Unit 10 - Week 7

Assignment 7

The due date for submitting this assignment has passed.
As per our records you have not submitted this assignment.

Due on 2019-09-18, 23:59 IST.

Kindly download the paper from given link:
https://www.cv-foundation.org/openaccess/content_cvpr_2015/papers/Long_Fully_Convolutional_Networks_2015_CVPR_paper.pdf

1) In the context of Fully Convolutional Networks (FCNs) which of the following statements are true? ***1 point***

- During inference of FCNs, the size of the input image feed need not be the same as that used during training
- The size of the output feature maps of FCNs is independent of the input image size
- For image segmentation task FCNs are computational inefficient when compared to patch-based CNNs
- When comparing between FCNs and patch-based CNNs for segmentation task, FCNN computation is highly amortized over the overlapping regions of those patches

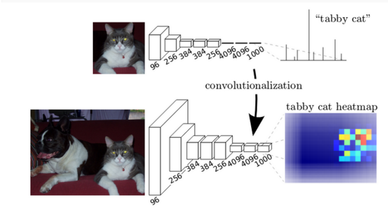
No, the answer is incorrect.
Score: 0
Accepted Answers:
During inference of FCNs, the size of the input image feed need not be the same as that used during training
When comparing between FCNs and patch-based CNNs for segmentation task, FCNN computation is highly amortized over the overlapping regions of those patches

2) One issue in this specific FCN is that by propagating through several alternated convolutional and pooling layers, the resolution of the output feature maps is downsampled. Therefore, the direct predictions of FCN are typically in low resolution, resulting in relatively fuzzy object boundaries. What methods are suggested by authors to gain original resolution? ***1 point***

- Bilinear Interpolation
- Deconvolution
- Backward convolution
- Forward convolution

No, the answer is incorrect.
Score: 0
Accepted Answers:
Bilinear Interpolation
Deconvolution
Backward convolution

3) This figure has been taken from the paper mentioned above. It illustrates the transformation of a fully connected network into a fully convolutional network (FCN). In the figure an illustration of feature map size variation along the network is shown for both fully connected (top) and FCN (bottom). In the figure, only the number of features in each layer is mentioned. The 5th layer of the FCN has 256 feature maps and say suppose the size of feature maps is 64x64 (height x width), then select the possible kernel sizes that can be used in the final three (6th, 7th and 8th) convolutional layers of FCN? ***1 point***



- 64x64x256x256, 1x1x4096x4096, 1x1x1000x1000
- 64x64x4096x256, 1x1x4096x4096, 1x1x1000x4096
- 64x64x256x4096, 1x1x4096x4096, 1x1x4096x1000
- 28x28x4096x256, 3x3x1024x4096, 2x2x1024x1000

No, the answer is incorrect.
Score: 0
Accepted Answers:
64x64x256x4096, 1x1x4096x4096, 1x1x4096x1000

4) Which of the following statement/s is/are True? ***1 point***

- Decreasing subsampling within a net is a tradeoff: the filters see finer information, but have smaller receptive fields and take longer to compute.
- The shift-and-stitch trick is another kind of tradeoff: the output is denser without decreasing the receptive field sizes of the filters, but the filters are prohibited from accessing information at a finer scale than their original design
- In this paper, the network is trained with a per-pixel multinomial logistic loss
- The network is validated with the standard metric of mean pixel intersection over the union, with the mean taken over all classes, without including the background

No, the answer is incorrect.
Score: 0
Accepted Answers:
Decreasing subsampling within a net is a tradeoff: the filters see finer information, but have smaller receptive fields and take longer to compute.
The shift-and-stitch trick is another kind of tradeoff: the output is denser without decreasing the receptive field sizes of the filters, but the filters are prohibited from accessing information at a finer scale than their original design
In this paper, the network is trained with a per-pixel multinomial logistic loss

5) According to the paper, the optimization parameter/s used in training the FCN-VGG16 is/are: ***1 point***

- Stochastic Gradient descent used as optimization algorithms
- Learning rate is 10-2
- For hyperparameter, Grid Search has been used
- Dropout has been not used in the network

No, the answer is incorrect.
Score: 0
Accepted Answers:
Stochastic Gradient descent used as optimization algorithms

6) In the context of deep learning, transfer learning depends on which of the following variables: ***1 point***

- The similarity of the task data to original data used for pre-training model weights
- Amount of data available
- Complexity of data
- Network Architecture

No, the answer is incorrect.
Score: 0
Accepted Answers:
The similarity of the task data to original data used for pre-training model weights
Amount of data available
Complexity of data

7) In this context of Transfer Learning which of the following statements are True? ***1 point***

- Transfer learning enables feature extraction with pre-trained deep learning models
- When target labels are scarce, the weights of pre-trained models are frozen (fix weights) so as to avoid underfitting.
- Fine-tuning of pre-trained model weights is generally preferred when target task labels are scarce
- Transfer learning works better when the tasks on which the networks are trained for are similar

No, the answer is incorrect.
Score: 0
Accepted Answers:
Transfer learning enables feature extraction with pre-trained deep learning models
Transfer learning works better when the tasks on which the networks are trained for are similar

8) Choose the correct statement with regard to different approaches used for hyper-parameter optimization. ***1 point***

- Grid search approach is usually very efficient, and each new guess is independent of the previous guess
- Random search approach is usually more efficient than grid search, and each new guess is independent of the previous guess
- In Bayesian optimization, each new guess is independent of the previous guess
- Both A and B

No, the answer is incorrect.
Score: 0
Accepted Answers:
Random search approach is usually more efficient than grid search, and each new guess is independent of the previous guess

9) Which of the following statements are true with respect to padding operation performed on the input feature maps before performing convolution operation? ***1 point***

- Padding aids in designing deeper networks
- Padding enables preserving the spatial size of the feature maps after convolution
- Padding increases performance by keeping information at borders of the feature maps
- Padding aids in the concatenation of feature maps generated after performing convolution operation with different kernel sizes

No, the answer is incorrect.
Score: 0
Accepted Answers:
Padding aids in designing deeper networks
Padding enables preserving the spatial size of the feature maps after convolution
Padding increases performance by keeping information at borders of the feature maps
Padding aids in the concatenation of feature maps generated after performing convolution operation with different kernel sizes

10)Which of the following are not examples of Hyper-parameters in the context of machine learning: ***1 point***

- The learning rate for training a convolution neural network
- The filter weights in the convolution neural network
- The penalty in Logistic Regression Classifier i.e. L1 or L2 regularization
- The coefficients in logistic regression

No, the answer is incorrect.
Score: 0
Accepted Answers:
The filter weights in the convolution neural network
The coefficients in logistic regression

11)Consider two different models: ***1 point***

$$Model1: \hat{y} := wx_1 + w_0$$

$$Model2: \hat{y} := \sum_{i=1}^{n} w_i x^i + w_0$$

What should be the value of the regularisation parameter (i.e., hyperparameter)?

- Low for both models
- High for both models
- High for model 1, low for model 2
- Low for model 1, high for model 2

No, the answer is incorrect.
Score: 0
Accepted Answers:
Low for model 1, high for model 2