Lab Two

MSDS Summer 2022

- Submit code via Canvas using Markdown Cells to **clearly** indicate which code answers which question and to answer short answer questions.
- Please put your name at the top of the Assignment as well.
- Failure to do either of the above will result in points deducted.

For this assignment download the bird species classification dataset here: `https://www.kaggle.com/gpiosenka/100-bird-species`. This is a Multi-Class classification problem (you'll need `nn.CrossEntropyLoss`). Since this dataset is so large, I suggest you select between 3 and 10 species of birds to classify. Also just use the `birds` folder; you can discard `birds_rev2`.

1. Hint for the following: use the csv file in here to your advantage, but be careful of the forward slashes!

(a.) Create a Dataset such that the output is tensor of size $(3, 128, 128)$, resizing "on-the-fly", meaning using `cv2.resize()` during `__getitem__`.

(b.) Is this a good idea? How can we save time during training?

2. Create the following CNN model and train it on the training data (or a subset if it's too big) for at least 5 epochs. Print the accuracy of the resulting model (Hint: You will need to understand `nn.CrossEntropyLoss` and `nn.LogSoftmax`).

- The CNN has convolution layers with filter size of 3×3 and same padding.
- Convolutions are followed by an activation function and a pooling layer with filter size 2×2 and stride 2.
- The size of the input tensor is $(N, 3, 128, 128)$ and the size of the tensor after convolution+pooling layers is $(N, 64, 8, 8)$.
- The CNN ends with two Fully Connected linear layers.

3. Read the Albumentations GitHub ReadMe. Pick your favorite augmentation (pick a weird one for fun). Implement it with probability $p = 1$ and see how it affects an image.

4. Fix a reasonable suite of data augmentations for your training data. Compare the performance of three models on the bird species dataset:

- partially frozen VGG-16
- finetuned VGG-16
- VGG-16 from scratch (set `pretrained=False`)