WhaleMatch

A Hack-for-the-Sea project
September 2018

Rich Bean, Eric and Rhianon Brown,
Ted Greene, Isaac Vandor, Chris Zadra

The Problem

- The "SnotBot" is a drone that follows whales and collects samples from the blowhole

- Whale identification is traditionally done visually using the tail, but not all whales show their tails when surfacing

- The "SnotBot" should be able to identify whales in order to more easily match DNA data with visual imagery
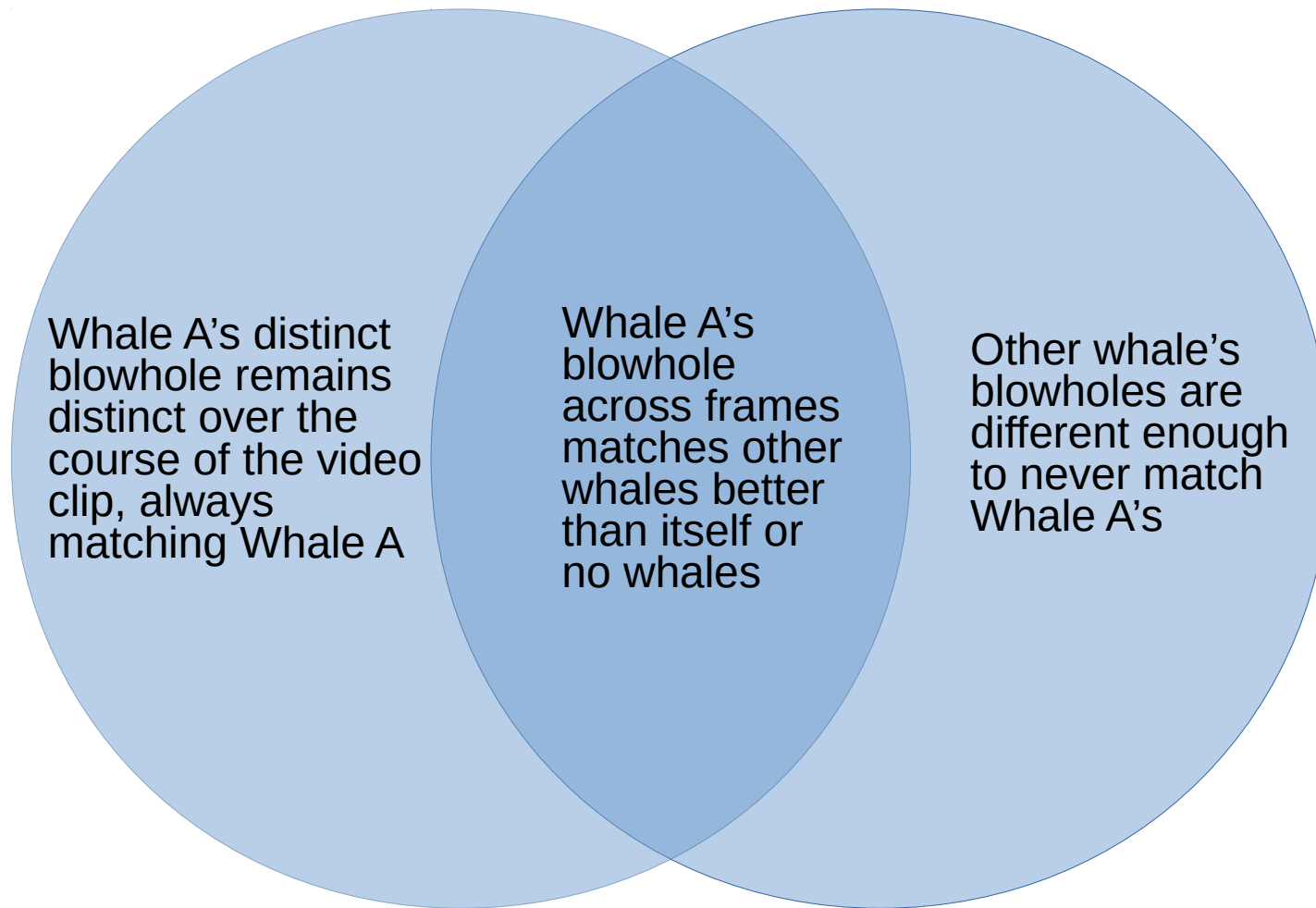
Possible Solution

- Whales have to show their blowholes
- The SnotBot collects video directly above these blowholes
- Is it possible to use this for identification?
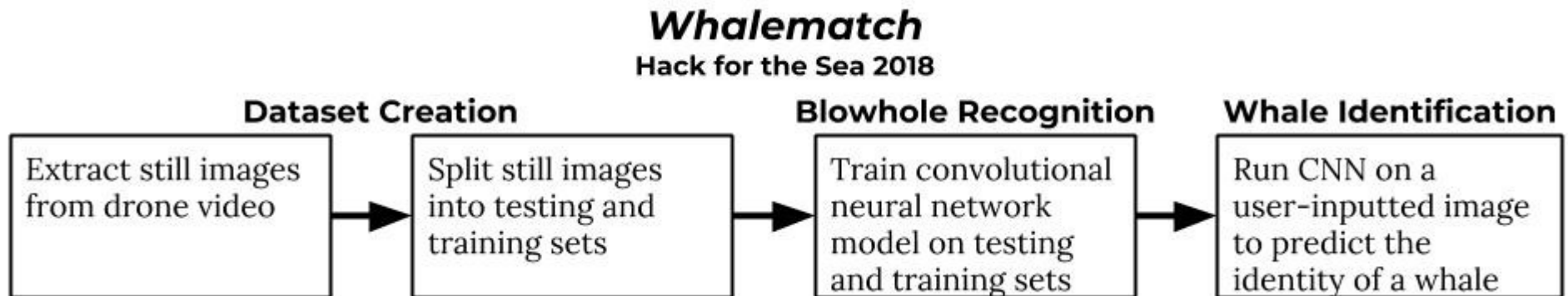 - It's like identifying a person using only nostrils

Challenges

- Whales are moving targets
- Spray from blowholes kills visibility
- Blowholes open and close (to keep water out)
- Skin color very similar to surrounding ocean
- Self and cross-subject variability of blowhole shape
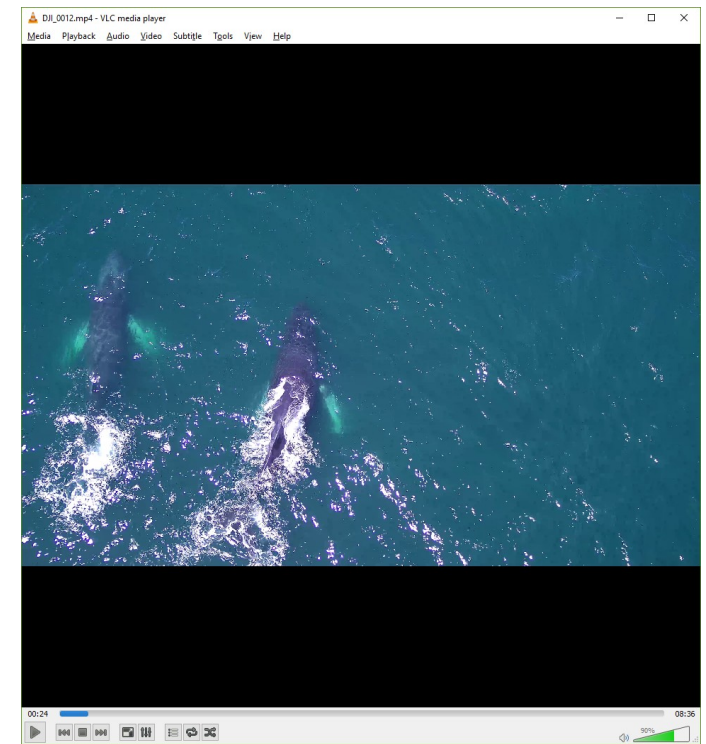
Varying Shape Causes Errors

Whale A's distinct blowhole remains distinct over the course of the video clip, always matching Whale A

Whale A's blowhole across frames matches other whales better than itself or no whales

Other whale's blowholes are different enough to never match Whale A's

Whalematch Block Diagram

Whalematch
Hack for the Sea 2018

Dataset Creation

| Extract still images from drone video | → | Split still images into testing and training sets |

Blowhole Recognition

| → | Train convolutional neural network model on testing and training sets |

Whale Identification

| → | Run CNN on a user-inputted image to predict the identity of a whale |

Video Processing

- 3-8 minute long HD videos
- Extract and convert frames
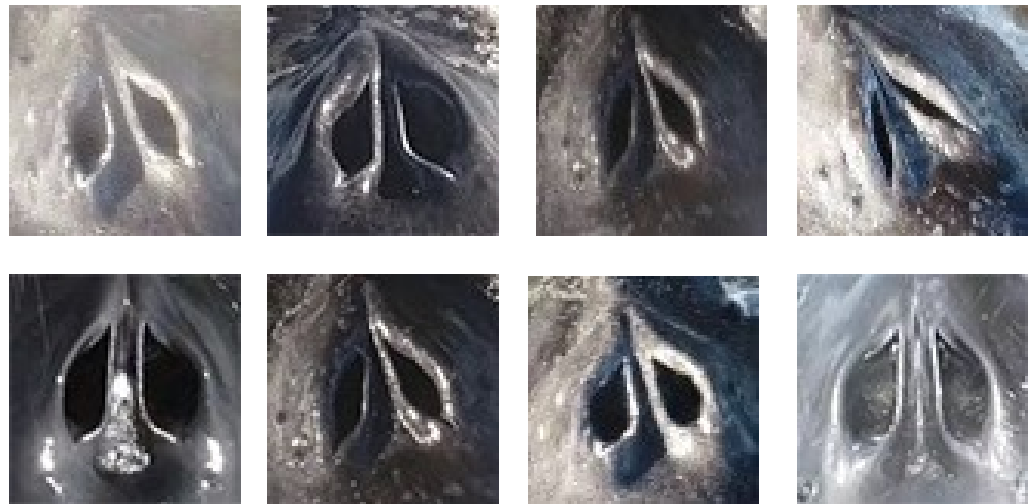

We got around 100K images!

Our (non-representative) Dataset

- Took random 200x200px samples from video stills, resized to 75x75
- Negative sample creation was semi-automated (hand reviewed)
- Blowholes were manually created
- Total Count: 727 negatives, 14 blowholes

Training Data (Negatives)
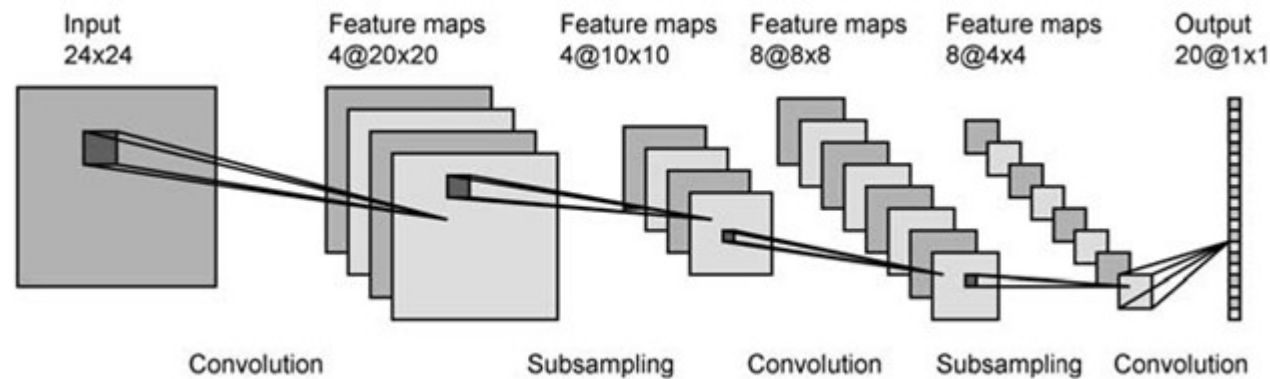

Testing Data (Blow Holes!)

The Machine Learning Approach

- Without doing much pre-processing, can we train a machine learning model to recognize and classify blowholes?

- A similar, but harder, problem to common ML examples like Iris detection, cats vs. dogs detection etc.

- Decided on a convolutional neural network approach to recognize and classify images

Convolutional Neural Networks

- Feature matching across a set of training data to inform recognition of images in the testing data

| Input 24x24 | Feature maps 4@20x20 | Feature maps 4@10x10 | Feature maps 8@8x8 | Feature maps 8@4x4 | Output 20@1x1 |

Convolution · Subsampling · Convolution · Subsampling · Convolution

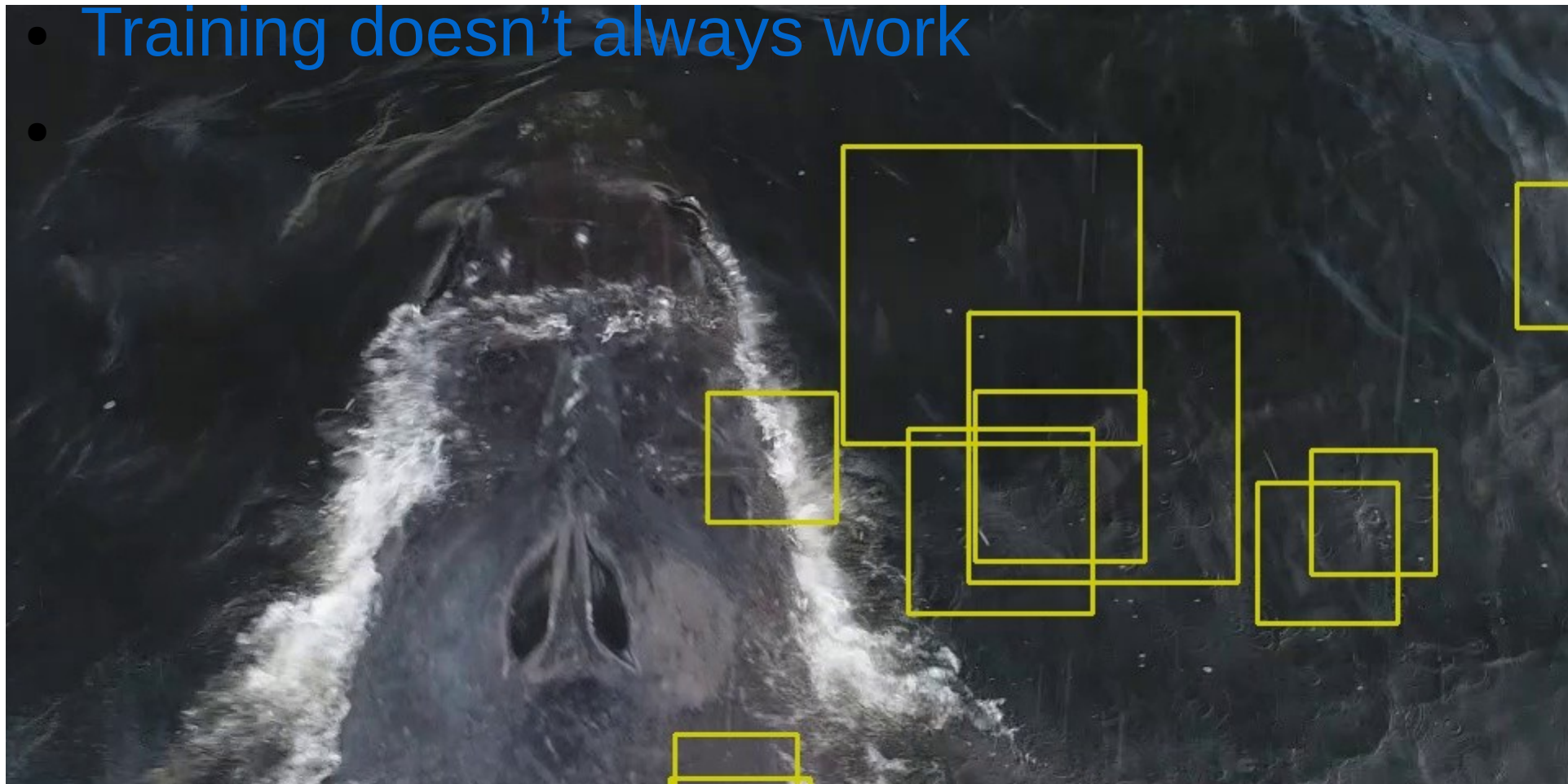- Very useful for image recognition since you can do multiple types of recognition as layers

- We tried three different techniques for detection
  - Two using OpenCV
  - One using Keras
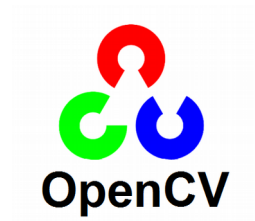
OpenCV False Positives

- Training doesn't always work
-

But Not Unreasonable Either!

- Close ups of the failures help explain them

Keras Results

Whalematch: Example Output
Hack for the Sea 2018

Input Image

Output Blowhole
Classified as 'Whale C'

The Road Ahead

- A larger dataset of blowhole images and non-blowhole images

- Tuning the ML model

- A database mapping blowhole images to whale names

For More info, Check our website/Github:
https://hackforthesea.github.io/whalematch/