Exercise May 16th

This exercise is about understanding transposed convolution using simple examples.

Question 1

a) *Standard Convolution in 2-D*

Suppose you want to apply standard convolution to transform a 4x4 array into a 1x1 array. If the convolution kernel is itself 4x4:

$$\begin{matrix} w_0 & \cdots & w_3 \\ \vdots & \ddots & \vdots \\ w_{12} & \cdots & w_{15} \end{matrix}$$

write the output y (the 1x1) array as a function of the input (the 4x4 array):

$$\begin{matrix} x_0 & \cdots & x_3 \\ \vdots & \ddots & \vdots \\ x_{12} & \cdots & x_{15} \end{matrix}$$

b) *Transposed Convolution in 2-D*

Now that you have identified the convolution structure in question a), its transposed convolution is defined too: it goes the other way, that is it transforms a 1x1 input (a real number) y into a 4x4 output

$$\begin{matrix} x_0 & \cdots & x_3 \\ \vdots & \ddots & \vdots \\ x_{12} & \cdots & x_{15} \end{matrix}$$

Write the expression of the output of the transposed convolution as a function of y and of the convolution kernel coefficients.

c) Now extend the result to the case where the input has two channels y_1 and y_2 and the 4x4 output has a single channel.

d) Now extend the result to the case where the input still has two input channels y_1 and y_2 but the 4x4 output also has 2 channels.

Question 2

a) Start from a simple example of convolution, where we take an image of size 4x4, with no padding, and apply to it a convolution kernel of dimension 3x3 with stride 1. Show, using the usual formula, that the output will be 2x2.

b) If we unroll - or reshape - the input and output, this convolution transforms a vector of 16 coordinates (4x4) into a vector of 4 coordinates (2x2). We know that the convolution operation can be written as a matrix multiplication. The matrix is obviously 4x16. Give the expression of this matrix as a function of the parameters of the convolution kernel:

$$\begin{matrix} w_0 & w_1 & w_2 \\ w_3 & w_4 & w_5. \\ w_6 & w_7 & w_8 \end{matrix}$$

c) When we take the transpose of this matrix and apply it now to an input vector of size 4, the output will be of size 16. Give the expression of the associated 4x4 output.

d) Now there would be another way to obtain the same result, and this is using a convolution (not a transposed convolution), but after applying some padding to the input.

   1. Show, using the same formula as question a), that if we apply a padding of 2 to the 2x2 input image and apply the convolution kernel W to it with a stride of one, we obtain a 4x4 image.

   2. Show that, with a proper choice of the convolution parameters, an identical result to that of question c) is obtained.