

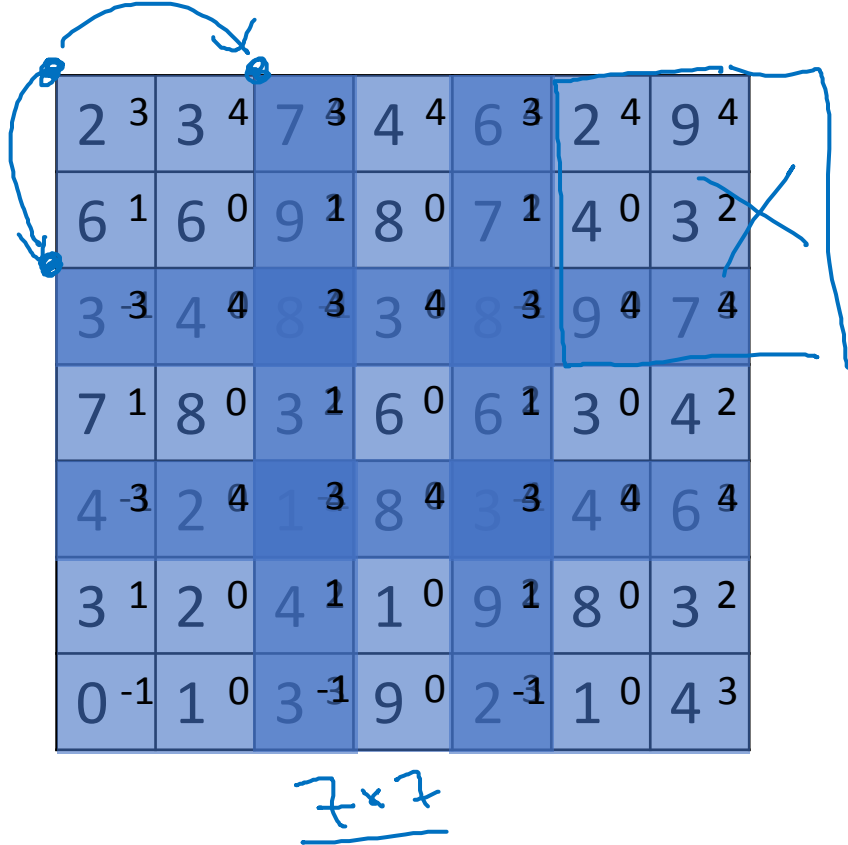


deeplearning.ai

Convolutional Neural Networks

Strided convolutions

Strided convolution




Then it turns out that that's the convention, that your 3x3 filter must lie entirely within your image plus padding region before there's a corresponding output generated.

3	4	4
1	0	2
-1	0	3

3x3

Stride = 2

=



91	100	83
69	91	127
44	72	74

3x3

$\lfloor z \rfloor = \text{floor}(z)$

$n \times n$ * $f \times f$
padding p stride s
 $s = 2$

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

round this down
$$\frac{7 + 0 - 3}{2} + 1 = \frac{4}{2} + 1 = 3$$

Summary of convolutions

$n \times n$ image $f \times f$ filter

padding p stride s

Output Size:

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

Technical note on cross-correlation vs. convolution

Convolution in math textbook:

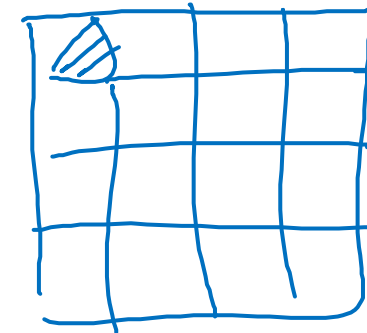
2 ⁷	3 ²	7 ⁵	4	6	2
6 ⁹	6 ⁰	9 ⁴	8	7	4
3 ⁻¹	4 ¹	8 ³	3	8	9
7	8	3	6	6	3
4	2	1	8	3	4
3	2	4	1	9	8

*

3	4	5
1	0	2
-1	9	7

7	2	5
9	0	4
-1	1	3

Taking the 3X3 filter and mirroring it both on the vertical and horizontal axes



$$(A * B) * C = A * (B * C)$$

Associativity