

# Humpback Whale Identification Challenge

Changyu Li, Jiayu He, Huang Lin, Wenda Chen

## Abstract

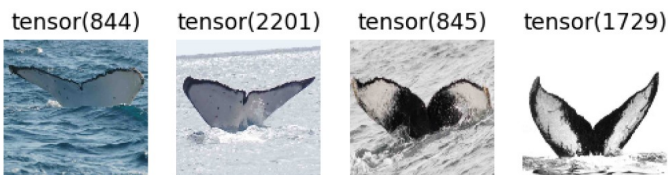
Humpback whale, which takes its common name from the distinctive hump on its back, is a mammal living in oceans around the world. This surface-active species is a favorite to whale watchers. The task of this project is to help scientists identify humpback whale species by its tail image. We got a very challenging dataset. Both VGG16 and ResNet pretrained model as well as the data augmentation optimization attempt, were used in this

## Data

The data set contains training data set and, validation data set and testing data set. This training data contains thousands of images of humpback whale flukes. Individual whales have been identified by researchers and given an ID.

- 9850 images with over 4000 classes as training set
- 3000 images as validation set

The purpose is to predict the 5 possible whales ID of these whales image. If any one predicted whales ID correspond to the correct whales ID, then we mark this as a successful prediction.



## Methods

### 1. Data Processing

Data Augmentation:



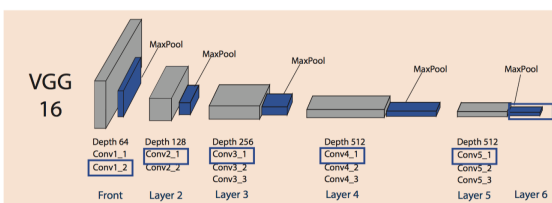
Data Augmentation

RGB -> Black & White

### 2. convolutional neural network

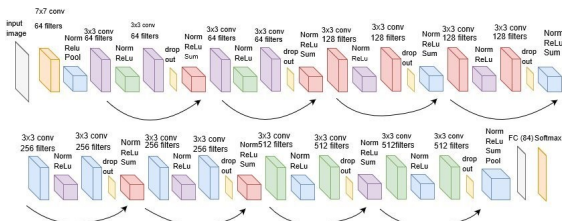
- VGG 16 Model

VGG is a convolutional neural network model proposed by K. Simonyan and A. Zisserman. It increases depth using an architecture with very small (3x3) convolution filters.



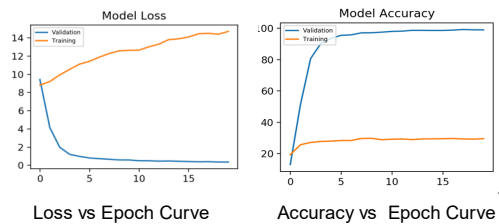
- ResNet 18 Model

ResNet means deep residual network. The residual learning framework eases the training of deep networks. ResNet is much deeper than their plain counterparts, yet it requires a similar number of parameters.



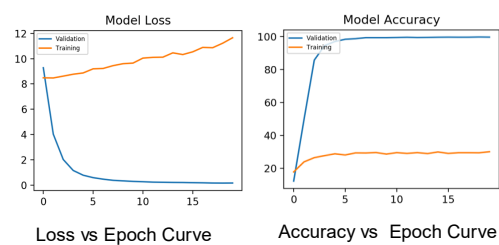
## Result

- VGG model



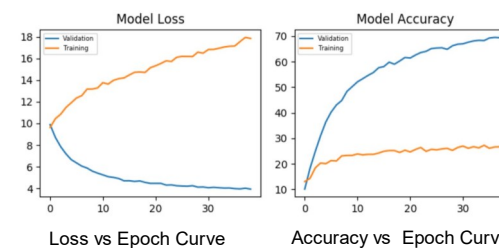
Optimizer: Adam  
Mini-batch size: 16  
Epochs: 20  
Accuracy: 29.52%

- Resnet Model

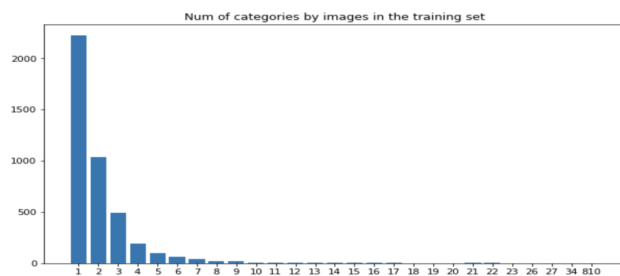


Optimizer: Adam  
Mini-batch size: 16  
Epochs: 20  
Accuracy: 29.97%

- Data Augmentation Attempt



Optimizer: Adam  
Mini-batch size: 16  
Epochs: 40  
Accuracy: 27.43%



As we can see, the vast majority of classes only have a single image in them. This is going to make predictions very difficult for most conventional image classification models.

## Conclusion

VGG16 and ResNet approach showed similar results although ResNet improves a little. Adding data augmentation and gray-scaling all the images didn't help increase prediction accuracy. The model reached a high accuracy for training set but low for validation set. The system shows signs of overfits to the training set because the sample distribution is unbalanced containing 2220 one-sample classes and one class with 880 samples. The model only learned distribution of training set but not validation set. Also, there are many classes in validation set that never appeared in training set thus prediction in validation set is very difficult to be accurate.

## Discussion

By comparing the result in the results section, after data augmentation the didn't perform better than the previous model. The reason is unknown. In addition, the evaluation loss is keep increasing since the training begin. Given more time we would also try to do more data preprocessing and see if we can improve the prediction accuracy.

## References

Eigen-Distortions of Hierarchical Representations, <http://www.cns.nyu.edu/~lcv/eigendistortions/ModelsVGG.html> ; Proposed Modified ResNet-18 architecture for Bangla HCR ,[https://www.researchgate.net/figure/Proposed-Modified-ResNet-18-architecture-for-Bangla-HCR-In-the-diagram-conv-stands-for\\_fig1\\_323063171](https://www.researchgate.net/figure/Proposed-Modified-ResNet-18-architecture-for-Bangla-HCR-In-the-diagram-conv-stands-for_fig1_323063171) ; Data Augmentation using Fastai ,Amrit Virdee,May 3, 2018,<https://becominghuman.ai/data-augmentation-using-fastai-aefa88ca03f1> ; <https://www.mathworks.com/matlabcentral/mlc-downloads/downloads/submissions/9107/versions/1/screenshot.jpg>