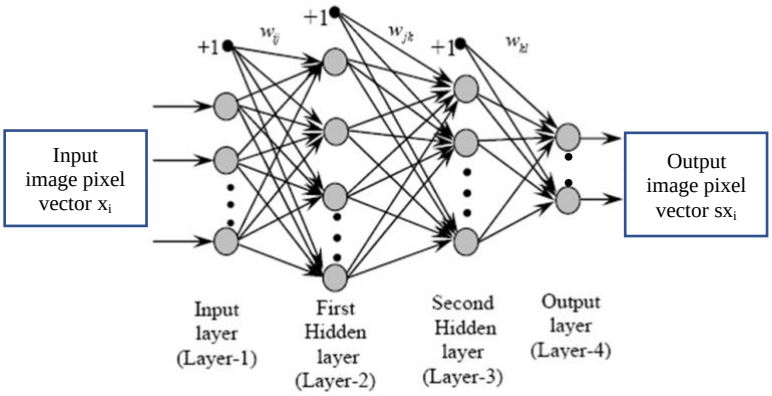
COMP4432 Machine Learning

Tutorial Questions on Convolutional Neural Networks (with answers)

1. Super resolution is the process of upscaling and/or improving the details within an image.

a) Suppose you are asked to use a multilayer perceptron (MLP) neural network as shown in Fig.1 below to learn a mapping from a given lower resolution grayscale (i.e. no color) image x_i to an upscaled grayscale image sx_i.

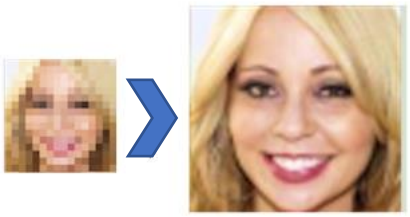


Assume that the input image size is 20×20 pixels and the upscaled image size is 50×50 pixels.

(i) How many learnable parameters, i.e. interconnecting weights, will be involved if there are N_{L2} hidden neurons in Layer-2 and N_{L3} hidden neurons in Layer-3? Note that there exist bias weights with fixed input +1 as shown in the MLP above.

(ii) How training data should be collected and used in the MLP based super resolution model?

b) Suppose now the MLP in part (a) is enhanced with convolutional layers and pooling layers so that a CNN is resulted to carry out **color** image super resolution. Assume that 3 (color) channels are used to represent a color image.



(i) Show the convolution results of the following 6x6 image plane (1 plane only) with the associated 2 3x3 filters using stride=1. Here, no zero padding to the input image is applied.

1	0	1	0	0	2
0	3	0	0	1	0
1	0	1	1	0	0
1	0	0	0	1	0
0	1	0	0	2	0
0	0	1	0	3	0

6x6 image plane

1	-1	-1
-1	1	-1
-1	-1	1

Filter 1

-1	1	-1
-1	1	-1
-1	1	-1

Filter 2

(ii) For the feature maps generated from part (b-i), apply a 2x2 max pooling and show the result. Recall that a 2x2 max pooling is to select the maximum value from a group of 2x2 windowed values. Here, stride=2 is assumed.

(iii) For the following table of CNN architecture, how many learnable parameters are there in each of the specified layers? Show the formula or calculations in your answers.

Layer in CNN	Specification	Number of learnable parameters (formula answer is acceptable)
Input Layer	20x20 color images (3 channels)	
1st Convolutional Layer	16 3x3x3 filters; stride=1; no zero padding	
1st Max Pooling Layer	2x2 window; stride=2	
2nd Convolutional Layer	64 3x3x16 filters; stride=1; no zero padding	
Input layer of fully connected (fc) feedforward network	Just the flattened output from previous layer	
1st hidden layer of fc feedforward network	N_{L2} hidden neurons	
2nd hidden layer of fc feedforward network	N_{L3} hidden neurons	
Output layer	50x50 color images (3 channels)	