ImageNet

Label images using a CNN trained on the ImageNet dataset.

Chapter Goals:

- Learn about the ImageNet dataset

A. ImageNet dataset

The dataset used in the ImageNet Challenge contains 1.4 million images, with 1.2M in the training set, 50,000 in the validation set, and the remaining 150,000 in the (unreleased) test set. The images are distributed across 1,000 categories, which are listed here.

B. Image preprocessing

We first perform image preprocessing and data augmentation. However, while the CIFAR-10 images were all 32x32, the ImageNet images have varying heights and widths.

In order to have some consistency with image sizes, we resize each image such that its new minimum dimension has size equal to some fixed value (`min_aspect_dim`). When resizing, we maintain the image's aspect ratio, i.e. the ratio between its height and width dimensions.

C. Other models

Since the AlexNet model broke the mold in 2012, the ImageNet Challenge has been dominated by deep CNNs. The model that won in 2014 (the year prior to ResNet) was *GoogLeNet*, developed by a team of Google researchers. GoogLeNet is also known as the Inception model due to its *inception module*, which is pretty similar to the fire module of SqueezeNet.