

Transfer Learning

Mihir Patel

December 2017

1 Introduction

We have learned a lot about image classification and using convolutional neural networks for a host of problems. However, we also know that training all of these structures takes a lot of time and computational power we don't have. So how can we utilize these structures?

2 ImageNet

As we've mentioned before, most of these structures are trained and tested on ImageNet. It turns out that this is very, very useful. Since ImageNet contains a very large variety of classes and labels, it can be hypothesized that the features it learns at the very lowest of levels are simple, generalizable features that are not problem specific. It turns out that this can be empirically demonstrated as if these networks are trained to learn additional classes, they converge much faster than random weight initialization.

3 Transferring Knowledge

How is this useful? In humans, if we see a new object we try to describe it with certain terms we already know, using things such as smooth, red, etc. which are all features we know. In the same regard, neural networks can transfer over generic image-based knowledge to new patterns. By using the weights that other organizations have trained and obtained on ImageNet, we can have a baseline that already gives us the core image analysis power needed. This knowledge can easily be exploited by us! In fact, many modern machine learning libraries such as keras already have several common models such as VGG and ResNet built in for transfer learning.

4 Applying Our Data

In general, the process for transfer learning is pretty straightforward. We take the weights, and instead of randomly initializing, we use the pretrained weights

and begin from there. We then run as normal. However, often this is either too slow or we don't have enough data. In that case, we do something called freezing. When layers are "frozen", their weights won't change throughout backpropagation. A common approach is to freeze the convolutional layers, which are image specific, and add on a new top portion consisting of fully connected layers that we train. This way, we can use transfer learning as a method to go from image to some extracted features we can train on.