1 point	1.	What do you think applying this filter to a grayscale image will do? $ \begin{bmatrix} 0 & 1 & -1 & 0 \end{bmatrix} $
		$\begin{bmatrix} 1 & 3 & -3 & -1 \\ 1 & 3 & -3 & -1 \\ 1 & 3 & -3 & -1 \\ 0 & 1 & -1 & 0 \end{bmatrix}$
		Detect horizontal edges
		Detect image contrast
		Detect 45 degree edges
		Detect vertical edges
1 point	2.	Suppose your input is a 300 by 300 color (RGB) image, and you are not using a convolutional network. If the first hidden layer has 100 neurons, each one fully connected to the input, how many parameters does this hidden layer have (including the bias parameters)?
		9,000,001
		9,000,100
		27,000,001
		27,000,100
1 point	3.	Suppose your input is a 300 by 300 color (RGB) image, and you use a convolutional layer with 100 filters that are each 5x5. How many parameters does this hidden layer have
		(including the bias parameters)?
		2600
		7500
		7600
1 point	4.	You have an input volume that is 63x63x16, and convolve it with 32 filters that are each 7x7, using a stride of 2 and no padding. What is the output volume?
		16x16x16
		© 29x29x32
		16x16x32
		29x29x16
	5.	You have an input volume that is 15x15x8, and pad it using "pad=2." What is the
1 point	J.	dimension of the resulting volume (after padding)?
		19x19x12
		17x17x8
		19x19x8

1 point	6.	You have an input volume that is 63x63x16, and convolve it with 32 filters that are each 7x7, and stride of 1. You want to use a "same" convolution. What is the padding?				
		<u> </u>				
		O 2				
		3				
		O 7				
1 point	7.	You have an input volume that is 32x32x16, and apply max pooling with a stride of 2 and a filter size of 2. What is the output volume?				
		16x16x8				
		32x32x8				
		16x16x16				
1 point	8.	Because pooling layers do not have parameters, they do not affect the backpropagation (derivatives) calculation.				
		True				
		False				
1 point	9.	In lecture we talked about "parameter sharing" as a benefit of using convolutional networks. Which of the following statements about parameter sharing in ConvNets are true? (Check all that apply.)				
		lt reduces the total number of parameters, thus reducing overfitting.				
		It allows gradient descent to set many of the parameters to zero, thus making the connections sparse.				
		It allows a feature detector to be used in multiple locations throughout the whole input image/input volume.				
		It allows parameters learned for one task to be shared even for a different task (transfer learning).				
1 point	10	. In lecture we talked about "sparsity of connections" as a benefit of using convolutional layers. What does this mean?				
		Each layer in a convolutional network is connected only to two other layers				
		Regularization causes gradient descent to set many of the parameters to zero.				
		Each filter is connected to every channel in the previous layer.				
		Each activation in the next layer depends on only a small number of activations from the previous layer.				
this co	urse or (ndelwal, understand that submitting work that isn't my own may result in permanent failure of deactivation of my Coursera account.				

Submit Quiz



