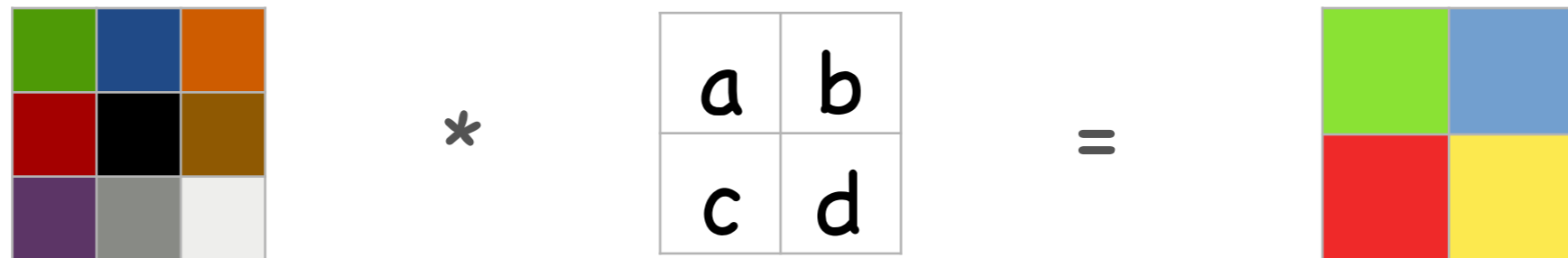


# Convolutional operators and their structure

© 2019 Philipp Krähenbühl and Chao-Yuan Wu

# Output size

- Input:  $\mathbf{X} \in \mathbb{R}^{H \times W \times C_1}$
- Kernel:  $\mathbf{w} \in \mathbb{R}^{h \times w \times C_1 \times C_2}$
- Output:  $\mathbf{Z} \in \mathbb{R}^{(H-h+1) \times (W-w+1) \times C_2}$



# Padding

- Add  $p_w, p_h$  zeros in each dimension
- Input:  $\mathbf{X} \in \mathbb{R}^{H \times W \times C_1}$
- Kernel:  $\mathbf{w} \in \mathbb{R}^{h \times w \times C_1 \times C_2}$
- Output:  $\mathbf{Z} \in \mathbb{R}^{(H-h+2p_h+1) \times (W-w+2p_w+1) \times C_2}$

0	0	0	0	0
0	green	blue	orange	0
0	red	black	brown	0
0	purple	gray	light gray	0
0	0	0	0	0

\*

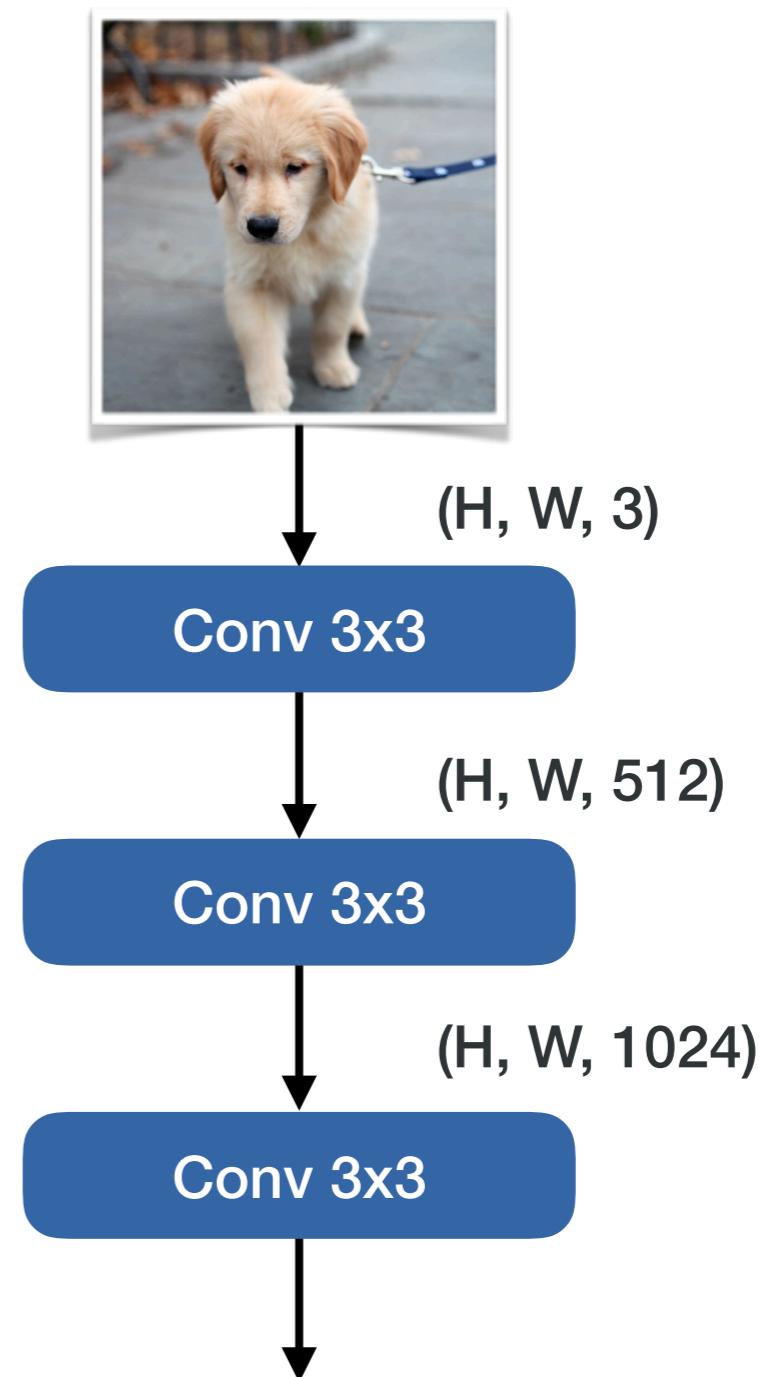
a	b
c	d

=

green	blue	gray	blue
red	purple	gray	green
purple	red	orange	brown
red	orange	yellow	light gray

# Output resolution

- High output resolution
- Slow computation



# Striding

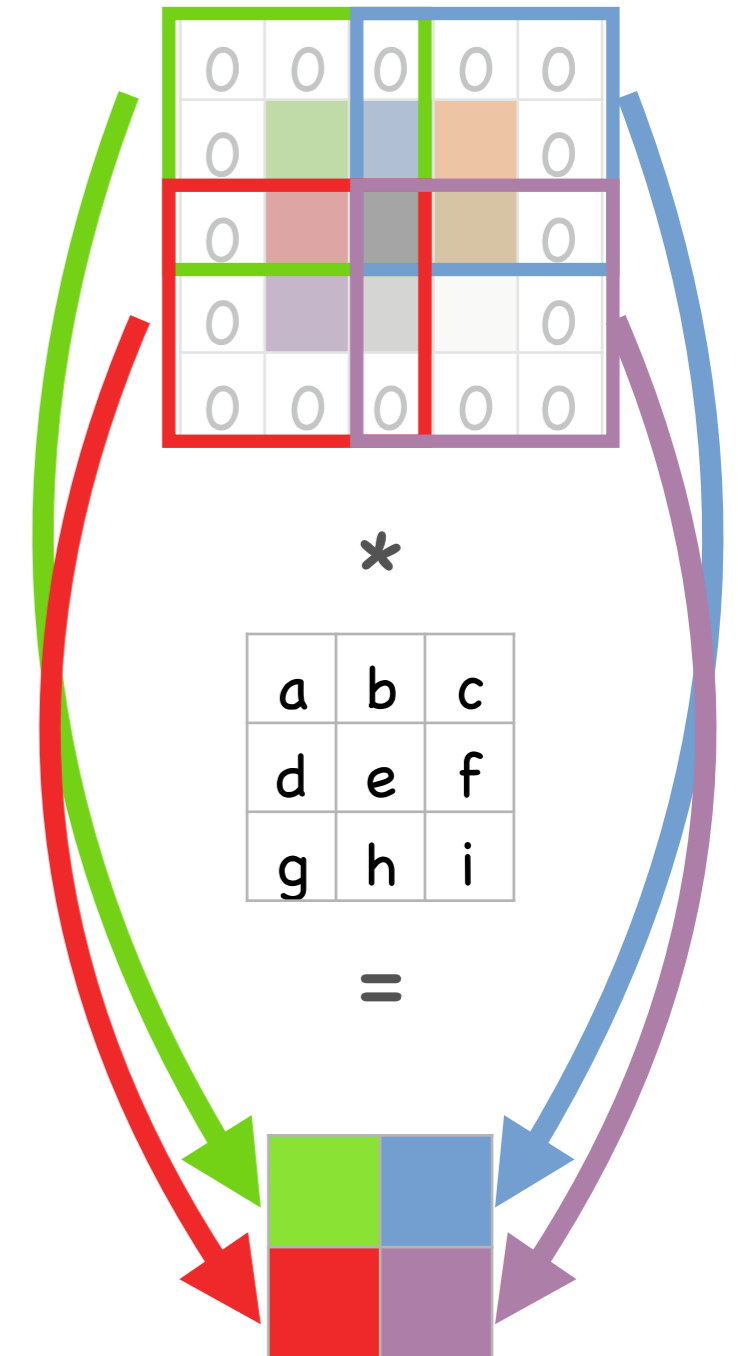
- Only compute every n-th output:  $s_w, s_h$

- Input:  $\mathbf{X} \in \mathbb{R}^{H \times W \times C_1}$

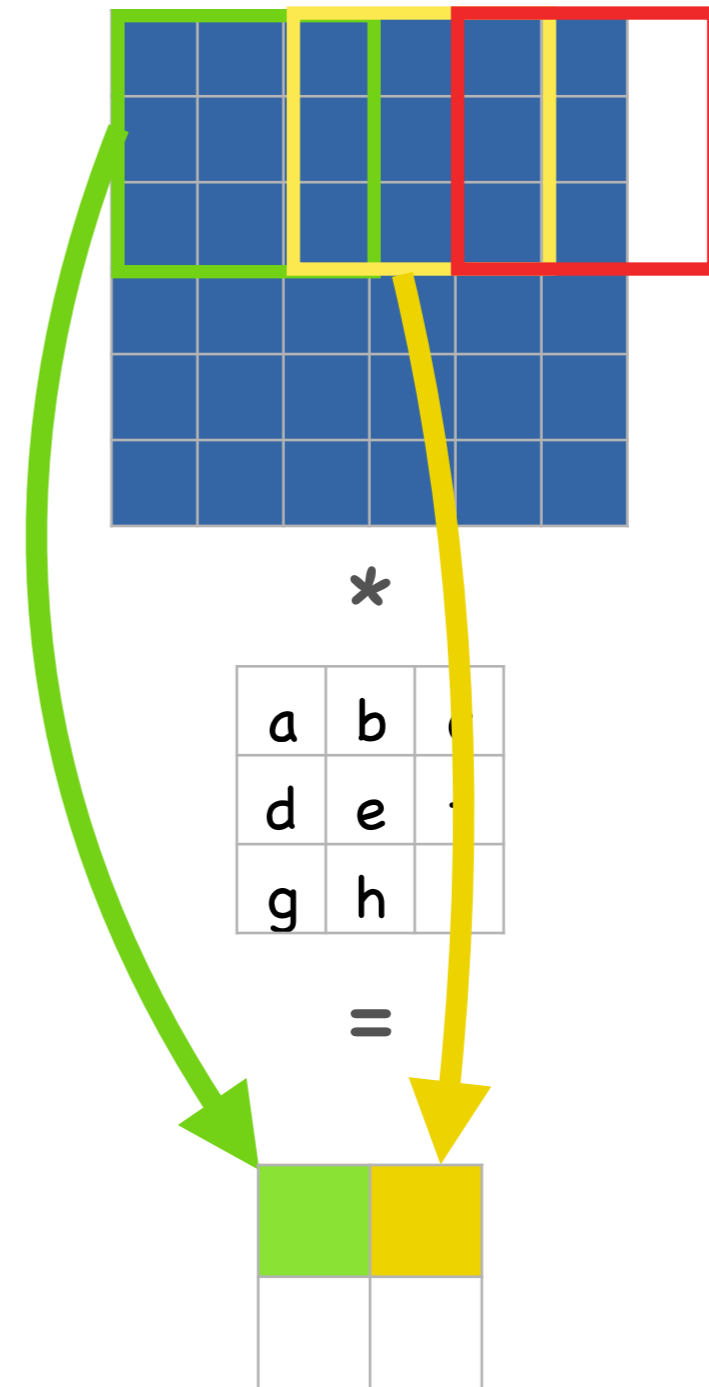
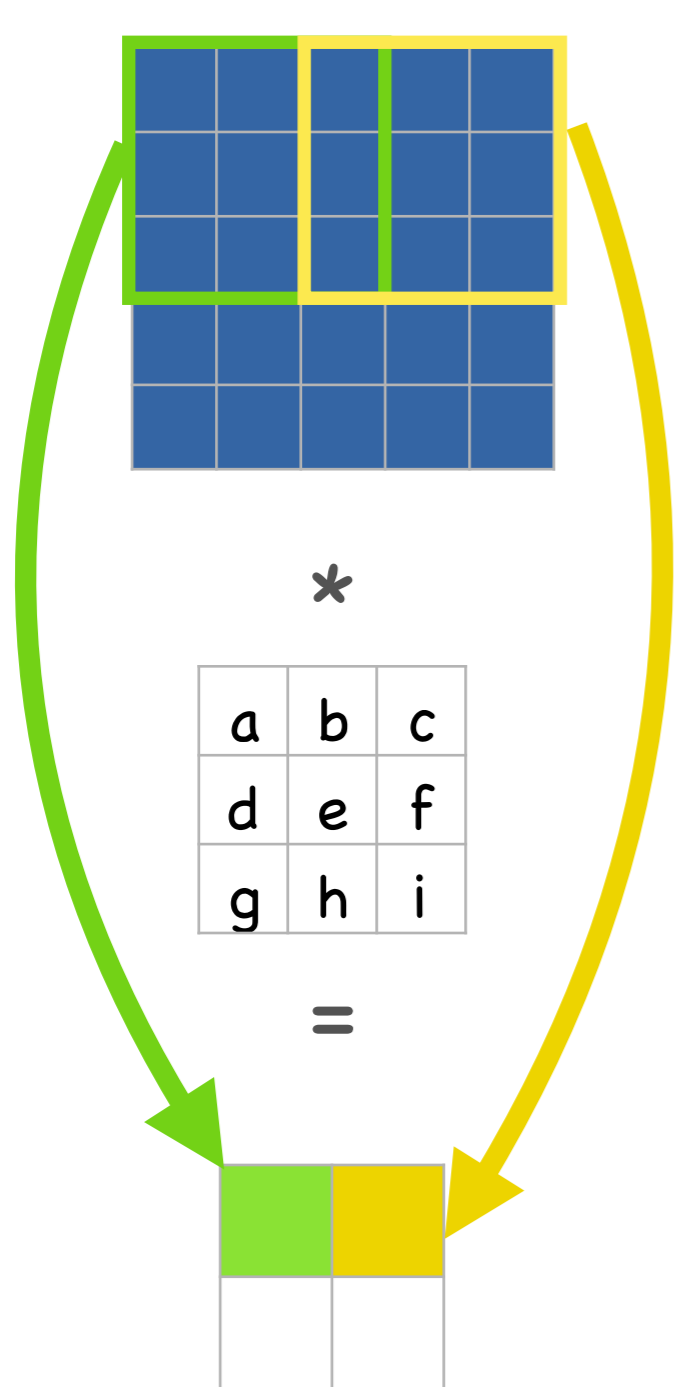
- Kernel:  $\mathbf{w} \in \mathbb{R}^{h \times w \times C_1 \times C_2}$

- Output:

$$\mathbf{Z} \in \mathbb{R}^{\left(\frac{H-h+2p_h}{s_h}+1\right) \times \left(\frac{W-w+2p_w}{s_w}+1\right) \times C_2}$$

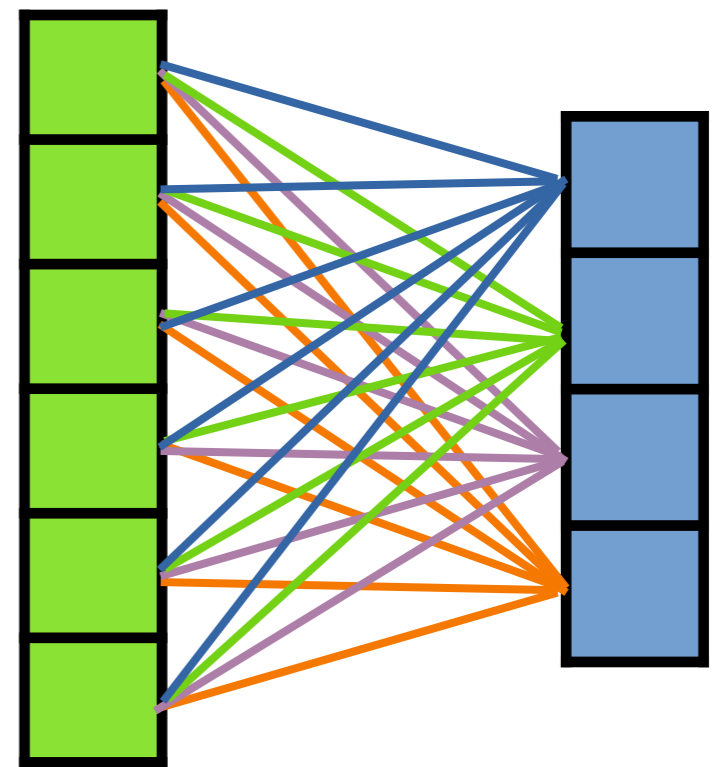


# Output size with striding



# Parameters

- Every input channel  $C_1$  is connected to every output channel  $C_2$

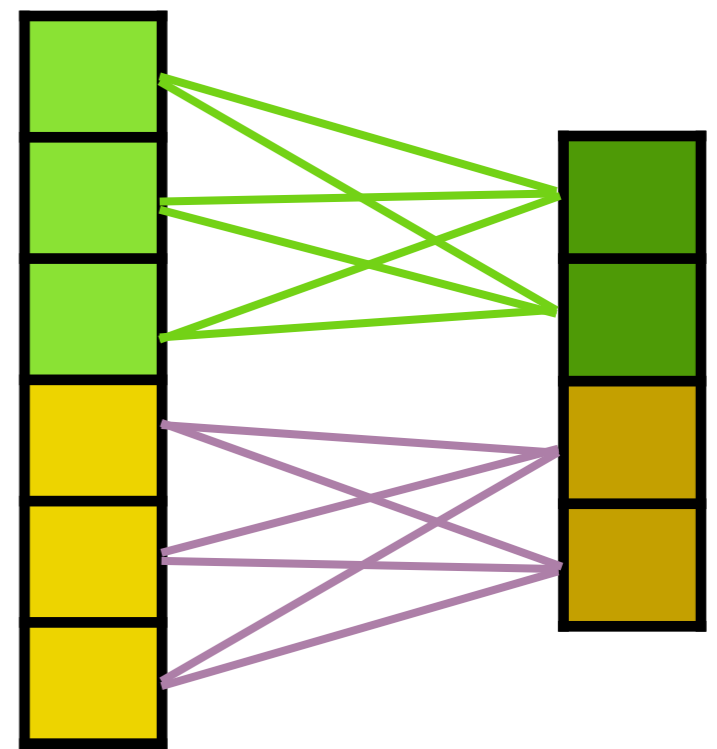


$$C_1 = 6$$

$$C_2 = 4$$

# Grouping

- Split channels into  $g$  groups
- Reduce parameters and computation by factor  $g$



$$C_1 = 6$$

$$C_2 = 4$$

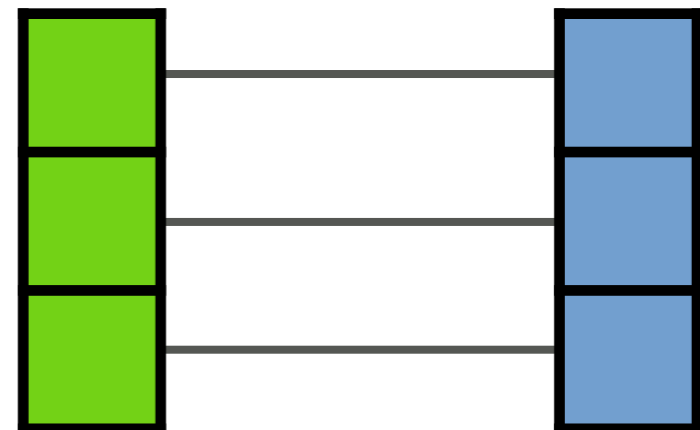


# Depthwise convolution

- Special grouping

- $C_1 = g$

- $C_2 = g$

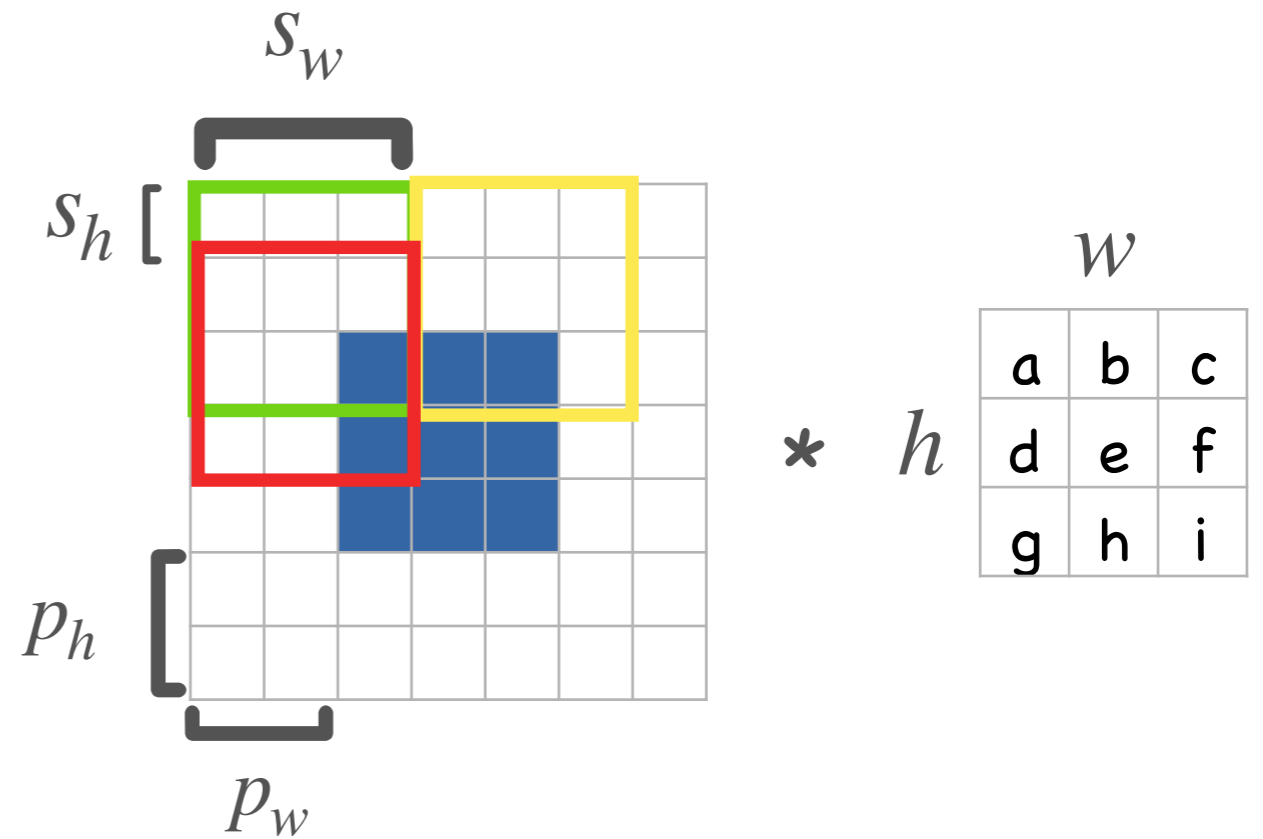


$$C_1 = 3$$

$$C_2 = 3$$

# Hyper-parameters of convolutions

- Kernel size:  $w \times h$
- Padding:  $p_w, p_h$
- Stride:  $s_w, s_h$



# Convolutional operators

- Run arbitrary operation  $f(\mathbf{x})$  "over" image

$f(\mathbf{x})$



\*

a	b	c
d	e	f
g	h	i

=

