Week 7 Exercise: Neural Networks

Note: An indicative mark is in front of each question. The total mark is 10. You may mark your own work when we release the solutions.

2   1. Using the definitions for **o** and **h** on slide 10 of Lecture 7 to show that if the activation function is linear such that $g(a) = a$, then the one-hidden-layer on that slide encodes a linear relationship between the input **x** and output **o**. Include all steps.

2   2. In Slide 38: we change the 3×3 kernel to $\begin{bmatrix} 0 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & 0 \end{bmatrix}$. What will be the 3×3 convolved features? What features can this kernel detect?

2   3. For the kernel in Question 2 above, a) show the convolved features with pad=1; and b) show the convolved features with stride=2.

4   4. We have a 256×256 colour image. We apply 80 7×7 filters with stride 3, and pad 3 to obtain a convolution output. What is the output volume size? How many parameters are needed for such a layer?