CNN Homework

- Assume a traditional CNN with an initial input image of 16x16, followed by a convolutional layer with 8 feature maps using 5x5 receptor fields, followed by a max pooling layer with 2x2 receptive fields, followed by a convolution layer with 10 feature maps using 3x3 receptor fields. Those outputs go straight into (no additional pooling layer) a fully connected MLP with 20 hidden nodes followed by 3 output nodes for 3 possible output classes. Assume no zero-padding and stride=1 for convolution layers, no overlap and no trainable weights for the one pooling layer, and convolutional maps connect to all maps in the previous layer. Sketch the network. For each layer state a) What is the size of the maps in the layer (e.g. the input layer is 16x16), b) how many unique trainable weights are there per layer, and c) total connections in the layer. Show your work and explain your numbers in each case (similar to the previous slide).

CNN Homework Solution

Layer	Trainable Weights per layer	Total Connections
C1 (12x12)	(25+1)*8 = 208	208*12*12 = 29,952
S1 (6x6)	0	4 (2x2 links) *8*6*6 = 1152
C2 (4x4)	(3*3*8+1)*10 = 730	730*4*4 = 11,680
Hidden	(4*4*10+1)*20 = 3220	Same since fully connected MLP
Output	(20+1)*3 = 63	Same since fully connected MLP

- C1 is 12x12 because 5x5 filter causes dropping 2 features on each side
- S1 is 6x6 because 2x2 pooling filter halves the length of each side
- C2 is 4x4 because 3x3 filter causes dropping 1 feature on each side