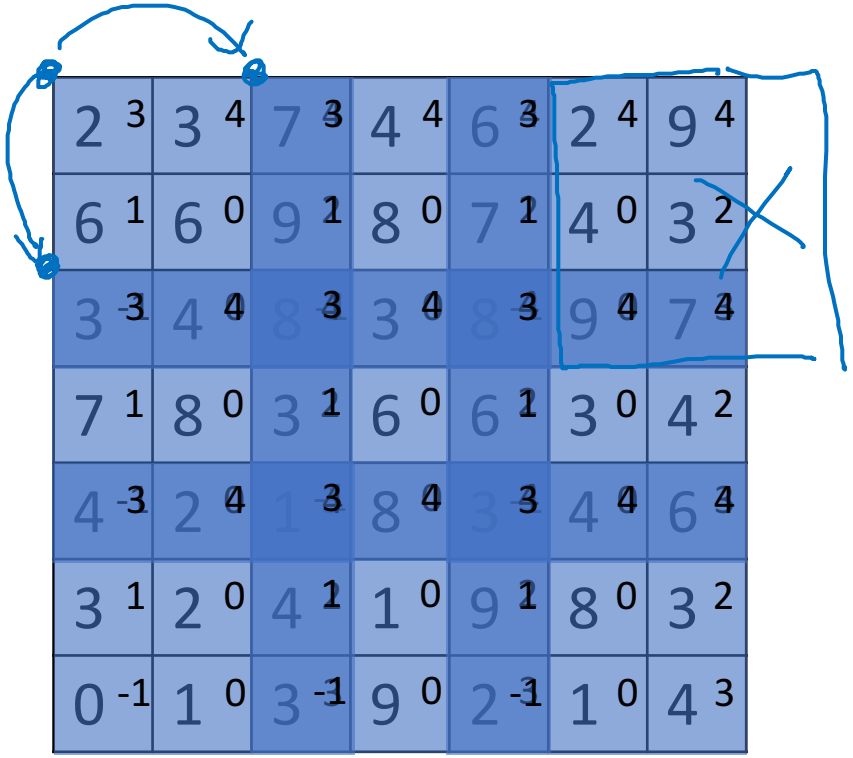
deeplearning.ai

Convolutional Neural Networks

Strided convolutions

Strided convolution



7×7

*

3	4	4
1	0	2
-1	0	3

3×3

=

91	100	83
69	91	127
44	72	74

3×3

stride = 2

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

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

padding p

stride s

$s = 2$

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

Andrew Ng

Summary of convolutions

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

padding p $\quad$ stride s

Output size:

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

Technical note on cross-correlation vs. convolution

Convolution in math textbook:

2	3	7	4	6	2
6	6	9	8	7	4
3	4	8	3	8	9
7	8	3	6	6	3
4	2	1	8	3	4
3	2	4	1	9	8

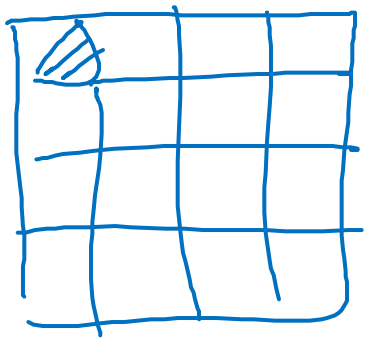
시그널 프로세싱에서 컨벌루젼은 product랑 summing 하기 전에 필터를 vertical/horizontal axis로 flip해주는 과정이 있다. 근데 우리는 컨벌루젼을 정의할 때 이 mirroring operation을 스킵했다.

정확히 말하면 우리가 한건 convolution이 아니라 cross-correlation이었음. 근데 딥러닝에서는 그냥 구분없이 쓴다.

*

3	4	5
1	0	2
-1	9	7

=



7	2	5
9	0	4
-1	1	3

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

flipping을 적용하는 컨벌루션은 이러한 associativity 특성을 만족하며 이 특성은 어떤 시그널 프로세싱 애플리케이션에 대해 nice한 경우가 있어서 많이 쓰인다. 근데 딥러닝에서는 상관없음.

Andrew Ng