**Expected features for detecting individual whales**

Citation:

North Atlantic Right Whale Catalog. (1997). Retrieved December 27, 2015, from http://rwcatalog.neaq.org

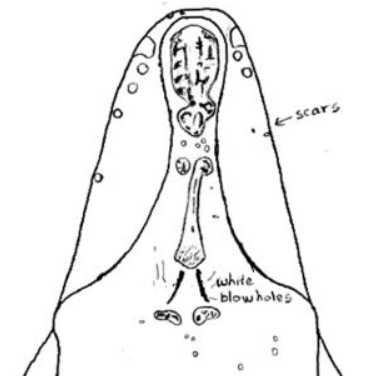


Figure X: Possible facial features of an individual whale specimen

The feature that we are most interested is called the callosity pattern which includes the facial markings on top of the head of the whale and the white markings above the blowholes. These features are unique to each whale.

Some of the features that we decided to ignore were the shape of the tail, dorsal fins and side flippers which may also have been useful, but would have increased the complexity of the detection.

**The dataset**

From the original dataset [cite dataset] of 11469 images only 4542 images were labeled, the labels included 427 unique individuals. Some of the labels only included a single image of the whale. We could not process a dataset this sparse and decided to extract a new dataset – α-whales from the labeled data. Taking only whales which have 20 or more labeled images. This gave us a set of 924 images of 38 unique whales, which is what we based our classifier on.

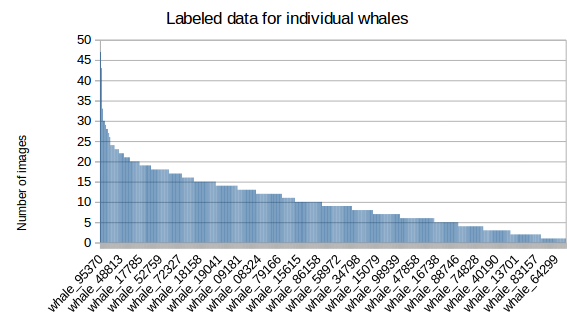


Figure X: Histogram of label distribution in original Kaggle data

**Dataset Preprocessing**

Citation:

Piccardi, Massimo. "Background subtraction techniques: a review." *Systems, man and cybernetics, 2004 IEEE international conference on*. Vol. 4. IEEE, 2004.

The raw images from the dataset were a very large size and resolution. Operating on such images would require massive processing power. In addition much of the visible area of each image was taken up by the water. A large amount of noise with respect to the ROI (region of interest) was added by the waves and splashes around the whale.



Figure X: Image of whale with noise from water

To preprocess the data we had tried to segment the ROI of the whale from the water. We managed to discard the majority of the water pixels by segmenting the histogram of the saturation of the image.

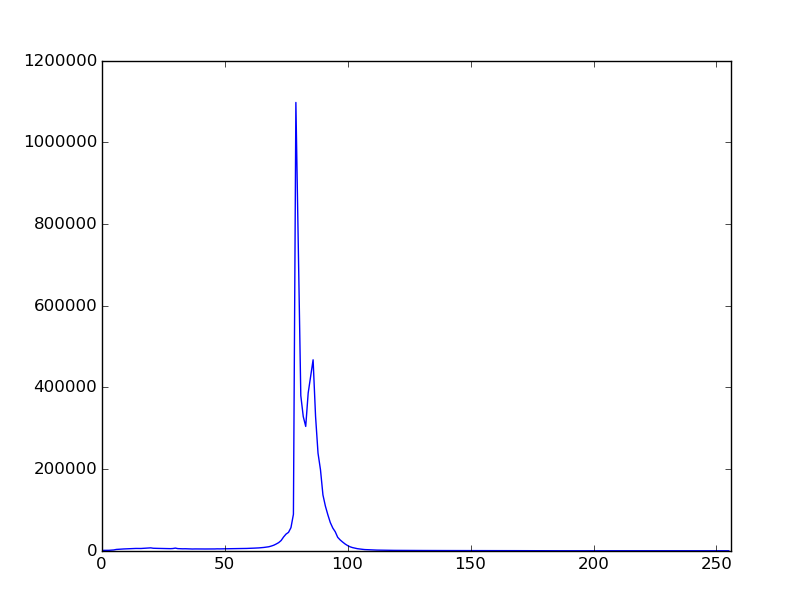
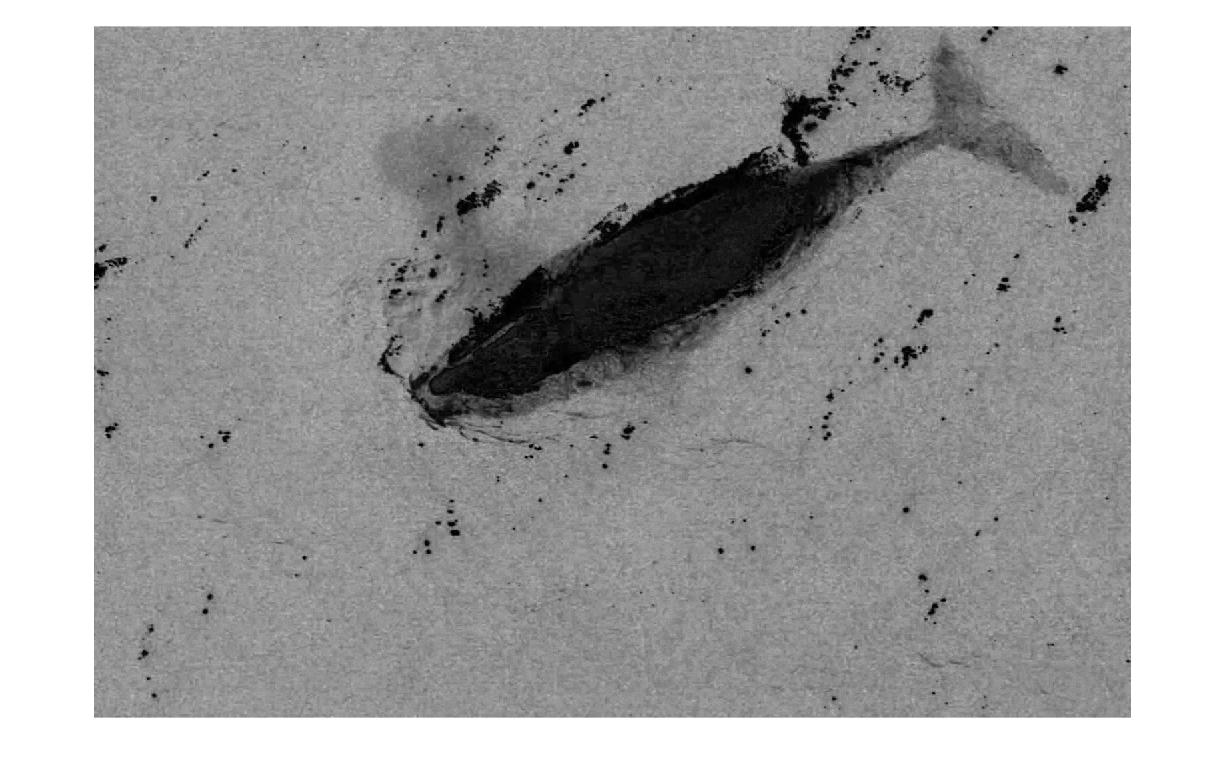


Figure X: Saturation from image and histogram of saturation channel

Ideally the saturation histogram should have two visible peaks (as above). The first and greater maxima is a marker for the water pixels which make up the majority of the image. The second and lesser maxima is therefore expected to be the whale or the foreground of the image. We can threshold the image using the minima that can be found between the two local maxima points – which would leave us with the pixels corresponding to the whale and the surrounding noisy pixels of waves/splashes.

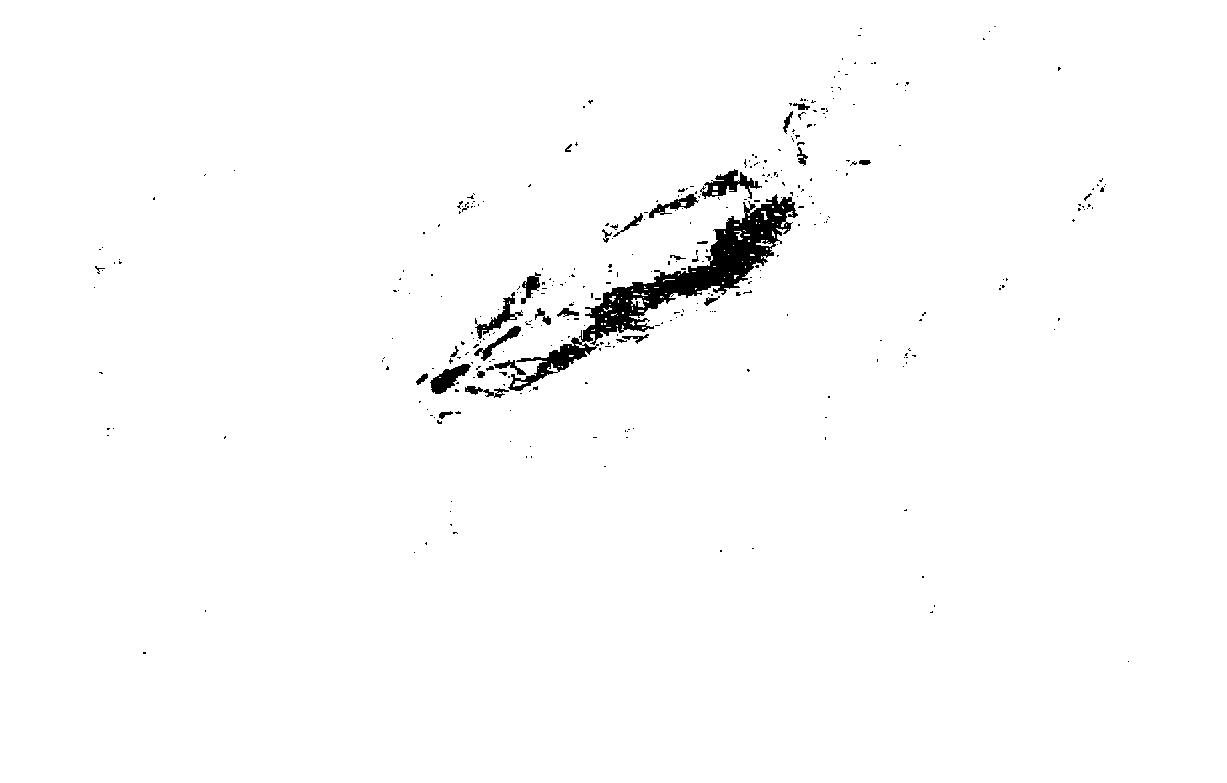


Figure X: Extracted mask of ROI from thresholding the saturation channel

Using this ROI polygon we can fix yet another problem with the data – the fact that the whales all face random directions. A solution to this particular problem is to enscribe an ellipse into the ROI polygon. The major axis of this polygon will roughly coincide with the major axis (head to tail) of the whale. Knowing the angle of rotation we can use an affine transformation to rotate the image to our desired location. The whale will now be facing either up or down.

The entire process works well about 4 out of 5 times. The other times the noise in the water will mess up the histogram which will create an incorrect ROI which in turn will result in a misleading rotation and crop of the image. Even in the successful runs of the preprocessor we were left with the problem of not knowing whether the whale was facing up or down – which mean that we could not further crop out the remainder of the whales body since the face is the most interesting feature.

Since we were not able to reliably preprocess the raw images using conventional Computer Vision techniques we turned to manual preprocessing of what we desired the output to be:

* Rotate the image to have the whale facing right.
* Cut out a square area with the whale dead center.
* Resize image to 256x256

These steps ensured that we obtained the “passport” photos of that the CV preprocessor should have made.

Move to improvements section:

A possible solution we could have implemented would have been to train a simple classifier on this dataset which could have detected the position of the whales head. We could have then used that information to crop out the smallest possible images of the whale heads on the entire dataset.