Convolution operator

$3 \cdot 1 + 1 \cdot 1 + 2 \cdot 1 + 0 \cdot 0 + 5 \cdot 0 + 7 \cdot 0 + 1 \cdot -1 + 8 \cdot -1 + 2 \cdot -1 = -5$

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

6 × 6

convolution

$*$

1	0	-1
1	0	-1
1	0	-1

3 × 3

filter / kernel

$=$

-5	-4	0	8
-10	-2	2	3
0	-2	-4	-7
-3	-2	-3	-16

4 × 4

Vertical edge detection with convolution

1	1	1	0	0	0
1	1	1	0	0	0
1	1	1	0	0	0
1	1	1	0	0	0
1	1	1	0	0	0
1	1	1	0	0	0

6 × 6

$*$

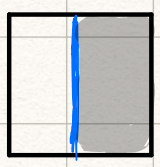
1	0	-1
1	0	-1
1	0	-1

3 × 3

$=$

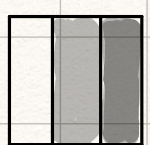
0	3	3	0
0	3	3	0
0	3	3	0
0	3	3	0

4 × 4



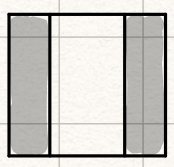
image

$*$



filter

$=$



edge region

image

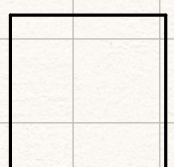
$*$

w_1	w_2	w_3
w_4	w_5	w_6
w_7	w_8	w_9

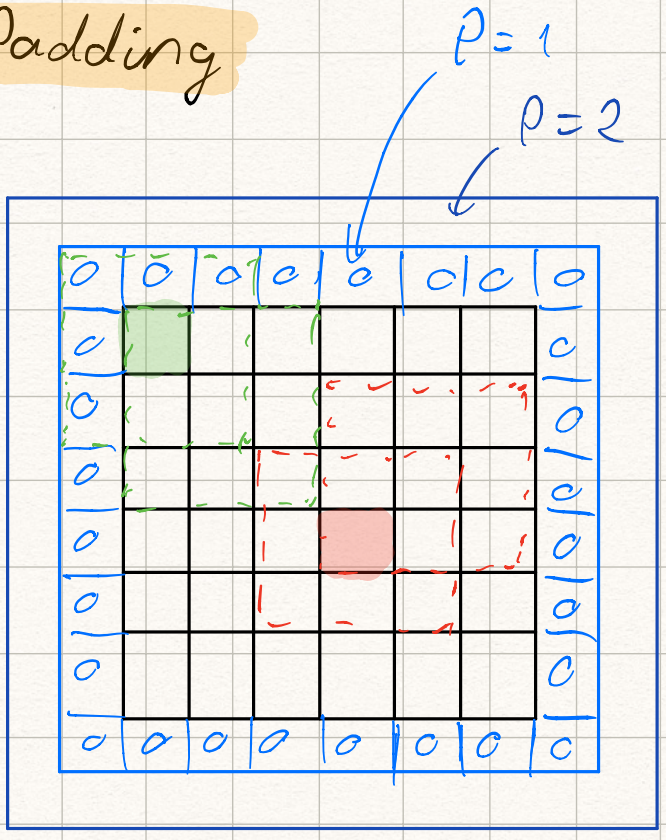
learnable parameters

filter

$=$



Padding



6 × 6 => 8 × 6

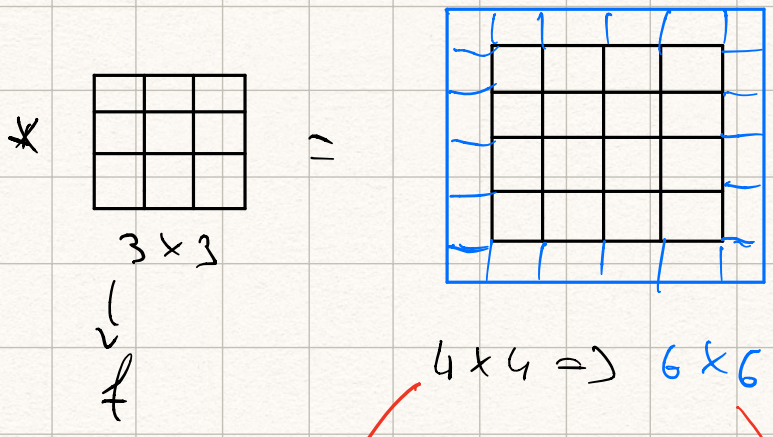
n

p = 1

padding size

Convolution

- shrinking output
- throw away edge information



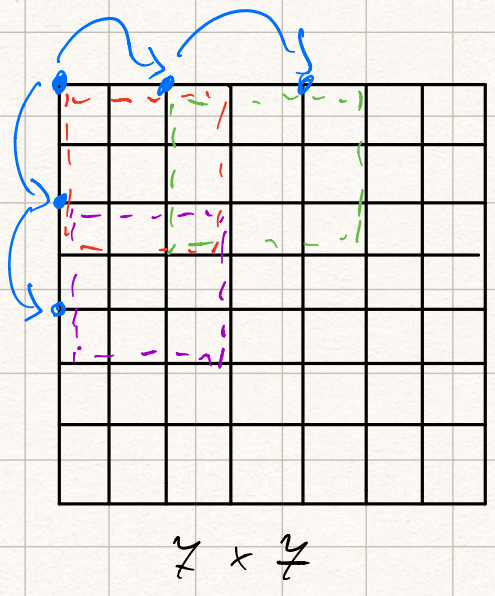
$(n-f+1) \times (n-f+1)$

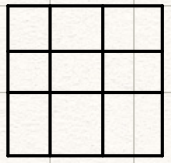
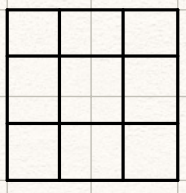
$(n+2p-f+1) \times (n+2p-f+1)$

Padding	image		filter		output
Valid	$n \times n$	*	$f \times f$	=	$n-f+1 \times n-f+1$
	6×6	*	3×3	=	4×4
Same	$n \times n$	*	$f \times f$	=	$n+2p-f+1 \times n+2p-f+1$
	6×6	*	3×3	=	6×6

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

Strided convolution



$*$  $=$ 

3 × 3 3 × 3

Stride = 2

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

padding p stride s

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