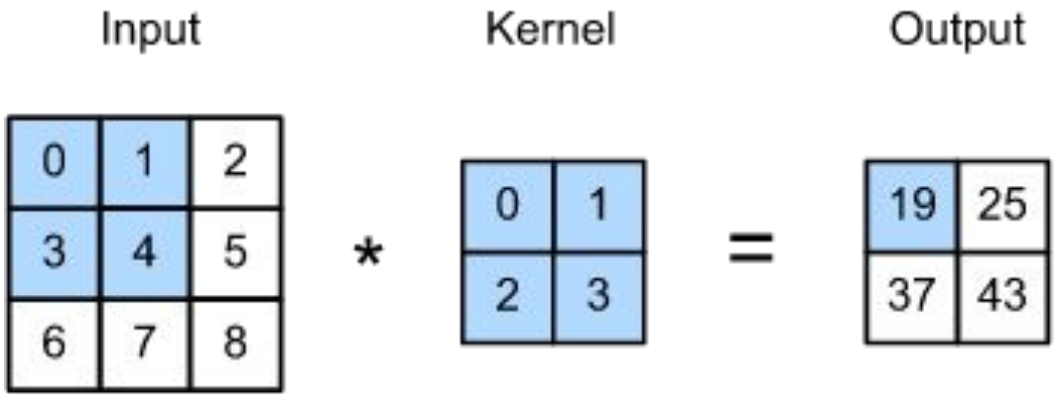
Convolution in Deep Learning

Convolution

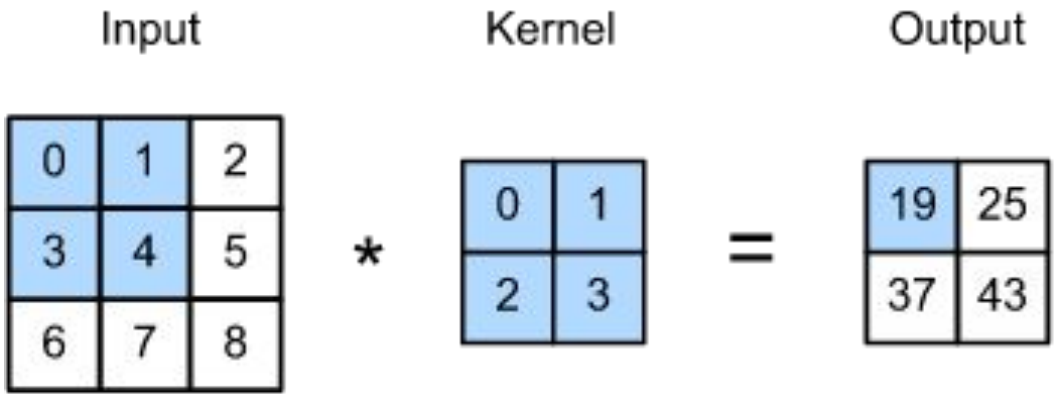
- Convolution is the two-dimensional cross-correlation operation.
- (0 x 0) + (1 x 1) + (3 x 2) + (4 x 3) = 19



https://d2l.ai/chapter_convolutional-neural-networks/conv-layer.html

Convolution

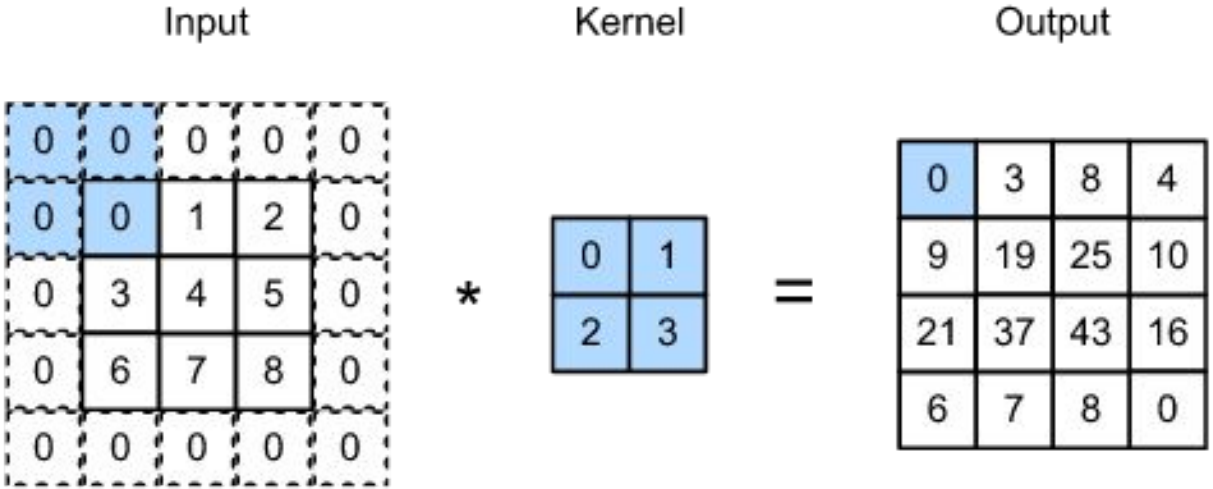
- Input size: m_h x m_w kernel size: z_h x z_w
- Output size: (m_h - z_h + 1) x (m_w - z_w + 1)



https://d2l.ai/chapter_convolutional-neural-networks/conv-layer.html

Padding

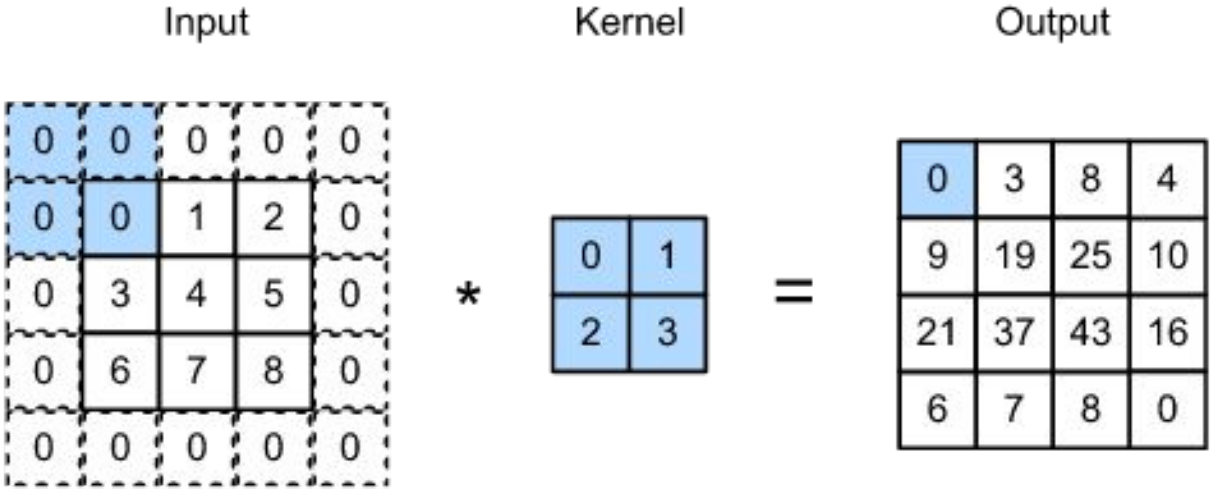
- Padding adds pixels around the boundary of the input.
- (0 x 0) + (0 x 1) + (0 x 2) + (0 x 3) = 0



https://d2l.ai/chapter_convolutional-neural-networks/padding-and-strides.html

Padding

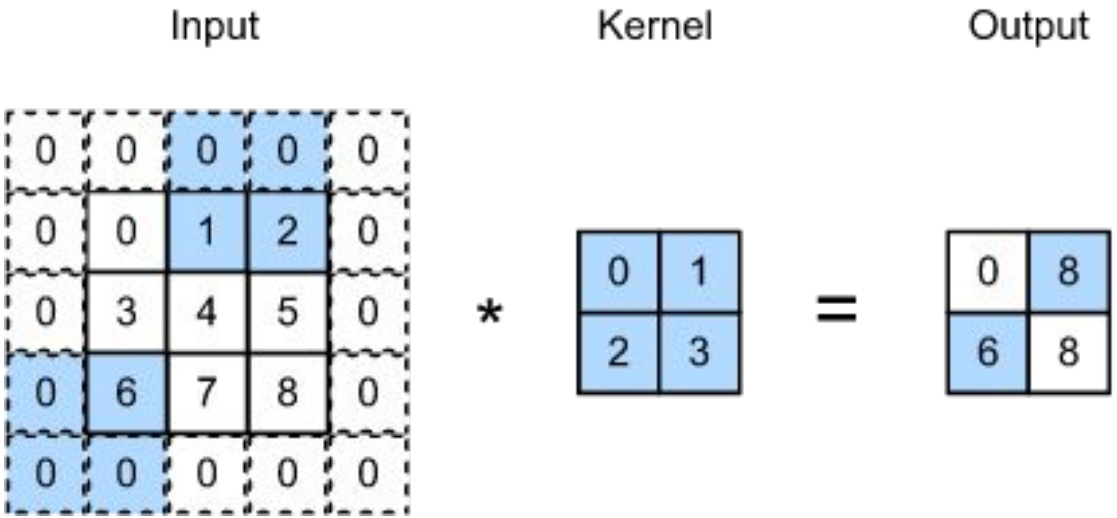
- Input size: m_h x m_w kernel size: z_h x z_w padding size: t_h x t_w
- Output size: $(m_h - z_h + t_h + 1) \times (m_w - z_w + t_w + 1)$



https://d2l.ai/chapter_convolutional-neural-networks/padding-and-strides.html

Stride

- Stride downsample the size of the input.
- Stride size: 3 x 2

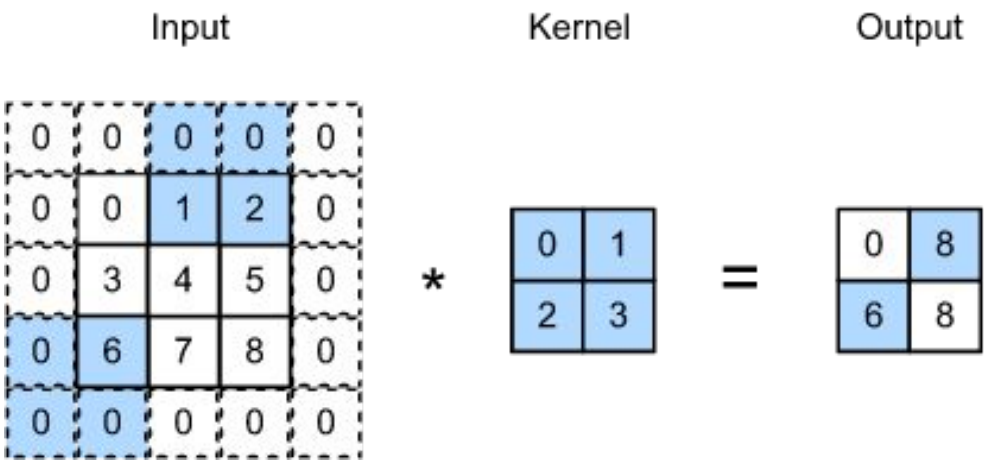


https://d2l.ai/chapter_convolutional-neural-networks/padding-and-strides.html

Stride

- Input size: m_h x m_w kernel size: z_h x z_w padding size: t_h x t_w Stride size: d_h x d_w
- Output size: $[(m_h - z_h + t_h + d_h)/d_h] \times [(m_w - z_w + t_w + d_w)/d_w]$

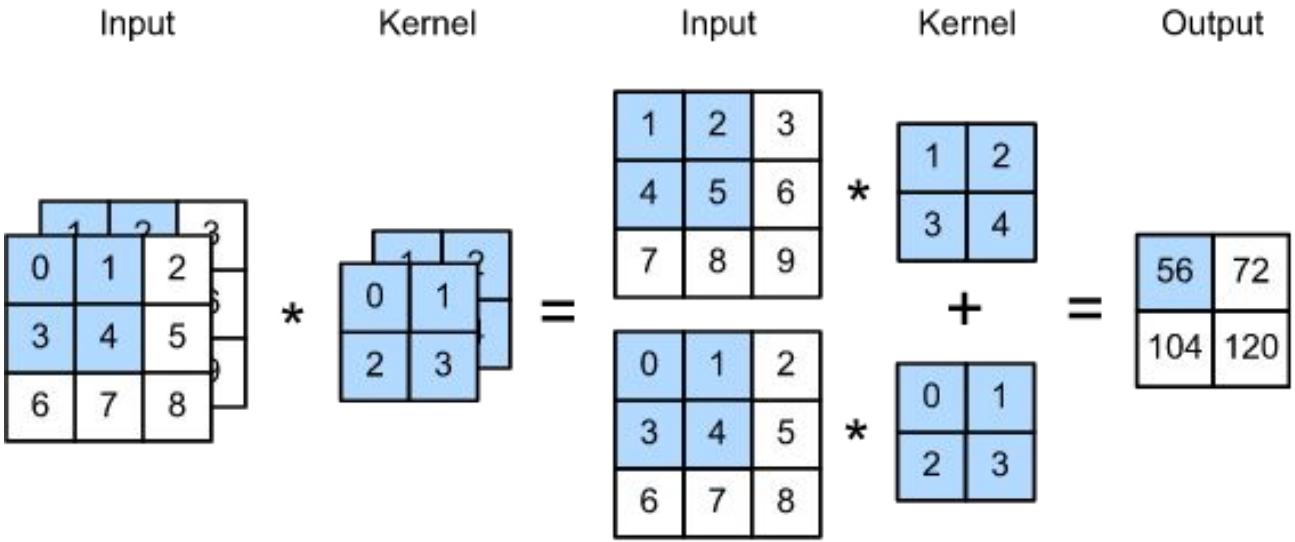
$$= [(m_h - z_h + t_h)/d_h + 1] \times [(m_w - z_w + t_w)/d_w + 1]$$



https://d2l.ai/chapter_convolutional-neural-networks/padding-and-strides.html

Convolution on multiple channels

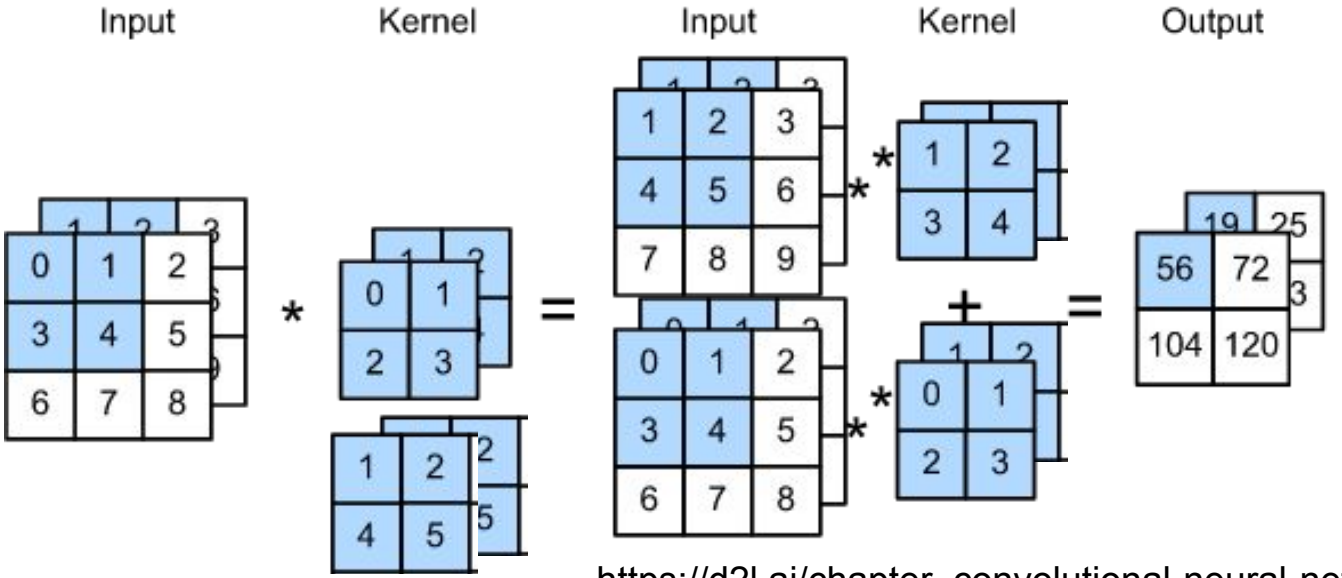
- Input size: 3 x 3 x 2 Kernel size: 2 x 2 x 2
- Output size: 2 x 2 x 1



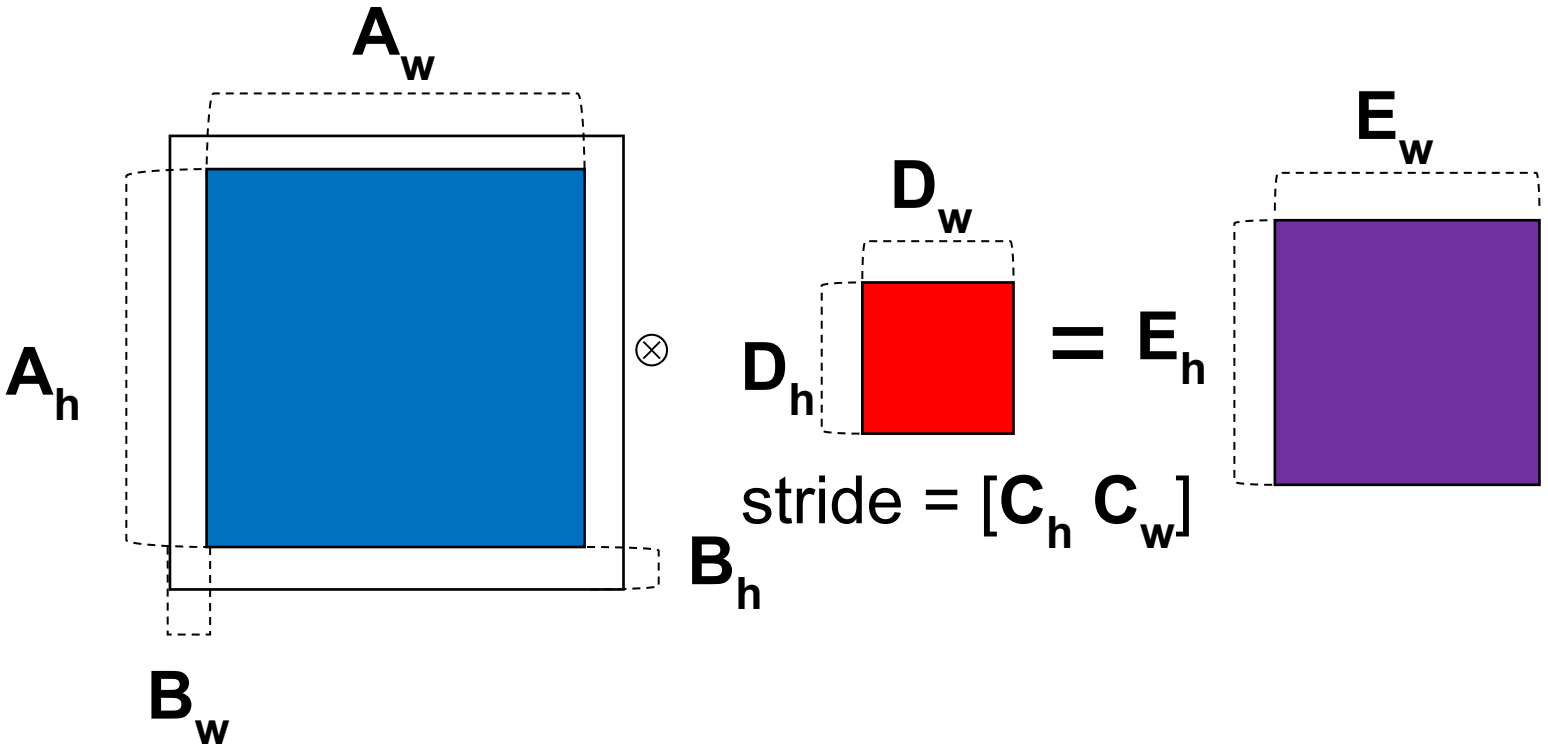
https://d2l.ai/chapter_convolutional-neural-networks/channels.html

Convolution on multiple channels

- Input size: 3 x 3 x 2 Kernel size: 2 x 2 x 2 x **2**
- Output size: 2 x 2 x **2**



https://d2l.ai/chapter_convolutional-neural-networks/channels.html


A_w
A_h
B_w
B_h
D_w
D_h
E_w
E_h
stride = [C_h C_w]