Machine Learning Engineer Nanodegree

Capstone Proposal

Kishan Kanhaiya
December 08, 2018

Proposal

Domain Background

Nearly half of the world depends on seafood for their main source of protein. In the Western and Central Pacific, where 60% of the world's tuna is caught, illegal, unreported, and unregulated fishing practices are threatening marine ecosystems, global seafood supplies and local livelihoods. The Nature Conservancy is working with local, regional and global partners to preserve this fishery for the future.

Currently, the Conservancy is looking to the future by using cameras to dramatically scale the monitoring of fishing activities to fill critical science and compliance monitoring data gaps. Although these electronic monitoring systems work well and are ready for wider deployment, the amount of raw data produced is cumbersome and expensive to process manually.

The Conservancy started a competition on the Kaggle community to develop algorithms to automatically detect and classify species of tunas, sharks and more that fishing boats catch, which will accelerate the video review process. Faster review and more reliable data will enable countries to reallocate human capital to management and enforcement activities which will have a positive impact on conservation and our planet.

Problem Statement

We have to detect which species of fish appeared on the boat based on images that has been captured from the boat cameras of various angles.

Our goal is to predict the likelihood of the fish species in each picture. Eight target categories are available in the dataset: Albacore tuna, Bigeye tuna, Yellowfin tuna, Mahi Mahi, Opah, Sharks, Other (meaning that there are fish present but not in the above categories) and No Fish (meaning that no fish is in the picture). Each image has only one fish category, expect that there are sometimes very small fish in the image that are used as a bait.

Datasets and Inputs

For this competition we have around 4000 images for training dataset and 1000 images for testing dataset, captured from boat cameras of various angles.

Eight Target Categories are available in the dataset

1. Albacore tuna
2. Bigeye tuna
3. Yellowfin tuna
4. Mahi Mahi
5. Opah
6. Sharks
7. Other (meaning that there are fish present but not in the above categories)
8. No Fish (meaning that no fish is in the picture)

The dataset was compiled by The Nature Conservancy in partnership with Satlink, Archipelago Marine Research, the Pacific Community, the Solomon Islands Ministry of Fisheries and Marine Resources, the Australia Fisheries Management Authority, and the governments of New Caledonia and Palau.

Solution Statement

In this project we will use deep learning approach to tackle the problem of classifying the fish in a given image. For this problem, we will train a Convolutional Neural Network using transfer leaning. Transfer learning is a process that focuses on storing knowledge gained while solving one problem and applying it to a different but related/similar problem.
There are some pretrained neural networks on ImageNet dataset such as Inception-V3, Resnet-50, VGG-16, VGG-19 are available, whose weights can be used for developing the solution for this classification problem.

Benchmark Model

For the Benchmark, we will use a Convolutional Neural Network using Conv2D layers with the increasing numbers of filters to progressively detect more complex patterns. We will also insert MaxPooling and Dropout layers in between Conv2D layers to reduce the input dimensions and finally the fully connected Dense layers. We will train this model on the training dataset and evaluate the predictions made on the test dataset using the log-loss function.

Evaluation Metrics

The Evaluation metrices used for this classification problem is multi-class logarithmic loss.

$$log\text{-}loss = -\frac{1}{N}\sum_{i=1}^{N}\sum_{j=1}^{M} y_{ij}\, log(p_{ij})$$

Where N is the number of images in the test set, M is the number of class labels of image, log is the natural logarithm, y_{ij} is 1 if the observation i belongs to class j and 0 otherwise, p_{ij} is the predicted probability that observation i belongs to class j.

The probabilities for a given image is not required to sum to one as they are scaled prior to being scored, each row is divided by the row sum. In order to avoid the extremes of the log function, predicted probabilities are replaced by $\max(\min(p, 1-10^{-15}), 10^{-15})$.

Project Design

1. Import the required datasets for analysis and classification of the fish.
2. Explore data via Visualization such as distribution of various categories of fishes, color distribution of various images etc.
3. Split the dataset into training, testing and validation set. And then we will do some required preprocessing like converting the images into tensors, normalization etc.
4. We will create a benchmark model using a mixture of Conv2D, MaxPooling2D, Dropout and Dense layers.
5. Train the benchmark model on training set and then evaluate its prediction using log-loss.
6. We will create one or more models (as computationally possible) with transfer learning approach using the pre-trained models like Inception-V3, RESNET-50, VGG-16, VGG-19. We will extract the features from the images using pre-trained networks and then use a fully connected dense layers with 8 nodes in the end, as there are 8 categories of the fishes.
7. We will try to tune the model using different optimizers and adjusting the various parameters.
8. Finally, we will compare our results with the result of the benchmark model and observe how our model have performed.

References:
https://www.kaggle.com/c/the-nature-conservancy-fisheries-monitoring
https://www.conserveca.org/
https://www.nature.org/en-us/
https://satlink.es/en/
https://www.archipelago.ca/
https://www.spc.int/
https://fisheries.gov.sb/
https://www.afma.gov.au/
https://gouv.nc/
https://palaugov.pw/executive-branch/ministries/natural-resources/
https://en.wikipedia.org/wiki/Transfer_learning
https://in.udacity.com/