

Homework 6

Submission: This homework is not to be handed in for a mark. However, it is not optional, in the sense that you are still responsible for the content. We will release solutions, but you will get more out of it if you attempt the problems yourself first.

For this assignment, you will first read the following paper:

A. Krizhevsky, I. Sutskever, and G. E. Hinton. ImageNet classification with deep convolutional neural networks. In *Advances in Neural Information Processing Systems (NIPS)*, 2012.

<http://papers.nips.cc/paper/4824-imagenet-classification-with-deep-convolutional-neural-networks>

This is a highly influential paper (nearly 10,000 citations on Google Scholar!) because it was one of the first papers to demonstrate impressive performance for a neural network on a modern computer vision benchmark. It generated lots of excitement both in academia and in the tech industry. The architecture presented in this paper widely used today, and is known as “AlexNet”, after the first author. Reading this paper will also help you review a lot of the important concepts from this class.

1. They use a conv net architecture which has five convolution layers and three fully connected layers (one of which is the output layer). Your job is to count the number of units, the number of weights, and the number of connections in each layer. I.e., you should complete the following table:

	# Units	# Weights	# Connections
Convolution Layer 1			
Convolution Layer 2			
Convolution Layer 3			
Convolution Layer 4			
Convolution Layer 5			
Fully Connected Layer 1			
Fully Connected Layer 2			
Output Layer			

You can ignore the pooling layers when doing these calculations, i.e. you don’t need to consider the units in the pooling layers or the connections between convolution and pooling layers. You can also ignore the biases. Note that the paper gives you the answers for the numbers of units in the caption to Figure 2. You should work out the values by yourself, though.

2. Now suppose you’re working at a software company and want to use an architecture similar to AlexNet in a product. Your project manager gives you some additional instructions; for each of the following scenarios, based on your answers to Part 1, suggest a change to the architecture which will help achieve the desired objective. I.e., modify the sizes of one or more layers. (These scenarios are independent.)
 - (a) You want to reduce the memory usage at test time so that the network can be run on a cell phone; this requires reducing the number of parameters for the network.
 - (b) Your network will need to make very rapid predictions at test time. You want to reduce the number of connections, since there is approximately one add-multiply operation per connection.