PROBLEM 27

3 Convolutional Network Architecture**

Consider the following 1D input and two convolutional filters:

(a) Convolve the first filter with the input.

(b) Convolve the second filter with the input.

(c) What would the output of the convolutional layer look like with both filters?

(d) In the previous part, the spatial size of the output decreased from that of the original input -ie. it had fewer neurons. In some settings, we preserve the size by applying zero padding. What would the output of the convolutional layer look like in this case?

(e) Continuing to use zero padding of size 1, we now use a stride length of 2. What is the output of the convolutional layer now? Compare how this relates to the case with no stride.

(f) If our filters were size 5, rather than size 3, with stride 1, how much padding would you apply to maintain output size?

(g) What is a general expression for the output size (number of neurons) in a convolutional layer, in terms of the filter size (F), stride length (S), amount of padding (P), and the input size (W)?

(h) Consider a 10x10 grayscale image input (no RGB channel). If you made a 1-layer fully connected network, with 1 hidden, how many weight parameters would be required in the hidden layer including bias terms (ignore the output layer)?

(i) Consider a 10x10 grayscale image input (no RGB channel). If you made a 2-layer fully connected network, with 10 and 5 hidden units respectively, how many weight parameters would be required in these hidden layers including bias terms (ignore the output layer)?

(j) With the same 10x10 input image, you now use a convolutional layer with F = 2 (square filters), S = 2 (in both dimensions), and P = 0. How is the output volume of the first convolutional layer with 2 filters, including bias terms?

(l) One advantage of convolutional networks is that the number of free parameters can be controlled by parameter sharing for a filter across spatial dimensions. Each filter is replicated across the entire visual field (image). That is, if a filter is used at one spatial coordinate (x_1, y_1), the same weights are also used in a different position (x_2, y_2). Adopting the parameter sharing scheme, how many parameters does the convolutional layer have?