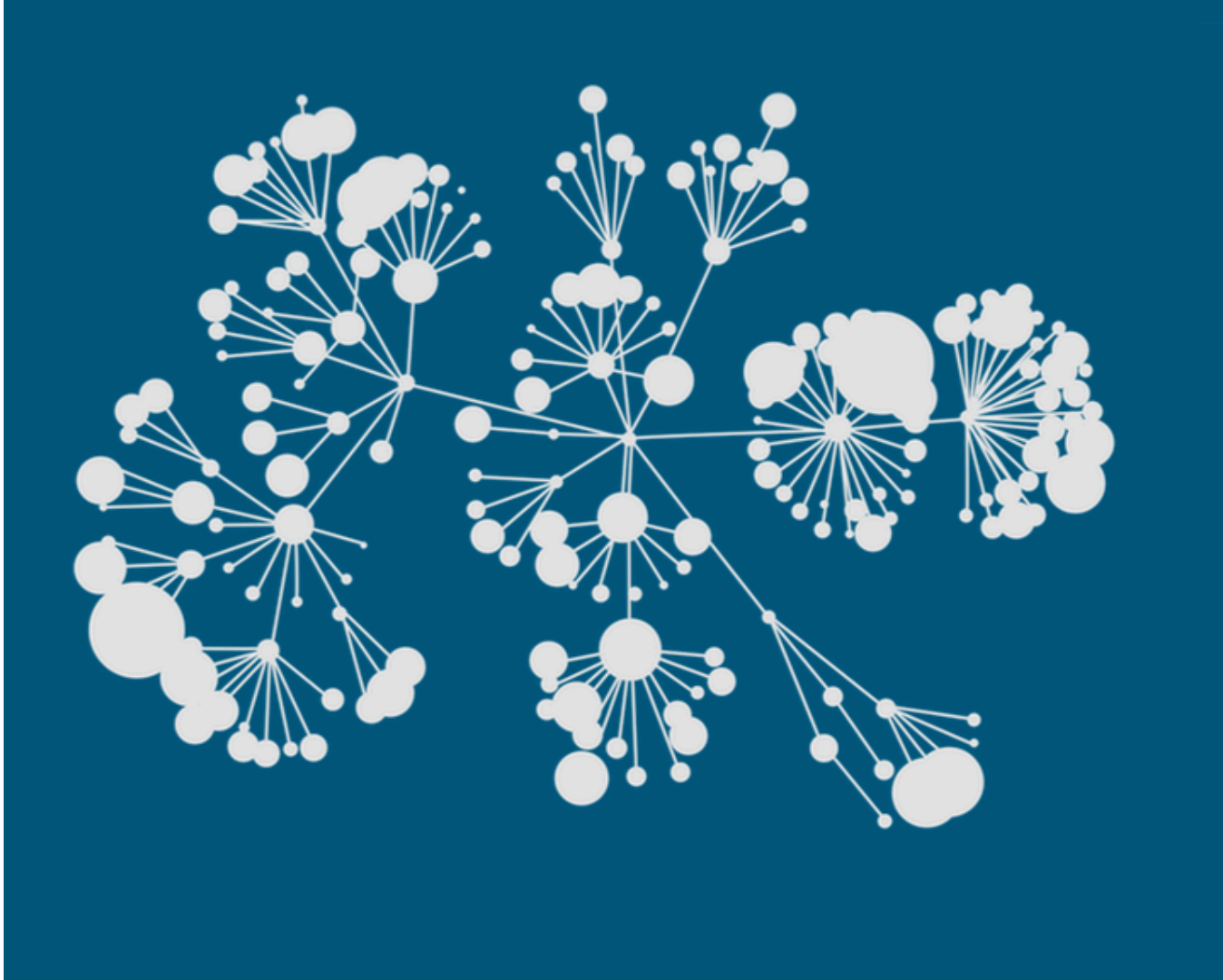
Kaggle

The Nature Conservancy Fisheries Monitoring

3rd prize Winner Presentation

Frank Zeng
5/17/2017

kaggle



Agenda

1. Background
2. Summary
3. Feature selection & engineering
4. Training methods
5. Important findings
6. Simple model

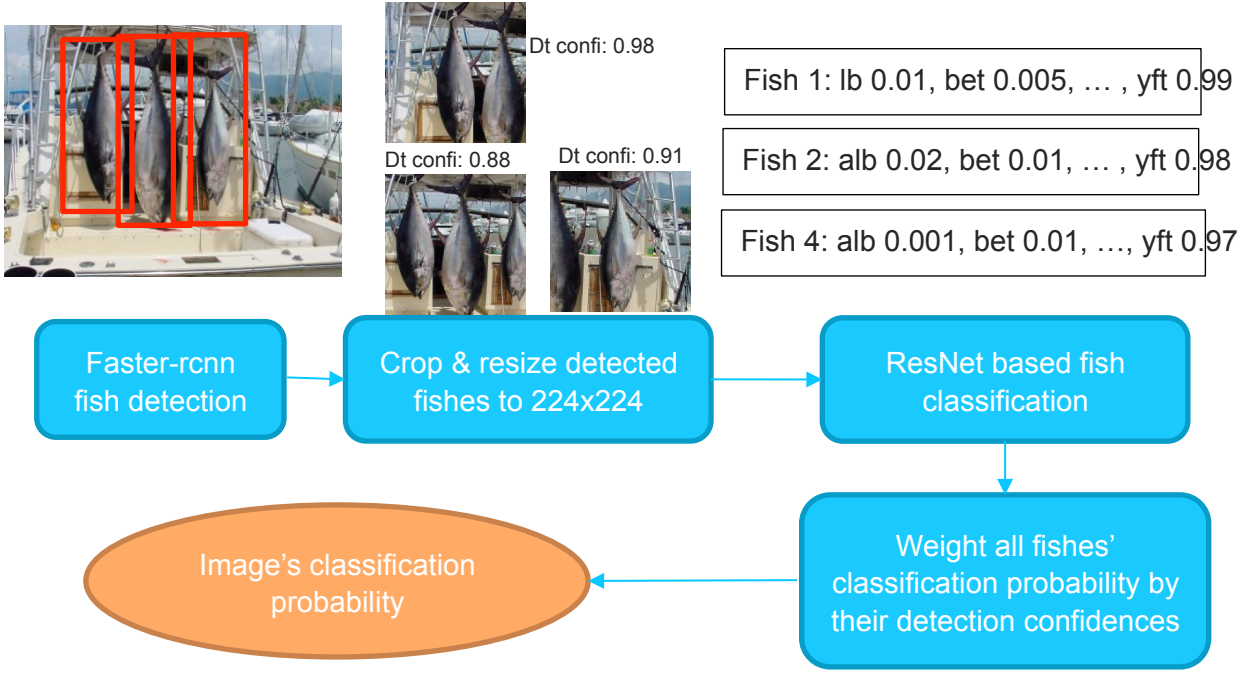
Background

- have a PhD degree in Information and Signal Processing
- have been in image processing/computer vision industry for 10 years
- 1.5 years’ experience in deep learning based detection/classification

Summary

- deep CNN based fish detection and classification
- All features are learned and embedded with deep convolution network
- Python + Caffe
- Detection: Faster Rcnn (Region cnn) with ResNet50 model
- Classification: ReNet50 + ResNet101

Features Selection / Engineering



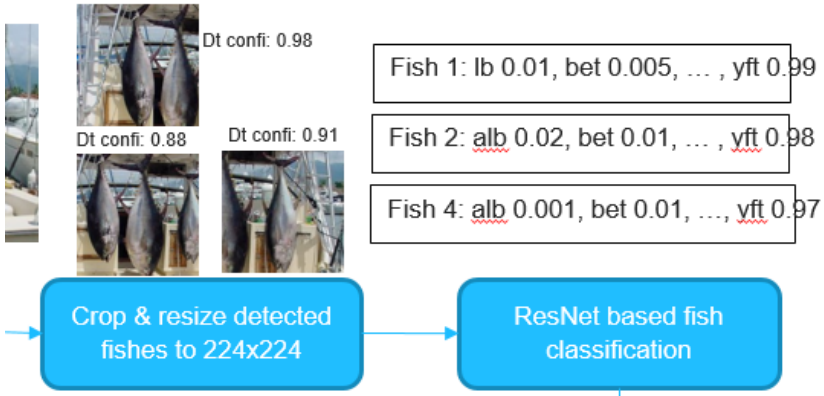
The main detection & classification flowchart

Features Selection / Engineering

- faster-rcnn is the state-of-the-art object detection method
- can use different pre-trained model as backend (e.g. VGG16, ResNet50 or ResNet101)
- Use ResNet50 as it has good accuracy and moderate model complexity

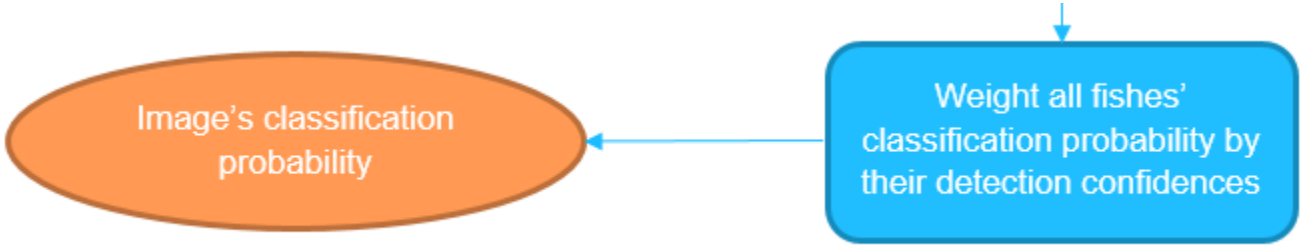
Faster-rcnn fish detection

Features Selection / Engineering



- Bounding box which has a detection score greater than some threshold is fed into 8 classes fish classifier (including NoF)
- detected bounding box is cropped with the same width & height by adding some context, and then resize it to 224x224
- Context is important as sometimes the detector can miss some part of the fish, such as the fin.
- ResNet50 and ResNet101 are used for fish classification.

Features Selection / Engineering



- one image may contain multiple fishes.
- give the image's final probability by combining each fish's classification probability with weighting.
- weighting is based on the detection score.

Fish detector training

- Training label preparation. Label fish bounding boxes in training images from kaggle forum and do some cleanup.
- Init from Pre-trained model. I use ResNet50 model trained on imagenet.
- Data augmentation. Left right mirror and top down mirror to augment training set to 4X.

Training Methods

Fish classifier training

- Training label preparation. Manually label all fishes in each image into its true class. For unsure type, don't use in training.
- Init from Pre-trained model. I use ResNet50 and ResNet101 model trained on imagenet.
- Data augmentation. random rotation / shift, Gaussian smoothing, crop-paste
- Additional training data. 444 cropped fishes from kaggle forum.
- Early stopping. Split into training & validation set by nearest neighbor clustering.

Training Methods

Model ensemble

- Faster rcnn based fish detector is good enough, so I only use one single model.
- trained 10 ResNet50 models and 10 ResNet101 models for average based ensemble by split training and validation dataset in 10 folds.
- inter-image ensemble.
    - similar images and fishes should have similar classification result
    - average these neighboring images' classification result weighted by similarity scores with current image's classification result as final result.
    - Image similarity scores is calculated in two stages, i.e. image-level and fish level.

Important and Interesting Findings

- There is no single tricks that works most importantly. The accuracy is improved step by step.
- A fish detection and classification framework is the start step of a successful method.
- Data augmentation is very important as training sample size is small.
- Be cautious when doing augmentation as not all augmentation works well

Simple Model

- remove model ensemble. Remove inter-image ensemble and use only 1 classification model instead of 20. accuracy will suffer. stg2 score drops from 1.13805 to 1.32678 (rank drops from 3 to 12) without model ensemble.
- add more training samples for training detector (at least 20,000) and classifier (at least several thousand for each fish type).
- Use simpler and less computation-intensive model as backend model, both for detection and classification. (e.g. change from ResNet50 to ResNet34 or even SqueezeNet).

Q & A

kaggle™