Whale Detection

Team members

Juhana Kangaspunta: juhana@stanford.edu
Youssef Ahres: yahres@stanford.edu

Problem and Goals

As we speak, only 500 North Atlantic right whales are left in the world. To ensure the survival of this endangered species, marine biologists are tracking all of them to know their status and health at all times. However, manual recognition is tricky and very few researchers can perform it on the fly.

In this project, we would like to implement a whale recognition system that would allow researchers to reliably identify the whales from aerial photographs.

This project was proposed on Kaggle and the data is readily available and labeled.
The training data consists of 447 individual whale and 4545 aerial photographs that contain only one headshot each. This data comes with notable challenges: on average, an individual whale has only 3 photographs, which might be very small to correctly classify it. Second, as seen below, the whale color can be somehow close to the natural ocean color, which makes it hard to detect or to get its full edges.
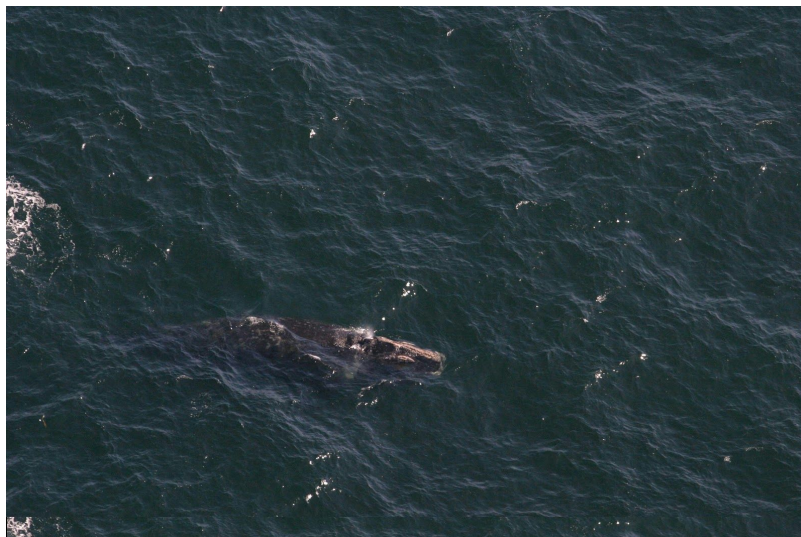
Proposed Methodology

As pointed out in the project presentation, the data consists of aerial photos that contain one headshot in each photo.

We will start by identifying the whale's head in the image. Initially, we will use a basic cascade classifier for detection [1]. Given that the whales may face different directions, we will train 4 different detectors, one for each direction. Every image will be inputed the four detectors, which will output one bounding box each around what it believes is the whale's head. We will then select the detected head that has the highest confidence. If necessary, we will apply more complex approaches for detection such a Histogram of Oriented Gradients (HoG) detector [2]. The previously described detection step outputs a bounding box around the whale's head as well as its direction on the image. The latter information allows us to rotate the image, making all the whales facing the same direction.

Using the detected and rotated whales' heads, we will perform recognition techniques to assign each photo to an individual whale (447 classes). First, we will try eigen / Fisher images and move on to more complex approaches such as using keypoint detection and using keypoints as features or convolutional neural network based approaches.

The main challenge in the project is having 447 classes and only 4545 training images. Because of this limitation, training data augmentation is required.

[1] Viola, Paul, and Michael Jones. "Rapid object detection using a boosted cascade of simple features." *Computer Vision and Pattern Recognition, 2001. CVPR 2001. Proceedings of the 2001 IEEE Computer Society Conference on*. Vol. 1. IEEE, 2001.

[2] Dalal, Navneet, and Bill Triggs. "Histograms of oriented gradients for human detection." *Computer Vision and Pattern Recognition, 2005. CVPR 2005. IEEE Computer Society Conference on*. Vol. 1. IEEE, 2005.