Whale Identification – Few shot learning

PAUL G. ALLEN SCHOOL OF COMPUTER SCIENCE & ENGINEERING

Jingkun Zhang & Tianyuan Fu | Computer Vision CSE455 Fall 2019 | University of Washington

Problem Identification

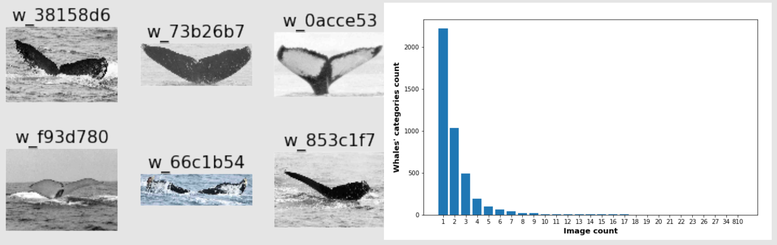
Scientists are using the photo surveillance system to log the whale activities and tail shape in the them to determine the species. Th manual identification is laborious and lacks accuracy. An automated system is needed to replace the process.

Dataset Overview and Challenges

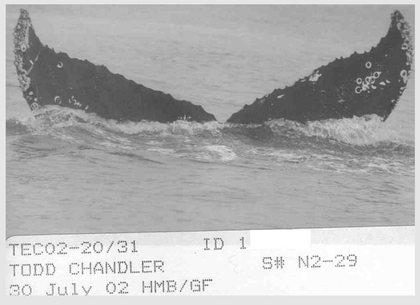
The training data contains 9850 distinct pictures of whale tails, paired with 4251 different kinds of whale species. A sample record is like this:

Challenges:

1. The distribution of picture numbers among whale classes are uneven. For few number of classes, the training data size can be over 2000 while for some there can be only one picture for training. Here is the distribution diagram for all pictures.



2. Some pictures are given along with notations and comments, which may affect our feature extraction.



Proposed Solution

Based on the condition that: 1) the distribution of training data to classes are uneven and the majority of classes only has one training picture; 2) some pictures contains elements other than whale tail details, we propose the usage of 1) Siamese network and 2) Bounding Box Model, to tackle the issues respectively.

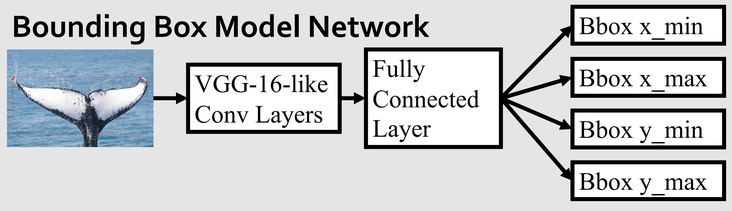
Also we tired classical way like Densely Connected Convolutional Networks (DenseNet) to compare the results.

Technical Details

Data Preprocessing

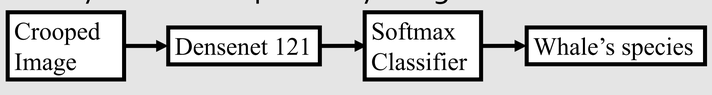
- Exclude 'new whale' class from dataset
- Horizontal flip and Affine

Bounding Box Model Network



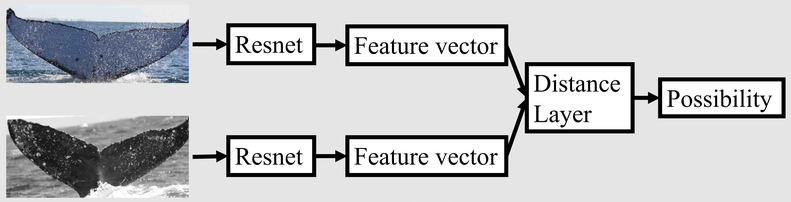
DenseNet Network

We implemented the pretrained densenet121 network to classify the whale species by images



Technical Details (con'd)

Siamese Neural Network(SNN)



We used pretrained resnet50 for feature extraction and L1 for calculate distance in distance layer.

For constructing training data for SNN, firstly we constructed the matching pairs by using the right order of a specific whale category and the derangement order of this category. Then for mismatching pairs, we just simply pick up two different images from sperate categories randomly.

Evaluation Metrics

It is evaluated by the Mean Average Precision @ 5 (MAP@5).Type equation here.

$$MAP@5 = \frac{1}{U}\sum u = 1^{U}\sum k = 1^{min(n,5)}P(k)$$

Results and Conclusion

DenseNet	0.40514
Siamese Neural Network	0.59073

Obviously, using SNN can boost the result greatly for the few-shot learning. For the future work, we will try other data augmentation and different backbone for SNN to achieve better result.

References

1. Huang, Gao, et al. "Densely connected convolutional networks." *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2017.
2. Koch, Gregory, Richard Zemel, and Ruslan Salakhutdinov. "Siamese neural networks for one-shot image recognition." *ICML deep learning workshop*. Vol. 2. 2015.