Image Classification Project 3

Train a neural net that is able to classify images from the Open Images Dataset into three categories: **Alpaca, Camel, Dog**

Dataset

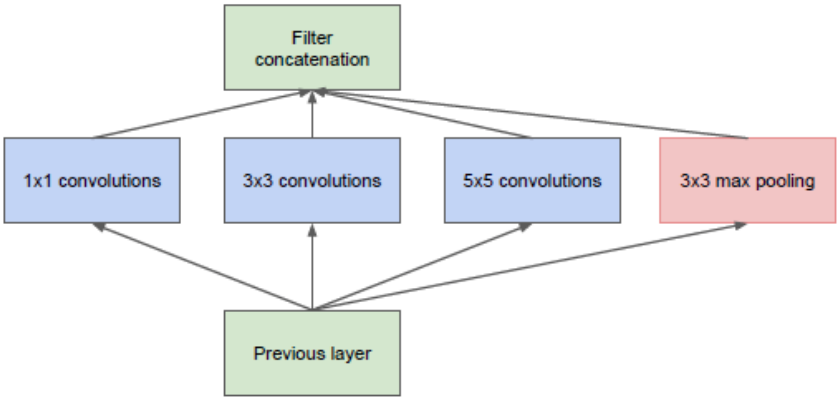
https://storage.googleapis.com/openimages/web/visualizer/index.html?type=detection

Base model

VGG 19

Tasks

- Preparation: Split dataset into a **80/20** Train/test split
- **Experiment transfer learning**: Use a imagenet pretrained **VGG19** architecture, train the model and estimate the testset accuracy
- **Experiment data augmentation:** Add data augmentation and train again, discuss results
    - Random flip
    - Random contrast
    - Random translation
- **Experiment with architecture:** Rebuild VGG19. After layer block4_conv4 (32, 32, 512):
    - Add a naive inception layer (output filter size should be 512, each padding same, activations leaky relu)



    - Add conv layer (kernel 3x3, filters 512, padding valid, stride 2, activation relu)
    - Add conv layer (kernel 1x1, filters 640, padding valid, stride 1, activation relu)
    - Freeze conv2 layers and before
- Test a few of your own images from the internet and show the **activation maps** (how did the neural net come to its conclusion)

- Answer the following questions:
    - What accuracy can be achieved? What is the accuracy of the train vs. test set?
    - On what infrastructure did you train it? What is the inference time?
    - What are the number of parameters of the model?
    - Which categories are most likely to be confused by the algorithm? Show results in a confusion matrix.

Compare the results of the experiments.