CS 410/510: Deep Learning
Assignment 3
Winter 2022
Due Feb. 25th by 11:59pm

There are two parts to this assignment.

PART 1:

You will learn *transfer learning* where you take a pre-trained model and fit it to your classification problem.

Reading and sample code:

1. https://pytorch.org/tutorials/beginner/transfer_learning_tutorial.html
2. https://d2l.ai/chapter_computer-vision/fine-tuning.html

1. I want you to start with a pre-trained Resnet50 model. Next, replace the output layer with an output of 10 neurons. Randomly initialize the weights and biases coming into this layer from the previous layer (i.e., last hidden layer). When you train this network, **you need to freeze all parameters except the weights and biases coming from the last hidden layer to the output layer**.
    a. You will train this network on MNIST. I want you to train the network with train sizes of 10*10, 30*10, 50*10, 70*10, 90*10 (to explain, 10*10 means ten images of each of the ten categories). Use whatever learning rate you think is appropriate and as many epochs as you think is sufficient. Use cross-entropy loss. The goal is to get the highest **test** accuracy for each case. Plot the lowest test error vs size of train set. Plot the number of epochs for each to get the best test accuracy.
2. Repeat the above with VGG19

Notes: MNIST images need to be scaled up to 224x224

PART 2:

For this you will experiment with the Spatial Transformer Network discussed in class. I suggest you read the paper as well.

1. Download and run the code in this tutorial: https://pytorch.org/tutorials/intermediate/spatial_transformer_tutorial.html# List the deformations that the model learned and show examples of each. Quantify test accuracy of this algorithm.
2. Download https://pypi.org/project/elasticdeform/ and then add images to the test and training datasets for MNIST using the random deformation (first example in the package) and zoom deformation (zoom out 4x). Your dataset will go from 60k to 180k. Run the STN on this new data and plot the train and text accuracy. Show via images that the network has (or has not) learned these two deformations. Explain what you observe.