CS444 Convolutional Neural Network Exercises

1. CONVOLUTIONAL NEURAL NETWORKS

Consider a convolutional neural network that is used to classify images into two classes. The structure of the network is as follows:

- INPUT: 100x100 grayscale images.
- LAYER 1: Convolutional layer with 100 5x5 convolutional filters.
- LAYER 2: Convolutional layer with 100 5x5 convolutional filters.
- LAYER 3: A max pooling layer that down-samples Layer 2 by a factor of 4 (from 100x100 $\rightarrow$ 50x50)
- LAYER 4: Dense layer with 100 units
- LAYER 5: Dense layer with 100 units
- LAYER 6: Single output unit

(a) How many weights does this network have?

(b) What is "stochastic" about stochastic gradient descent?

(c) Assume that this network is trained using stochastic gradient descent on a fixed training set. Is there a guarantee that the algorithm will converge to an optimal set of weights given enough epochs of training?

(d) As with any learning algorithm, designers of deep neural networks face a bias/variance trade-off. Given your answer to part a), would you say that high bias or high variance is a bigger concern? How might we address the concern?