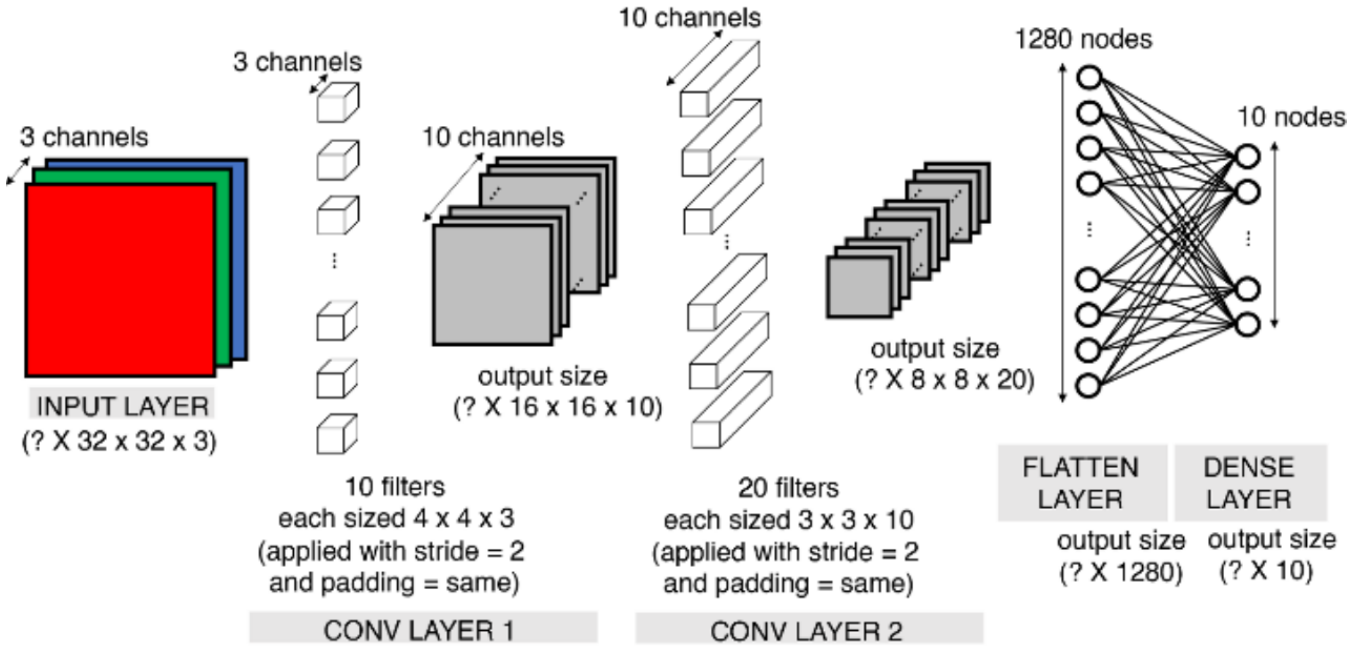
COMP 474/6741 Intelligent Systems (Winter 2021)

Worksheet #10: Introduction to Deep Learning



Task 1. Consider the following matrix that represents an image. This image will be fed into a convolutional neural network (CNN):

1	1	2	2	2	0	0
2	0	1	1	2	1	2
0	1	0	0	1	1	2
0	2	1	2	0	2	2
1	2	0	0	1	0	1
0	0	0	0	1	2	1
2	0	0	0	2	1	1

Assume that we use the following convolution filter with a stride of 2:

0	1	1
0	1	0
0	-1	-1

What will be the size of the activation map? .....

What will be the resulting activation map?

Task 2. The task here is to understand the structure of the CNN (shown on the first page) we are building: Our first convolution layer has 10 filters, each sized 4×4×3 (`kernel_size = (4,4)`), thus $(4 \times 4 \times 3 + 1) \times 10 = 490$ weights (parameters to train).

How do we obtain the output shape of this layer? The general formula you can use is (for `padding = "same"`, meaning the size of the kernel is the same as the input, padded with zeros):

$$\text{output\_shape} = \left(\text{None}, \frac{\text{input\_height}}{\text{stride}}, \frac{\text{input\_width}}{\text{stride}}, \text{filters}\right)$$

Now compute the output tensor shape of the first convolution layer using the formula above:
output_shape = ...

If you arrived at `[16, 16, 10]`, congratulations!

In the second convolution layer, we want to apply 20 filters of size 3×3 and a depth of 10.
How many weights do we have to train? .....

Compute again the output tensor shape using the formula above:
output_shape = ...

Task 3. What will be the output of a pooling layer with a size of 2×2 and a stride of 1, on the activation map of Task 1 above, if we use the following strategies:

1. Average pooling:

2. Max pooling: