

## CS 7150: Deep Learning — Summer-Full 2020 — Paul Hand

Week 4 — Preparation Questions For Class

Due: Monday June 1, 2020 at 12:00 PM Eastern time via [Gradescope](#)

Name: [Put Your Name Here]

Collaborators: [Put Your Collaborators Here]

You may consult any and all resources in answering the questions. Your goal is to have answers that are ready to be shared with the class (or on a hypothetical job interview) as written. **Make sure to tag each question when you submit to Gradescope.**

**Directions:** Read the articles ‘[ImageNet Classification with Deep Convolutional Neural Networks](#)’ (AlexNet) and ‘[Deep Residual Learning for Image Recognition](#)’ (ResNets) .

**Question 1.** *What is dropout? What evidence is there that it works? Why does it work?*

**Response:**

**Question 2.** *What is weight decay? Why would one use it?*

**Response:**

**Question 3.** *The AlexNet paper used a learning rate schedule where the learning rate was lowered when validation error stopped improving. Why is it reasonable to have a schedule where learning rate decreases? Why wait until validation error stops improving?*

**Response:**

**Question 4.** *The authors in the AlexNet paper said that the kernels on one GPU were “largely color-agnostic,” whereas the kernels on the other GPU were largely “color-specific.” Further, they said this specialization occurs “during every run and is independent of any particular random weight initialization.” Can you explain why this would be more likely to happen than having the color-agnostic and color-specific features interspersed on each GPU?*

**Response:**

**Question 5.** *Explain the different data augmentation strategies used in the AlexNet paper? What do these strategies accomplish?*

**Response:**

**Question 6.** *The ResNet paper reports 3.57% error on the ILSVRC. Some people would claim this performance is superhuman. Look up the rate of error achieved by humans. Why is the human error rate not 0%? (After all, wasn’t it labelled by humans?) Do you think it is fair to say that this net can achieve superhuman performance at image classification?*

**Response:**

**Question 7.** *Explain Figure 4 of the ResNet paper, including what the context is, what is being plotted, what is being observed, and what is being concluded. Make sure to explain why there are two sudden steep drops in error % in both plots.*

**Response:**

**Question 8.** *Estimate the number of weight parameters in the three nets depicted in Figure 3 of the ResNet paper.*

**Response:**