

Instituto Superior de Agronomia, ULisboa

Green Data Science

Practical Machine Learning/Aprendizagem Automática Aplicada

Questionnaire #5, March 31, 2023 – One possible resolution

Name: \_\_\_\_\_

Topic: Convolutional neural networks

1. A max pooling layer is used to reduce the size in the encoding process of a CNN. Consider the following input layer with  $nrows=ncolumns=6$ . If you apply max pooling with a  $2*2$  kernel with stride 2, what would be the  $3*3$  output?

|    |    |    |    |    |    |
|----|----|----|----|----|----|
| 9  | 7  | 11 | 5  | 6  | 10 |
| 4  | 13 | 11 | 8  | 9  | 9  |
| 14 | 7  | 11 | 10 | 6  | 3  |
| 3  | 8  | 11 | 9  | 3  | 6  |
| 22 | 22 | 22 | 22 | 22 | 6  |
| 9  | 4  | 8  | 3  | 3  | 8  |

|    |    |    |
|----|----|----|
| 13 | 11 | 10 |
| 14 | 11 | 6  |
| 22 | 22 | 22 |

2. Instead of applying “max pooling” to reduce the layer size, one can apply a convolution layer with stride 2. If you apply a  $2*2$  kernel with weights  $(1,0,0,0)$  to the same input, with stride 2, with no padding, what would be the output?

|    |    |    |    |    |    |
|----|----|----|----|----|----|
| 9  | 7  | 11 | 5  | 6  | 10 |
| 4  | 13 | 11 | 8  | 9  | 9  |
| 14 | 7  | 11 | 10 | 6  | 3  |
| 3  | 8  | 11 | 9  | 3  | 6  |
| 22 | 22 | 22 | 22 | 22 | 6  |
| 9  | 4  | 8  | 3  | 3  | 8  |

|    |    |    |
|----|----|----|
| 9  | 11 | 6  |
| 14 | 11 | 6  |
| 22 | 22 | 22 |

3. Suppose you have an input image with 3 channels and you apply a 3\*3 kernel to produce by convolution a new hidden layer with depth=10 (i.e. 10 is the number of channels or output features of the hidden layer).

Determine how many parameters this part of the network has if:

- a) you do not include additive weights (bias) in the convolutions

R: Each convolution involves  $3*3*3=27$  multiplicative weights. Since there are 10 output features, we have in total 270 parameters.

- b) you include additive weights (bias) in the convolutions

R: Each convolution involves  $3*3*3=27$  multiplicative weights and one additive weight. Since there are 10 output features, we have in total  $(27+1)*10=280$  parameters.

4. Given the following 4\*4 input (in red), suppose that one applies 1 pixel wide 0-padding (in gray) before a 3\*3 kernel convolution to obtain an output with the same 4\*4 size. Suppose that all kernel weights are positive. What would be the inconvenient consequence of using 0-padding to preserve the size.

|   |    |    |    |    |   |
|---|----|----|----|----|---|
| 0 | 0  | 0  | 0  | 0  | 0 |
| 0 | 35 | 19 | 25 | 6  | 0 |
| 0 | 13 | 22 | 16 | 53 | 0 |
| 0 | 4  | 3  | 7  | 10 | 0 |
| 0 | 9  | 8  | 1  | 3  | 0 |
| 0 | 0  | 0  | 0  | 0  | 0 |

R: Since padding is 0-padding and weights are positive, convolution outputs that involve more padded pixels will tend to have lower values. Therefore, corners after convolution will tend to be darker than pixels along the edges, than pixels in the middle of the image, so the image overall variability will increase.