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

HW 2
Due: Wednesday June 24, 2020 at 11:59 PM Eastern time via Gradescope

Name: [Put Your Name Here]
Collaborators: [Put Your Collaborators Here]

You may consult any and all resources. Note that these questions are somewhat vague by design. Part of your task is to make reasonable decisions in interpreting the questions. Your responses should convey understanding, be written with an appropriate amount of precision, and be succinct. Where possible, you should make precise statements. For questions that require coding, you may either type your results with figures into this tex file, or you may append a pdf of output of a Jupyter notebook that is organized similarly. You may use PyTorch, TensorFlow, or any other packages you like. You may use code available on the internet as a starting point.

**Question 1.** *Image denoising by ResNets*

Consider the following denoising problem. Let $x$ be a color image whose values are scaled to be within [0,1]. Let $y$ be a noisy version of $x$ where each color channel of each pixel is subject to additive Gaussian noise with mean 0 and variance $\sigma^2$. You will need to clip the values of $y$ in order to ensure it is a valid image. The denoising problem is to estimate $x$ given $y$.

(a) Look up the definition of Peak Signal-to-Noise Ratio (PSNR). Determine what value of $\sigma$ corresponds to an expected PSNR between $x$ and $y$ of approximately 20 dB.

**Response:**

(b) Create a noisy version of the CIFAR-10 training and test dataset, such that it has additive Gaussian white noise with PSNR approximately 20 dB. Show several pairs of images and their noisy version.

**Response:**

(c) Train a ResNet to denoise noisy CIFAR-10 images. Your net should take a noisy 32x32 px image as an input, and it should output a denoised 32x32 px image. Determine the mean and standard deviation of the recovery PSNRs over the noisy test set. Visually show the performance on several noisy images.

**Response:**

(d) Repeat the previous task but without the skip connections in your model.

**Response:**

**Question 2.** *Adversarial examples*

Obtain an image classifier for CIFAR-10. You may use the one you trained in HW 1.

(a) Select a couple images from the test set. For each image, select a target class that is different from the image's true class. For each image, compute an adversarial perturbation that is barely perceptible to the human eye and that results in the image being classified as the target class. Clearly state the algorithm that you used to generate the perturbation. Show the underlying images, the perturbed image, the perturbation, and the classifier's output.

**Response:**

(b) Compute a universal adversarial perturbation for the CIFAR-10 training dataset. Your goal is to compute a single perturbation that can be added to any CIFAR-10 image, resulting in misclassification. Your misclassification does not have to be targeted toward a specific class. Clearly state the algorithm that you used to generate the perturbation. Demonstrate the performance of your perturbation and compute the fraction of the training dataset for which the perturbation results in misclassification.

**Response:**