# CSCI-4160 Homework 9

Aidan Westphal, Anish Dhanasekaran, Abraham Alpuerto

December 2024

# Question 1

Our input dimension is $252 \times 189 \times 3$, we now trace through each convolutional layer:

```
CNN 1: 3 features -> 32 features, kernel 3: 3 * 32 * 3 * 3 = 864
CNN 2: 32 features -> 64 features, kernel 3: 32 * 64 * 3 * 3 = 18432
CNN 3: 64 features -> 128 features, kernel 3: 64 * 128 * 3 * 3 = 73728
CNN 4: 128 features -> 256 features, kernel 3: 128 * 256 * 3 * 3 = 294912
CNN 5: 256 features -> 512 features, kernel 3: 1179648

We then have a max pool with a stride of 2, so our image size is halved.
Due to odd dimensions, an infinite dimension is assumed, e.g. we round up
any odd numbers. Thus, we consider the pooling on an image shape of
[252 x 190 x 512]. Our output dimension is thus [126 x 96 x 512]. We then
use an adaptive average pool which reduces this down to a single vector,
e.g. [1 x 1 x 512]. Thus, our input to the FC layers is 512 features.

FC 1: 512 features -> 1024 features, 512 * 1024 = 524288
FC 2: 1024 features -> 11 features (out), 1024 * 11 = 11264

TOTAL PARAMS: 2103136 (2.1 million)
(ignoring batch norm parameters)
```

# Question 2

Our team improved the CIFAR-10 accuracy by enhancing the network architecture with an additional layer, more filters, batch normalization, and pooling layers (this helped to avoid overfitting on the training). Thus these changes allowed the model to improve training stability and generalization.