Lab Three
MSDS Summer 2021

- Submit code via GitHub Classroom 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.

1. Write a function to apply a general 2×2 filter (as a numpy array)

$$\begin{bmatrix} w_{11} & w_{12} \\ w_{21} & w_{22} \end{bmatrix}$$

with stride one (no padding) to another array.

2. Consider a filter of size $f \times f$ and a matrix of size $m \times n$ (where $f < m, n$).

(a.) What is the dimension of the output matrix, assuming stride 1, in terms of f, m, n?

(b.) If the padding of the matrix is given by an integer p, rewrite the formula now including p.

(c.) For both $(a.)$ and $(b.)$, use your answer from question one to double-check your formula.

3. Suppose you have a tensor of size $(N, 64, 128, 128)$ where N is the batch size and 64 is the number of channels.

(a.) Define the following hyperparameters so that the resulting tensor is of size $(N, 128, 64, 64)$. Note there may be more than one answer (Hint: first set filter size to 2×2).

   – Number of Filters:
   – Filter Size:
   – Stride:
   – Padding:

(b.) How many parameters does each filter have? How many parameters are there in total in this convolutional layer?

For the rest of the assignment download the bird spedies 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`.

4. 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?

5. 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.