## Unit 7 - Week 5

Course outline
How does an NPTEL online course work?
Week 0
Week 1
Week 2
Week 3
Week 4
Week 5
Convolutional Neural Networks     An Introduction - Part 01
Convolutional Neural Networks     An Introduction - Part 02
Backpropagation in CNNs
<ul> <li>Evolution of CNN Architectures for Image Classification-Part 01</li> </ul>
<ul> <li>Evolution of CNN Architectures for Image Classification-Part 02</li> </ul>
Recent CNN Architectures
<ul> <li>Finetuning in CNNs</li> </ul>
Lecture Materials
O Quiz: Weekly Quiz 5
<ul> <li>Deep Learning for Computer</li> <li>Vision: Week 5 Feedback Form</li> </ul>
Week 6
Week 7
Week 8
Week 9
Week 10
Week 11
Week 12
Download Videos
Live Session
Text Transcripts

he due date for submitting this assignment has passed. Is per our records you have not submitted this assignment.	Due on 2020-10-21, 23:59 IS
) Which of the following orders the deep architectures in the increasing number of parameters:	1
○ Lenet < AlexNet < InceptionNets < VGGNet ○ InceptionNets < VGGNet < Lenet < AlexNet	
Lenet < InceptionNets < AlexNet < VGGNet	
Lenet < AlexNet < InceptionNets < VGGNet     InceptionNets < VGGNet	
Score: 0 Accepted Answers:	
enet < InceptionNets < AlexNet < VGGNet	
) A DenseNet with $m{L}$ layers has the following number of direct connections:	1
L	
L+1	
$igcup_{L(L+1)}$	
(L(L+1))/2	
lo, the answer is incorrect. Score: 0	
accepted Answers: $L(L+1))/2$	
	1
Which among the following is False:     Dilated convolution increases the receptive field size when compared to standard convolution operator	1
O Dropout is a regularization technique	
Batch normalization ensures that the weight of each of the hidden layer of a deep network is normalised     Deep neural networks are translation invariant	
lo, the answer is incorrect.	
Accepted Answers: Batch normalization ensures that the weight of each of the hidden layer of a deep network is normalised	
) Which of the following is true?	1
Hyper-parameters of a model like learning-rate, number of hidden layers can be selected based on the performance.	
The performance evaluation for early-stopping is computed on the test set	
Adding a hidden layer of 10 neurons, would make a fully-connected feed-forward network with no hidden layer, connected feed-forward network with no hidden layer.	ecting 6 input units to 3 output units mor
Larger batch sizes reduces the variance in gradient estimation of SGD  lo, the answer is incorrect.	
core: 0 Accepted Answers:	
Hyper-parameters of a model like learning-rate, number of hidden layers can be selected based on the Performance on the validation split of the data	
arger batch sizes reduces the variance in gradient estimation of SGD	
In a layer of Convolutional Neural Network, assume the input size to be $256 \times 256 \times 3$ . Using the following values for the number of parameters respectively are? Stride 'S' = 2, Padding 'P' = 1, Filter size 'F' = 5, Number of Filters 'K' = 64	r hyper parameters, the output size 1
0	
126 × 126 × 64 ,4800	
125 × 125 × 32, 102400	
$127 \times 127 \times 64, 102400$	
127 × 127 × 64, 4800	
lo, the answer is incorrect. core: 0	
accepted Answers: 27 × 127 × 64, 4800	
) Which of the following statements is <b>False</b> ?	1
Number of parameters in CNNs are usually less than the number of parameters in Feed forward Neural Networks	
CNNs are prone to overfitting because of less number of parameters     There are no learnable parameters in Pooling layers	
In a max pooling layer, the unit that gets contributed(maximum entry) in forward propagation gets all the gradient while	e backpropagation
lo, the answer is incorrect. Score: 0	
Accepted Answers: CNNs are prone to overfitting because of less number of parameters	
) Which of the following statements is <b>true</b> about GoogLeNet?	1
O Inception modules without bottleneck layers reduce the number of parameters	
Bottleneck layers use 1x1 convolutions to increase the feature depth     Bottleneck layers use 1x1 convolutions to decrease the feature depth but still preserve spatial dimensions	
Residual blocks are used in GoogLeNet to avoid the problem of exploding gradients	
lo, the answer is incorrect. Score: 0 Accepted Answers:	
Bottleneck layers use 1x1 convolutions to decrease the feature depth but still preserve spatial dimensions	
) Which of the following statements is <b>False</b> ?	1
Residual blocks are used in ResNets to avoid the Vanishing Gradient problem	
Parameter sharing in CNNs helps in extracting features that are invariant to spatial position     Batch Normalization is used to prevent internal covariate shift	
Number of parameters in the pooling layer are more than the number of parameters in Fully connected layer	
lo, the answer is incorrect. Score: 0 Accepted Answers:	
lumber of parameters in the pooling layer are more than the number of parameters in Fully connected  ayer	
) Which of the following statements is True about parameter sharing in ConvNets?  OParameters learned for one task to be shared even for a different task (transfer learning)	1
Allows gradient descent to set many of the parameters to zero, thus making the connections sparse	
Allows a feature detector to be used in multiple locations throughout the whole input image/input volume     Using weights of each feature to find the best model for prediction, sharing the results and returning the average	
lo, the answer is incorrect. score: 0	
ccepted Answers: llows a feature detector to be used in multiple locations throughout the whole input image/input volume	
O) What does sparsity of connections mean as a benefit of using convolutional layers?	1
Each filter is connected to every channel in the previous layer	•
Each layer in a convolutional network is connected only to two other layers	
Each activation in the next layer depends on only a small number of activations from the previous layer     Regularization causes gradient descent to set many of the parameters to zero	
lo, the answer is incorrect. core: 0	
ccepted Answers: ach activation in the next layer depends on only a small number of activations from the previous layer	
1) Suppose you have an input volume of dimension 64x64x16. How many parameters would a single 1x1 convolutional f	ilter have (including the bias)? 1
O <sub>2</sub>	
○ 4097 ○ 1	
○ 17	
lo, the answer is incorrect. core: 0	
Accepted Answers: 7	
2) Which of the following statements are true for an input volume of dimension $H imes W imes C$ (Assuming stride of 1 and $R$	no padding) 1
1x1 convolutional layer can reduce $H,W$ and $C$	
1x1 convolutional layer can reduce $C$ but not $H$ and $W$	
Pooling layer can reduce $H,W$ but not $C$	
Pooling layer can reduce $H,W$ but not $C$ Pooling layer can reduce $H,W$ and $C$	