CMPSC 497 - Deep Learning for Computer Vision
Professor: Huijuan Xu
Homework 3
Due: February 27, 2023 @ 11:59 PM EST

1   Assignment Details

In homework 3 we will be creating an image captioning system using convnets and a RNN (LSTM in this case). We allow for the homework to be completed locally (with Jupyter), though we highly recommend students complete this in Google Colab if you do not have a GPU on your machine.

2   Image Captioner (100%)

Students will create and test an image captioning system in Pytorch which works with Microsoft's COCO dataset. We will get to train and test the network, and also load in weights from an already trained network whose weights' have been uploaded.

3   Rubric for Assignment Grading

Transform: 30%

Dataloader: 25%

Feed image: 10%

Create encoder and load pre-training weights: 35%

4   Getting Started + Submission

We suggest students complete the networks in Google Colab. If you'd like to complete the assignments in colab, you can visit the colab website and upload the notebook. To use a GPU, set your runtime to include a hardware accelerator. Students may also complete the homework locally with Jupyter, though training your network will be fairly slow on a CPU.

For submitting the assignment, simply upload the completed ipynb file. Be sure the cells have output from running your code. You do not need to include any other files (checkpoints, images, or h5py).