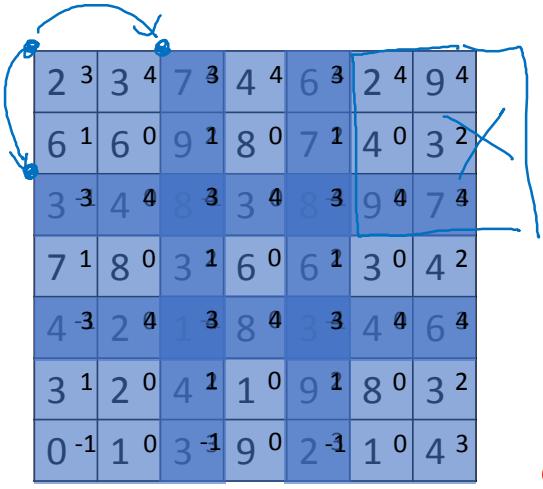
deeplearning.ai

# 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

→ Convention: filter must lie entirely within your image or the image plus padding region

stride = 2

$\lfloor z \rfloor$ = floor($z$)

$n \times n$ * $f \times f$

padding $p$ stride $s$

$s = 2$

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

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

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

in video

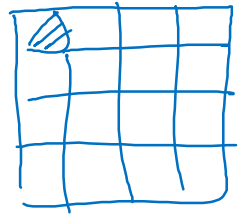
## Convolution in math textbook:

| | | | | | |
|---|---|---|---|---|---|
| 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 |

*

| | | |
|---|---|---|
| 3 | 4 | 5 |
| 1 | 0 | 2 |
| -1 | 9 | 7 |

=



| | | |
|---|---|---|
| 7 | 9 | -1 |
| 2 | 0 | 1 |
| 5 | 4 | 3 |

flipped

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