CNN

EDGE DETECTION

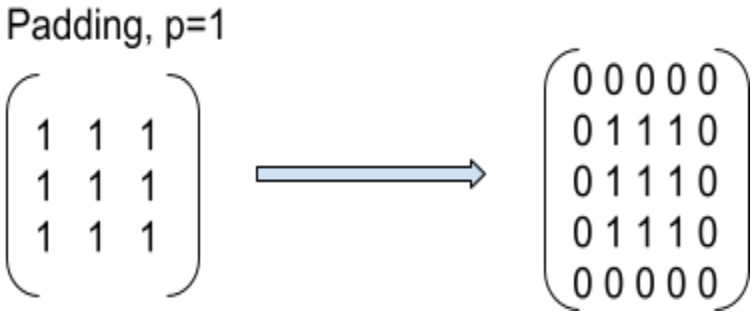
A standard way of applying convolution to, say, detect vertical edges in a matrix is to multiply it with a filter or kernel with specific values for the required output. The asterisk '*' denotes convolution or multiplication or element wise multiplication. Thus, to multiply the 3x3 'kernel' matrix, paste it on the first 3x3 block of the image matrix, and the sum of the element-wise product of those 9 blocks gives us the 1st element's value as our output matrix. To get the second element in the first row of the output, we shift the pasting of the 3x3 kernel 1 column to the right as shown.

$$\begin{pmatrix} 1 & 0 & -1 \\ 1 & 0 & -1 \\ 1 & 0 & -1 \end{pmatrix}$$ This is the kernel used for vertical edge detection.

PADDING

Padding is necessary for
1)keeping the original dimensions of the image and
2)ensuring an even contribution of all edge and corner pixels in the output image.

Our parameters are conventionally p=1 (i.e the image is padded on all sides with a border of 1), and padding_value=0 (i.e the value of the pixels in the padded region is conventionally kept 0).



STRIDED CONVOLUTIONS

Sometimes, if we hop to the right by two columns instead of one, i.e by a stride of 2, we get strided convolutions. This shrinks the size of our output image. So, if we convolve as shown, our output image's dimensions come out to be

Dimension of output after all this is:

((n+2p-f)/s) + 1 x ((n+2p-f)/s) + 1

CONVOLUTIONS OVER VOLUME

Doing convolutions over a mxnx3 matrix, where the 3 corresponds to the number of channels i.e. RGB colour channels. For this, we often use a 3x3x3 matrix filter, or a cube. Thus, this cube's 3 channels move over the 3 channels in the input image and we convolve as usual.

Also, it should be noted that convolving over volumes, even with the input image and kernel being 3-dimensional, we get a 2-dimensional normal output image.
If we use 2 different filters,say to detect vertical and horizontal edges, we end up with two 1-dimensional outputs that can be arranged to be a single output having a 3-rd dimension being equal to the number of filters.