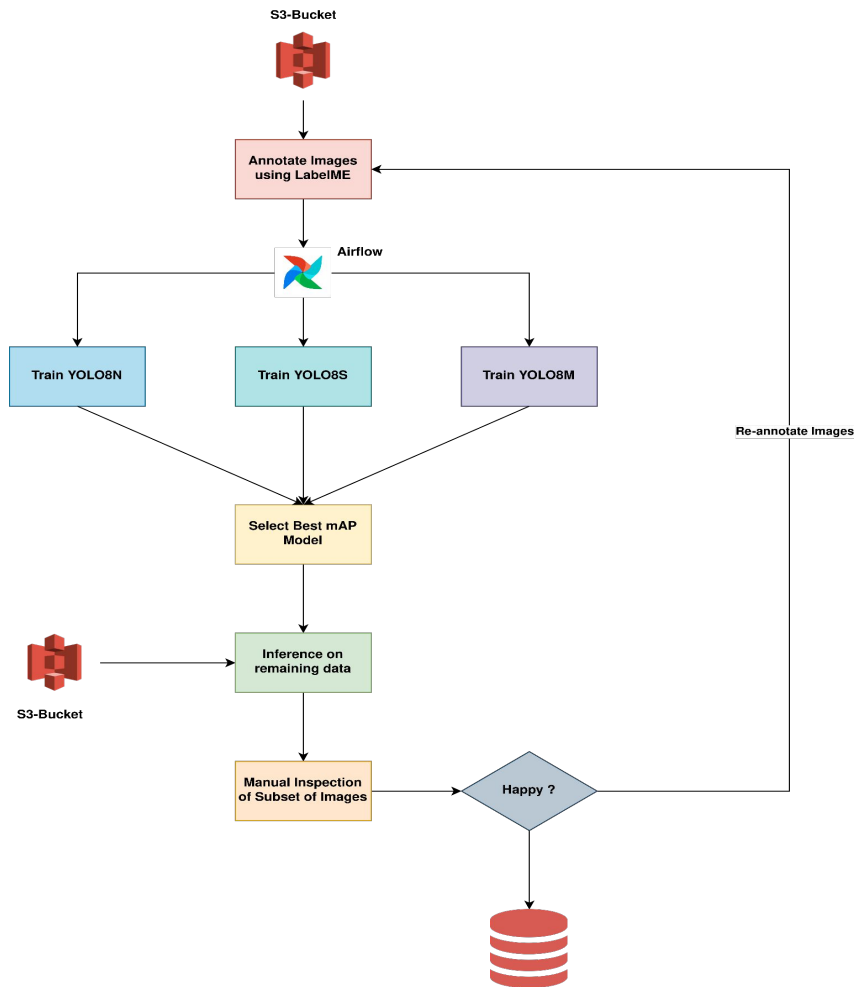
#5475

Ocean Research Education Society





- The maximal amount of images which belong to single class is 72
- **2073 Classes have a single image**
- There is a new category of class called “new_whale” which don’t have a tag yet.
- 9700 images belong to the “new_whale” class



Better Images via the help of YOLOv8

- Manually annotated -> 300 Images from train dataset.
- Trained 3 variants of YOLOv8 models with YOLOv8M performing the best



Tech Stack

- AWS S3/MinIO
- LabelME
- Airflow
- PyTorch/lightly
- Mlflow

Plan of Action

- Week 3 - Week 4: Dataset collection/Data cleaning & EDA
- Week 5 - Week 6: Model/Approach selection
- Week 7 - Week 10: Experimentation and Testing
- Week 11 - Week 12: Report Writing & Results presentation

References

- Jaisakthi, S. M., Palaniappan Mirunalini, and Rutuja Jadhav. "Automatic Whale Matching System using Feature Descriptor." In *CLEF (Working Notes)*. 2017.
- Marcos, Diego, et al. "A Whale's Tail-Finding the Right Whale in an Uncertain World." *International Workshop on Extending Explainable AI Beyond Deep Models and Classifiers*. Springer, Cham, 2022.
- Joly, Alexis, et al. "Unsupervised individual whales identification: spot the difference in the ocean." *CLEF: Conference and Labs of the Evaluation Forum*. No. 1609. 2016.

Thank you !

Any Questions ?

