Convolutional Neural Networks

The step size taken by a kernel while sliding over an input image is called as a Stride

We have our usual 6×6 matrix as before

$$A = \begin{pmatrix} 3 & 2 & 7 & 1 & 6 & 3 \\ 4 & 1 & 1 & 0 & 9 & 8 \\ 8 & 2 & 0 & 1 & 2 & 4 \\ 5 & 6 & 0 & 7 & 1 & 4 \\ 0 & 5 & 4 & 9 & 6 & 0 \\ 4 & 6 & 2 & 4 & 0 & 5 \end{pmatrix}$$

We have a convolutional kernel k as below.

$$\underline{\underline{K}} = \begin{pmatrix} 1 & 2 & 2 \\ 2 & 1 & 2 \\ 2 & 2 & 1 \end{pmatrix}$$

$n = 6$

$f = 3 \quad s = 3$

Let us consider a stride $(s) = 3$

The size of output matrix $= \left\lfloor \frac{n-f}{s} + 1 \right\rfloor \times \left\lfloor \frac{n-f}{s} + 1 \right\rfloor$

$$= \left\lfloor \frac{6-3}{3} + 1 \right\rfloor * \left\lfloor \frac{6-3}{3} + 1 \right\rfloor$$

$$= 2 \times 2.$$

We will have output matrix $H = \begin{pmatrix} h_{11} & h_{12} \\ h_{21} & h_{22} \end{pmatrix}$

Here, $h_{11} = 3 \times 1 + 2 \times 2 + 7 \times 2 + 4 \times 2 + 1 \times 1 + 1 \times 2 + 8 \times 2 + 2 \times 2 + 0 \times 1$

$$= \underline{\underline{52}}$$

$$A = \begin{pmatrix} \boxed{\begin{matrix} 3 & 2 & 7 \\ 4 & 1 & 1 \\ 8 & 2 & 0 \end{matrix}} & \begin{matrix} 1 & 6 & 3 \\ 0 & 9 & 8 \\ 1 & 2 & 4 \end{matrix} \\ \begin{matrix} 5 & 6 & 0 \\ 0 & 5 & 4 \\ 4 & 6 & 2 \end{matrix} & \begin{matrix} 7 & 1 & 4 \\ 9 & 6 & 0 \\ 4 & 0 & 5 \end{matrix} \end{pmatrix}$$

$$h_{12} = 1\times1 + 6\times2 + 3\times2 + 0\times2 + 9\times1 + 8\times2 + 1\times2 + 2\times2 + 4\times1 = 54$$

$$\begin{array}{c} \quad 1 \quad 2 \quad 3 \\ A = \begin{pmatrix} 3 & 2 & 7 & 1 & 6 & 3 \\ 4 & 1 & 1 & 0 & 9 & 8 \\ 8 & 2 & 0 & 1 & 2 & 4 \\ 5 & 6 & 0 & 7 & 1 & 4 \\ 0 & 5 & 4 & 9 & 6 & 0 \\ 4 & 6 & 2 & 4 & 0 & 5 \end{pmatrix} \end{array}$$

$$h_{21} = 5\times1 + 6\times2 + 0\times2 + 0\times2 + 5\times1 + 4\times2 + 4\times2 + 6\times2 + 2\times1 = 52$$

$$\begin{array}{cc} & \\ 1 & \\ 2 & A = \begin{pmatrix} 3 & 2 & 7 & 1 & 6 & 3 \\ 4 & 1 & 1 & 0 & 9 & 8 \\ 8 & 2 & 0 & 1 & 2 & 4 \\ 5 & 6 & 0 & 7 & 1 & 4 \\ 0 & 5 & 4 & 9 & 6 & 0 \\ 4 & 6 & 2 & 4 & 0 & 5 \end{pmatrix} \\ 3 & \end{array}$$

$$h_{22} = 7\times1 + 1\times2 + 4\times2 + 9\times2 + 6\times1 + 0\times2 + 4\times2 + 0\times2 + 5\times1 = 54$$

The output matrix is

$$h = \begin{pmatrix} 52 & 54 \\ 52 & 54 \end{pmatrix}$$

$6\times6,\ 3\times3,\ 3$

$h \rightarrow 2\times2$

Padding:- Adding extra layers of zeros along the edges of input image before performing convolution

In case, input size = output size we apply padding.

Input size = $\underline{\underline{n \times n}}$

filter size = $\underline{\underline{f \times f}}$

Output size = $\underline{\underline{n \times n}}$ if $\boxed{p = \frac{f-1}{2}}$

where $\underline{p \text{ is padding size}}$.

$f \times f = 3 \times 3$

$p = \frac{3-1}{2} = 1$

In our example, padded input A_p will be.

$$A_p = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 3 & 2 & 7 & 1 & 6 & 3 & 0 \\ 0 & 4 & 1 & 1 & 0 & 9 & 8 & 0 \\ 0 & 8 & 2 & 0 & 1 & 2 & 4 & 0 \\ 0 & 5 & 6 & 0 & 7 & 1 & 4 & 0 \\ 0 & 0 & 5 & 4 & 9 & 6 & 0 & 0 \\ 0 & 4 & 6 & 2 & 4 & 0 & 5 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

The padded input A_p will now be of size 8×8.

[Reason :- As $f = 3$; $p = \frac{3-1}{2} = 1$ [we pad by 1]]

$n = 8$

$f = 3$

$s = 1$

$$\text{output size} = \left\lfloor \frac{8-3}{1} + 1 \right\rfloor \times \left\lfloor \frac{8-3}{1} + 1 \right\rfloor$$

$$= 6 \times 6 = \text{same as input}$$

$$\left\lfloor \frac{n-f}{s} + 1 \right\rfloor = 6$$

$n = 6$

$n + 2p = 6 + 2(1) = 8$

For an input image of $n \times n$ size, a kernel of $f \times f$ will give a padding of $p = \frac{f-1}{2}$. The size of output image after applying kernel on padded input is $(\underbrace{n + 2p} - f + 1) \times (n + 2p - f + 1)$

Input image
(B, W, H, C)

↓

Kernel
$K_W \times K_H$

↓

output.
B, W', H', C_{out}

We have a batch of input images each of width W, height H & channel C.

We consider C_{out} kernels of size $K_W \times K_H$
↳ Number of kernels

output is of size

$$W' = \left\lfloor \frac{W - K_W}{s} + 1 \right\rfloor$$

$$H' = \left\lfloor \frac{H - K_H}{s} + 1 \right\rfloor$$

C_{out} = Number of kernels

Dilated Convolution: The convolution performed by inflating the kernel by inserting spaces between the kernel elements

We have input matrix of size $n \times n$

Kernel matrix of size $f \times f$

padding done of size p

stride of s

dilation rate $= d$

$$\left\lfloor \frac{n+2p-f}{s} + 1 \right\rfloor$$

$$-(f-1)(d-1)$$

The size of output $= \left\lfloor \frac{n+2p-f-(f-1)(d-1)}{s} + 1 \right\rfloor \times$

$$\left\lfloor \frac{n+2p-f-(f-1)(d-1)}{s} + 1 \right\rfloor$$

$$p = \frac{f-1}{2} = 1$$

We will consider our usual matrix A with padding $p=1$,

$n = 6$ [$\because$ A was of size 6×6], $f = 3$ [$\because$ K was 3×3 matrix]

We consider a stride $s = 1$ & dilation rate of 3 (i.e. $d = 3$)

We have size of output $= \left\lfloor \frac{6+2(1)-3-(3-1)(3-1)}{1} + 1 \right\rfloor \times \left\lfloor \frac{6+2(1)-3-(3-1)(3-1)}{1} + 1 \right\rfloor$

$= 2 \times 2$

We will have a 2×2 matrix as output

We will have padded matrix as below -

$$A_p = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 3 & 2 & 7 & 1 & 6 & 3 & 0 \\ 0 & 4 & 1 & 1 & 0 & 9 & 8 & 0 \\ 0 & 8 & 2 & 0 & 1 & 2 & 4 & 0 \\ 0 & 5 & 6 & 0 & 7 & 1 & 4 & 0 \\ 0 & 0 & 5 & 4 & 9 & 6 & 0 & 0 \\ 0 & 4 & 6 & 2 & 4 & 0 & 5 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix} = (a'_{ij})_{8\times 8}$$

$d = 3$

To calculate h_{11}, we have the below submatrix

$$A_p^{(11)} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 4 \\ 0 & 2 & 5 \end{pmatrix} = \begin{pmatrix} a'_{11} & a'_{14} & a'_{17} \\ a'_{41} & a'_{44} & a'_{47} \\ a'_{71} & a'_{74} & a'_{77} \end{pmatrix}$$

$$A_p = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 3 & 2 & 7 & 1 & 6 & 3 & 0 \\ 0 & 4 & 1 & 1 & 0 & 9 & 8 & 0 \\ 0 & 8 & 2 & 0 & 1 & 2 & 4 & 0 \\ 0 & 5 & 6 & 0 & 7 & 1 & 4 & 0 \\ 0 & 0 & 5 & 4 & 9 & 6 & 0 & 0 \\ 0 & 4 & 6 & 2 & 4 & 0 & 5 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

$$\begin{aligned} h_{11} = \; & 0\times 1 + 0\times 2 + 0\times 2 + \\ & 0\times 2 + 0\times 1 + 4\times 2 + \\ & 0\times 2 + 2\times 2 + 5\times 1 \\ = \; & 17 \end{aligned}$$

To calculate h_{12}, we have below submatrix

$$A_p^{(12)} = \begin{pmatrix} 0 & 0 & 0 \\ 8 & 1 & 0 \\ 4 & 4 & 0 \end{pmatrix} = \begin{pmatrix} a'_{12} & a'_{15} & a'_{18} \\ a'_{42} & a'_{45} & a'_{48} \\ a'_{72} & a'_{75} & a'_{78} \end{pmatrix}$$

$$h_{12} = 0\times 1 + 0\times 2 + 0\times 2 + 8\times 2 + 1\times 1 + 0\times 2 + 4\times 2 + 4\times 2 + 0\times 1$$

$$= 33$$

$$A_\beta = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 3 & 2 & 7 & 1 & 6 & 3 & 0 \\ 0 & 4 & 1 & 1 & 0 & 9 & 8 & 0 \\ 0 & 8 & 2 & 0 & 1 & 2 & 4 & 0 \\ 0 & 5 & 6 & 0 & 7 & 1 & 4 & 0 \\ 0 & 0 & 5 & 4 & 9 & 6 & 0 & 0 \\ 0 & 4 & 6 & 2 & 4 & 0 & 5 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

To calculate h_{21}, we have the below submatrix

$$A_\beta^{(21)} = \begin{pmatrix} 0 & 7 & 3 \\ 0 & 0 & 4 \\ 0 & 0 & 0 \end{pmatrix} = \begin{pmatrix} a'_{21} & a'_{24} & a'_{27} \\ a'_{51} & a'_{54} & a'_{57} \\ a'_{81} & a'_{84} & a'_{87} \end{pmatrix}.$$

$$h_{21} = 0\times 1 + 7\times 2 + 3\times 2 + 0\times 2 + 0\times 1 + 4\times 2 + 0\times 2 + 0\times 2 + 0\times 1$$

$$= 14 + 6 + 8 = 28$$

To calculate h_{22}, we have the below submatrix

$$A_\beta^{(22)} = \begin{pmatrix} a'_{22} & a'_{25} & a'_{28} \\ a'_{52} & a'_{55} & a'_{58} \\ a'_{82} & a'_{85} & a'_{88} \end{pmatrix} = \begin{pmatrix} 3 & 1 & 0 \\ 5 & 7 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

$$h_{22} = 3\times 1 + 1\times 2 + 0\times 2 + 5\times 2 + 7\times 1 + 0\times 2 + 0\times 2 + 0\times 2 + 0\times 1 = 22.$$

$$H = \begin{pmatrix} 17 & 33 \\ 28 & 22 \end{pmatrix} \Rightarrow \text{output matrix}$$

$$H = \begin{pmatrix} 17 & 33 \\ 28 & 22 \end{pmatrix}$$

Dilated convolution numerical example

Consider a 7x7 input and a 3x3 convolution kernel, as shown below. Point (i, j) corresponds to the i-th row (ranging 1, 2, ...) from top to bottom and j-th column (ranging 1,2, ...) from left to right.

0.5	1	1	1	1	1	1
1	0.5	0	0	0	0	0
0	0	0.5	1	1	1	1
1	1	1	0.5	0	0	0
0	0	0	0	0.5	1	1
1	1	1	1	1	0.5	0
0	0	0	0	0	0	0.5

$\rightarrow A$

1	0	-1
0	4	0
-1	0	-1

$\rightarrow K$

$p=0$ $d=3$

Assuming zero padding and 3-dilated convolution, what will be value at (1,1) in the output plane?

We have an input matrix A as below -

$$A = \begin{pmatrix} 0.5 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 0.5 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0.5 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 0.5 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0.5 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 & 0.5 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0.5 \end{pmatrix}$$

The Kernel is given as.

$$K = \begin{pmatrix} 1 & 0 & -1 \\ 0 & 4 & 0 \\ -1 & 0 & -1 \end{pmatrix}$$

We have $\underline{n = 7}$, $\underline{f = 3}$, $\underline{p = 0}$, $\underline{d = 3}$, $\underline{\underline{s = 1}}$ (assumed)

Output size $= \left\lfloor \frac{n + 2p - f - (f-1)(d-1)}{s} + 1 \right\rfloor \times \left\lfloor \frac{n + 2p - f - (f-1)(d-1)}{s} + 1 \right\rfloor$

$$= \left\lfloor \frac{7 + 2(0) - 3 - (3-1)(3-1)}{1} + 1 \right\rfloor \times \left\lfloor \frac{7 + 2(0) - 3 - (3-1)(3-1)}{1} + 1 \right\rfloor$$

$$= \underline{\underline{1 \times 1}}$$

We will just get a matrix $H = (\underline{h_{11}})$

The submatrix to calculate dilated convolution is

$$A'_{11} = \begin{pmatrix} 0.5 & 1 & 1 \\ 1 & 0.5 & 0 \\ 0 & 0 & 0.5 \end{pmatrix} = \begin{pmatrix} a_{11} & a_{14} & a_{17} \\ a_{41} & a_{44} & a_{47} \\ a_{71} & a_{74} & a_{77} \end{pmatrix}$$

$$h_{11} = 0.5 \times 1 + 1 \times 0 + 1 \times (-1) + 1 \times 0 + 0.5 \times 4 + 0 \times 0 + 0 \times (-1) + 0 \times 0 + 0.5 \times (-1)$$

$$= 1$$

Transposed Convolution

We have an input matrix $A = \begin{pmatrix} 1 & 3 & 2 \\ 2 & 5 & 6 \\ 7 & 1 & 3 \end{pmatrix}$ ✓

$n = 3$

$f = 2$

$s = 1$

We have an input matrix $X = \begin{pmatrix} 1 & 0 & 4 \\ 2 & 5 & 6 \\ 7 & 1 & 3 \end{pmatrix}$ ✓

$s = 1$

$$\boxed{n+(f-1) \times n+(f-1)}$$

$3 + (2-1)$

$= 4 \times 4$

The Kernel $K = \begin{pmatrix} 0 & 1 \\ 3 & 2 \end{pmatrix}$. We will calculate transposed convolution with stride 1 as below.

$$A_1 = \begin{pmatrix} 1\times0 & 1\times1 & 0 & 0 \\ 3\times1 & 2\times1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 3 & 2 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

$$A_2 = \begin{pmatrix} 0 & 3\times0 & 3\times1 & 0 \\ 0 & 3\times3 & 3\times2 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 & 3 & 0 \\ 0 & 9 & 6 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

$$A_3 = \begin{pmatrix} 0 & 0 & 2\times0 & 2\times1 \\ 0 & 0 & 2\times3 & 2\times2 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 & 2 \\ 0 & 0 & 6 & 4 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

$$A_4 = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 2\times0 & 2\times1 & 0 & 0 \\ 2\times3 & 2\times2 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 6 & 4 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

$$A_5 = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 5\times0 & 5\times1 & 0 \\ 0 & 5\times3 & 5\times2 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 5 & 0 \\ 0 & 15 & 10 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

$$A_6 = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 6\times0 & 6\times1 \\ 0 & 0 & 6\times3 & 6\times2 \\ 0 & 0 & 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 6 \\ 0 & 0 & 18 & 12 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

$$A_7 = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0\times7 & 1\times7 & 0 & 0 \\ 3\times7 & 2\times7 & 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 7 & 0 & 0 \\ 21 & 14 & 0 & 0 \end{pmatrix}$$

$$A_8 = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 3 & 2 & 0 \end{pmatrix}$$

$$A_9 = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 3 \\ 0 & 0 & 9 & 6 \end{pmatrix}$$

Transposed convolution output $= A_1 + A_2 + \cdots + A_9$

$$= \begin{pmatrix} 0 & 1 & 3 & 2 \\ 3 & 13 & 17 & 10 \\ 6 & 26 & 29 & 15 \\ 21 & 17 & 11 & 6 \end{pmatrix} \checkmark \qquad \underline{\underline{4\times4}}$$

If we consider a stride of 2, our output matrix will be of shape 6X6. We get following as an output matrix

$$\begin{pmatrix} 0 & 1 & 0 & 3 & 0 & 2 \\ 3 & 2 & 9 & 6 & 6 & 4 \\ 0 & 2 & 0 & 5 & 0 & 6 \\ 6 & 4 & 15 & 10 & 18 & 12 \\ 0 & 7 & 0 & 1 & 0 & 3 \\ 21 & 14 & 3 & 2 & 9 & 6 \end{pmatrix}$$

6×6

$$\boxed{S \times (n-1) + f}$$

$$2(3-1)+2$$

$$= 6.$$

$S = 2$

$$A = \begin{pmatrix} 1 & 3 & 2 \\ 2 & 5 & 6 \\ 7 & 1 & 3 \end{pmatrix} \qquad K = \begin{pmatrix} 0 & 1 \\ 3 & 2 \end{pmatrix}.$$