Convolutional Neural Networks Practice

Choose one answer per question. Circle A, B, C, or D.

1. Which operation lets a convolutional layer use the same learned weights to look for a feature in different parts of an image?
   A. Sliding a filter across the image
   B. Sorting the pixels by intensity
   C. Assigning independent weights to each image position
   D. Replacing all pixel values with their global mean

2. What does a 3 by 3 kernel describe for an ordinary two-dimensional convolution with dilation 1?
   A. Three passes over the entire image
   B. An output with three channels
   C. A window three pixels high and three pixels wide
   D. A step size of three pixels in each direction

3. Which setting says how far the filter window moves between neighboring positions?
   A. Bias
   B. Stride
   C. Batch size
   D. Learning rate

4. What gets added just outside an image when a layer uses zero padding?
   A. New learned pixel values
   B. The image-wide average
   C. A copy of each corner pixel
   D. Zeros

5. A max-pooling window contains 2, 7, 4, and 1. What single value does the layer return for that window?
   A. 2
   B. 7
   C. 3.5
   D. 14

6. An ordinary 2D convolution takes an RGB image and applies 12 learned filters. How many channels does its output have?

A. 12
B. 3
C. 4
D. 36

7. What spatial shape comes from 2 by 2 max pooling on a 14 by 14 feature map with stride 2 and no padding?

A. 13 by 13
B. 6 by 6
C. 12 by 12
D. 7 by 7

8. A 3 by 3 filter scans an 8 by 8 input with stride 1 and no padding or dilation gaps. What is the output’s spatial shape?

A. 7 by 7
B. 5 by 5
C. 6 by 6
D. 8 by 8

9. How many input cells must be padded on each side to preserve the spatial size with a 5 by 5 convolution, stride 1, and dilation 1?

A. 2
B. 1
C. 3
D. 4

10. A 3 by 3 convolution uses stride 2 and one cell of padding on every side. With dilation 1, what output width comes from an input width of 10?

A. 7
B. 4
C. 6
D. 5

11. An ungrouped 2D convolution maps 3 input channels to 8 output channels using 3 by 3 kernels and one bias per output channel. How many trainable values are there?

A. 216
B. 72
C. 224
D. 288

12. A 3 by 3 kernel uses dilation 2. Along one axis, what input span does the kernel cover, counting the gaps?

A. 3 cells
B. 5 cells
C. 6 cells
D. 7 cells

13. Using floor rounding, how many output positions along one axis does a 3-cell pooling window get from 10 input cells, stride 3, and no padding?

A. 5
B. 2
C. 4
D. 3

14. A 3 by 3 convolution maps 8 input channels to 12 output channels with 4 groups and no biases. How many weights does it have?

A. 216
B. 108
C. 228
D. 864

15. Two 3 by 3 convolutions are stacked with stride 1, dilation 1, and no pooling. What spatial region of the original input can affect one interior output location in the second layer?

A. 6 by 6
B. 3 by 3
C. 5 by 5
D. 9 by 9