

Fish species identification using a convolutional neural network and simulation-based learning

Vaneeda Allken, Tiffanie Schreyeck, Thomas Mahiout, Shale Rosen, Nils-Olav Handegard, Ketil Malde
Institute of Marine Research, P.O box 1870 Nordnes, NO-5817 Bergen

May 2017: Norwegian Spring spawning herring survey (Deep Vision data)

Aim: Automatic detection of blue whiting, mackerel or herring using DNN

Challenge: Limited number of clean labelled images

Idea: Train on simulated data

Questions to explore:

1. Can a deep neural network learn from simulated data?
2. How does performance scale with size of (simulated) dataset?



Simulated training data

Number of cropped fish images per species, C



Example of background

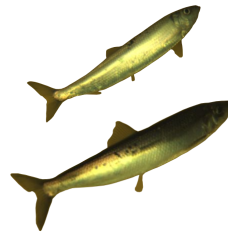
+



=



Blue whiting



=



Herring



=



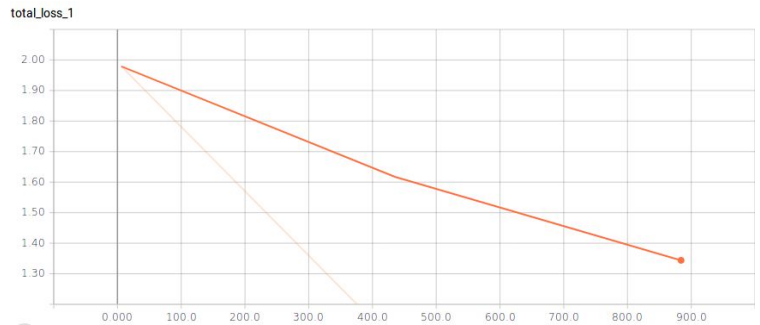
Mackerel



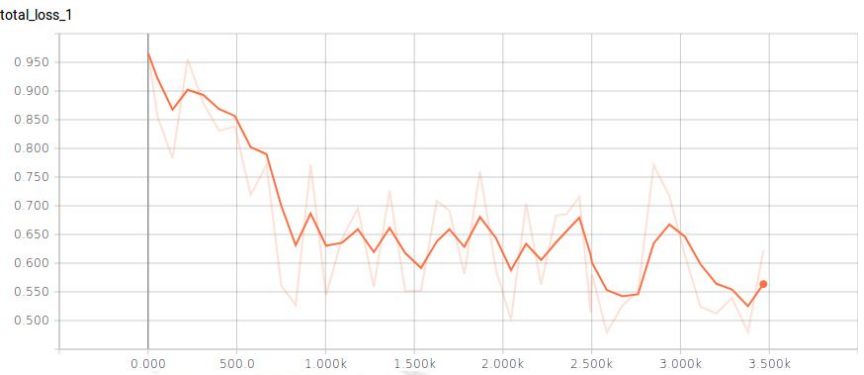
We use a TensorFlow implementation of Inception-v3, a state-of-the-art convolutional neural network model, pretrained on the ImageNet classification dataset

The training set consists of N simulated images built from C number of cropped fish per species

- Example of training loss while
 - retraining only last layer



- Finetuning the whole network



Testing on real images

- Example of correctly and incorrectly classified images

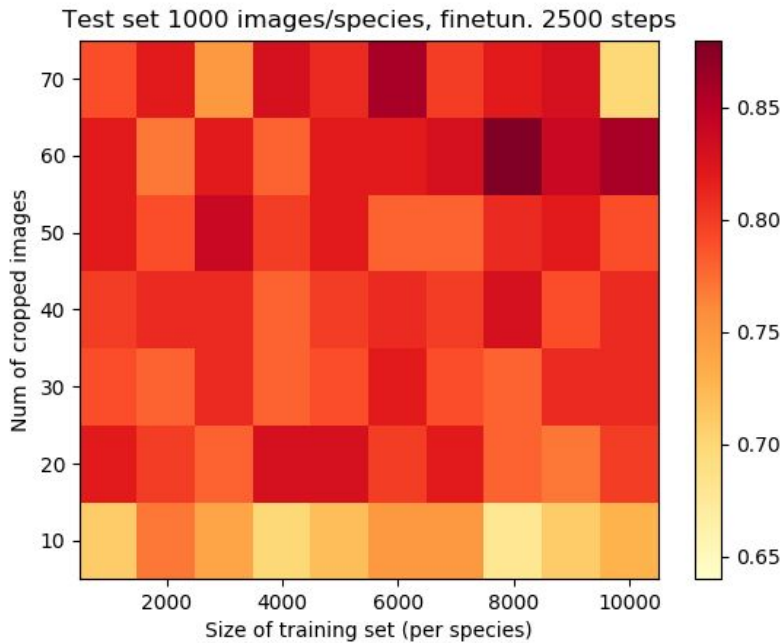
	Predicted as		
	Blue whiting	Herring	Mackerel
(Actual) Blue whiting			
(Actual) Herring			
(Actual) Mackerel			

We vary number of cropped images, C and number of simulated images, N

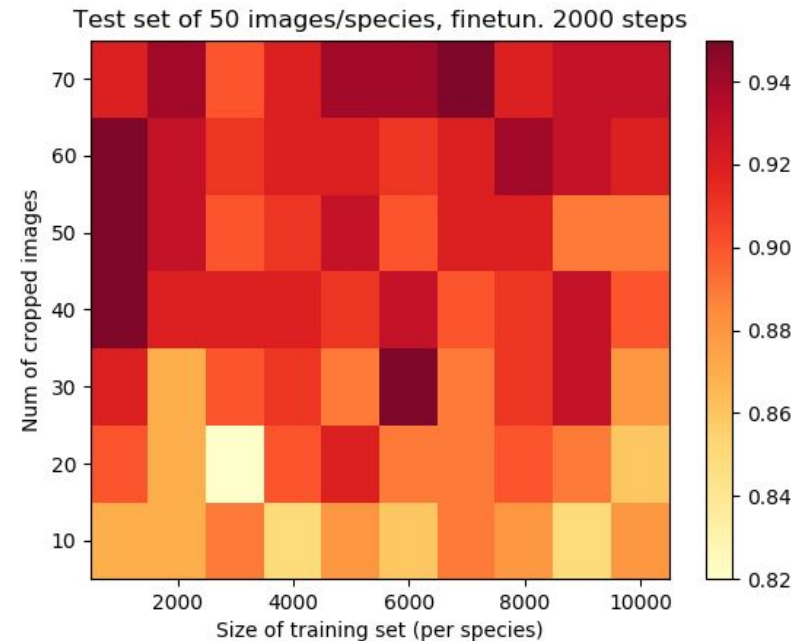
- Best accuracy on
 - Dataset 1: 87% (C = 60, N = 8000)
 - Dataset 2: 95% (C = 70, N = 7000)



Accuracy for varying C and N



Best accuracy: **87%**
for $C = 60$ and $N = 8000$



Best accuracy: **95%**
for $C = 70$ and $N = 7000$



HAVFORSKNINGSINSTITUTTET