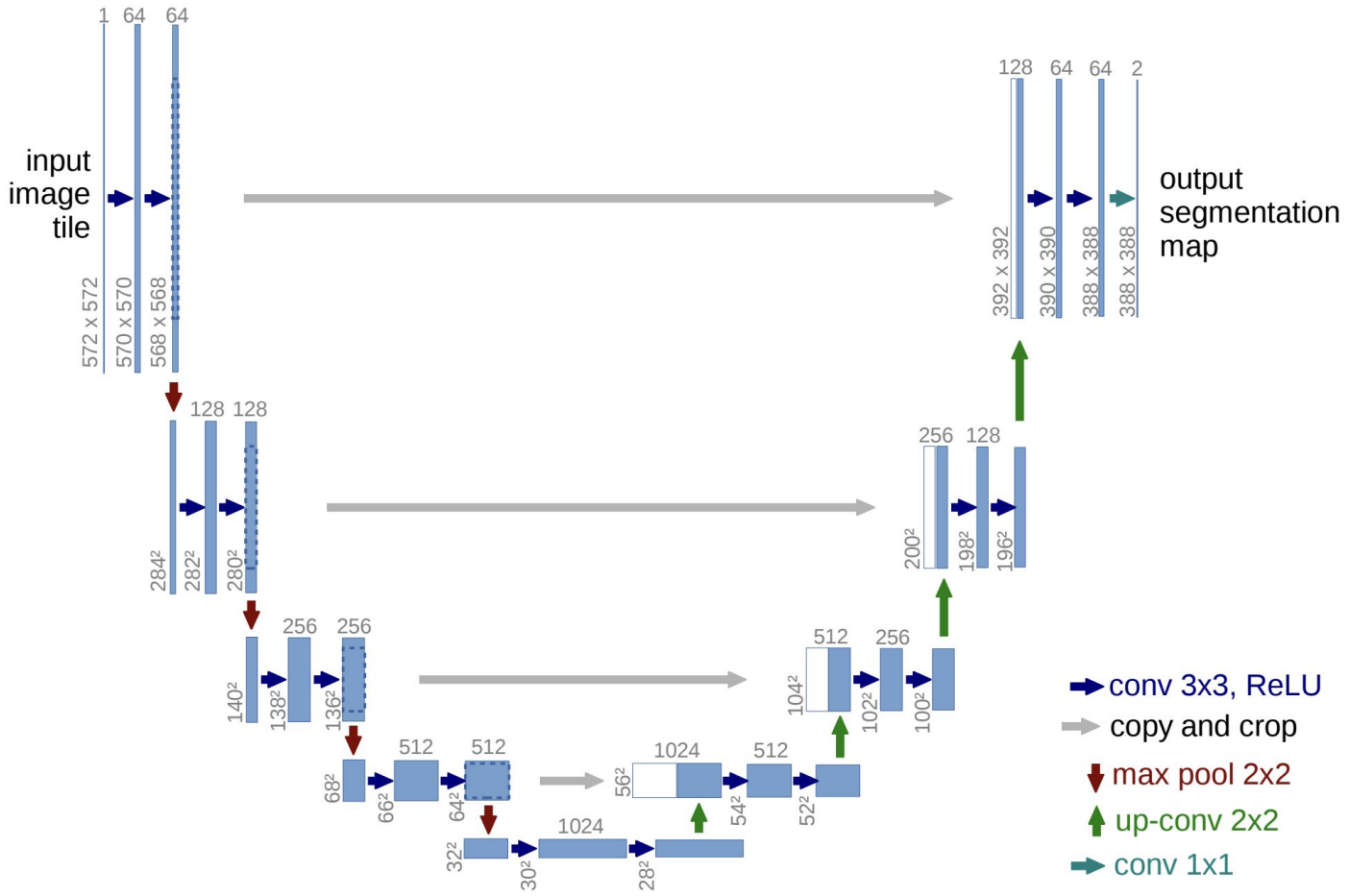
U-Net

Larissa Heinrich, HHMI Janelia

AI@MBL 2025

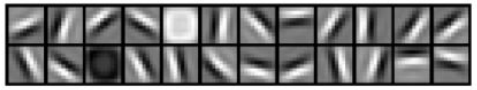
[slides adapted from Shalin Mehta and Anna Kreshuk]

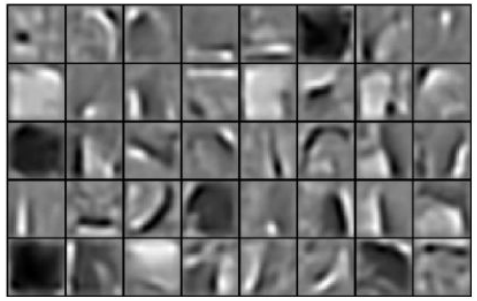
does this look familiar?

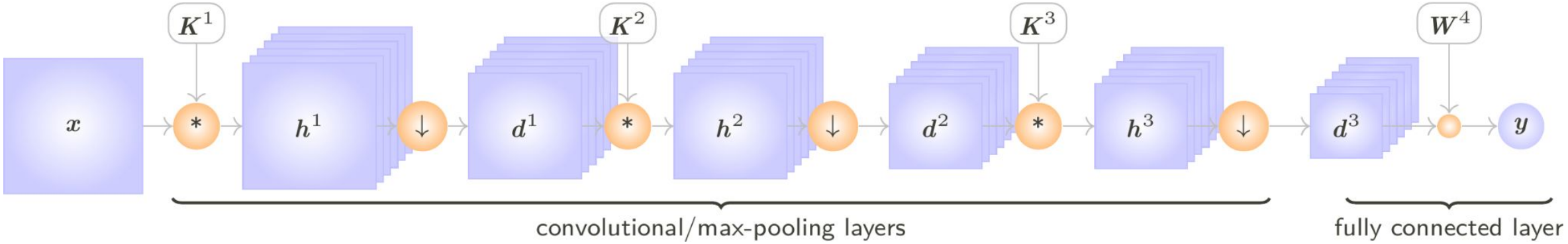


[Ronneberger et al., 2015]

remember yesterday?

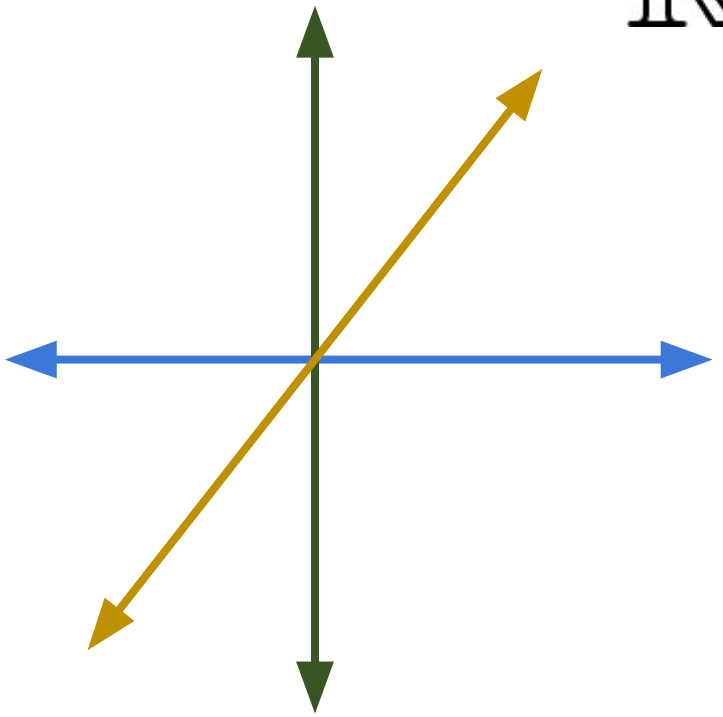






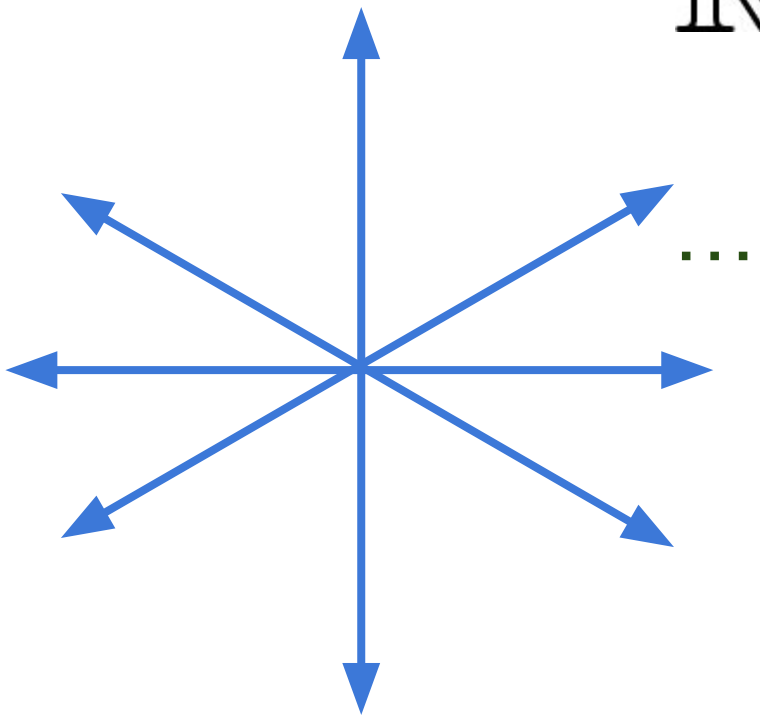
images

$\mathbb{R}^3$



images

$$\mathbb{R}^{H \times W}$$



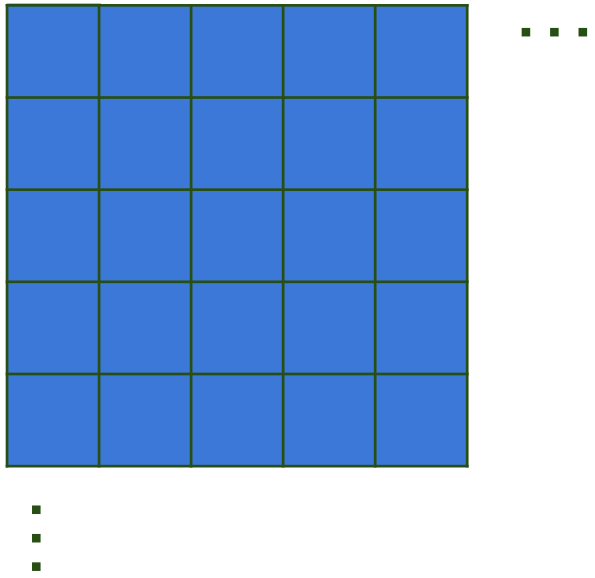


image classification

$$\mathbb{R}^{H \times W}$$

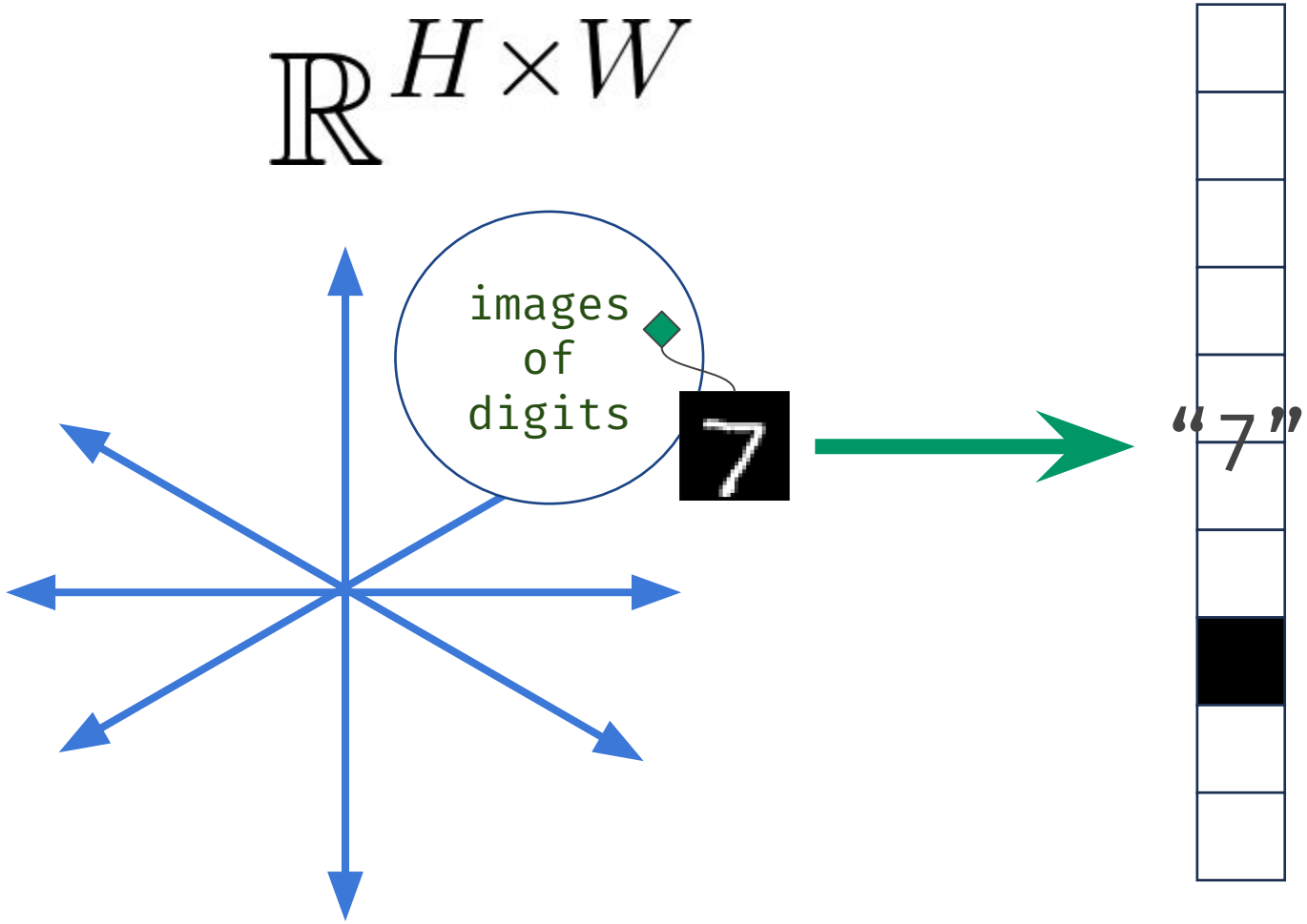
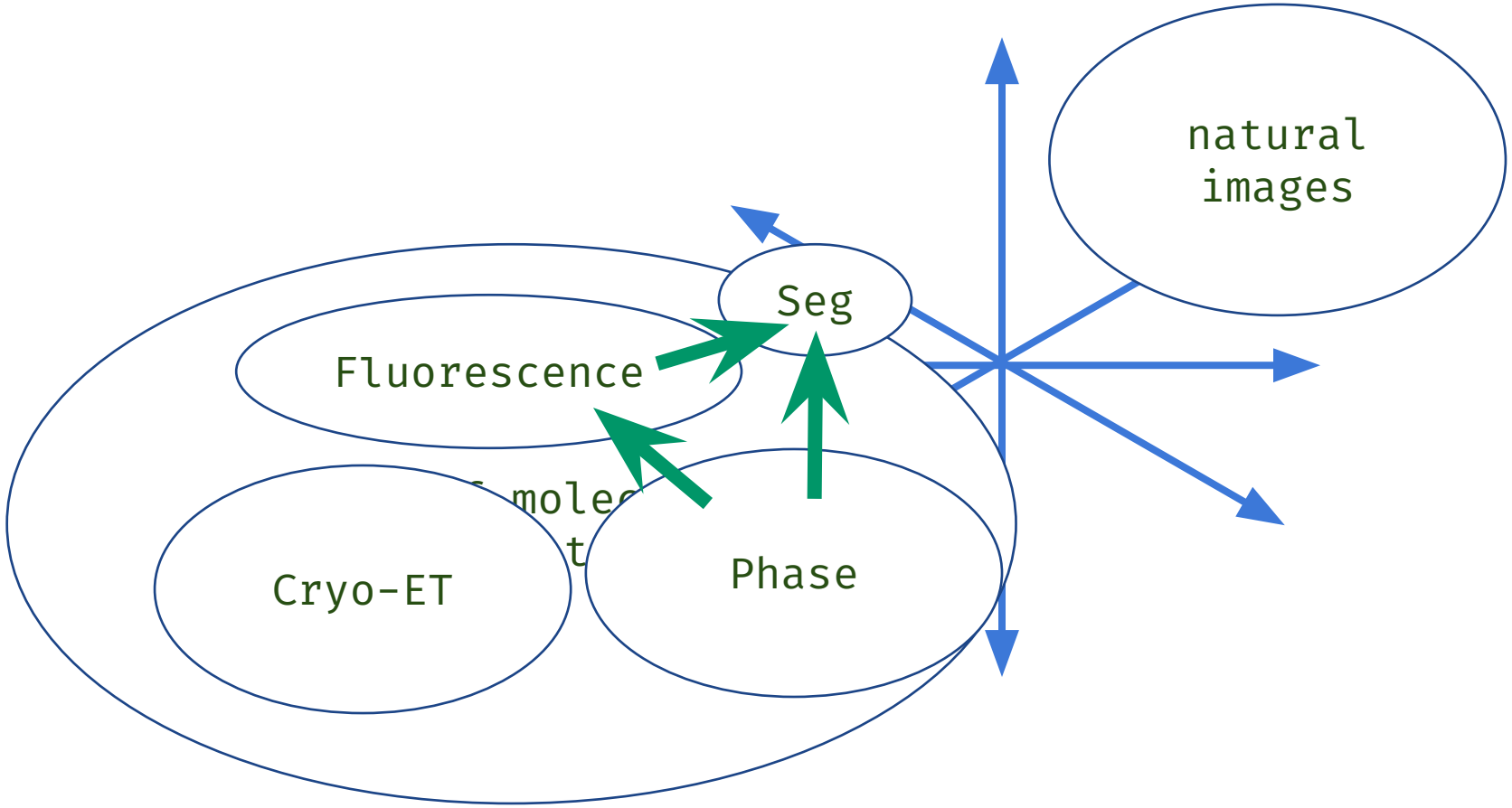
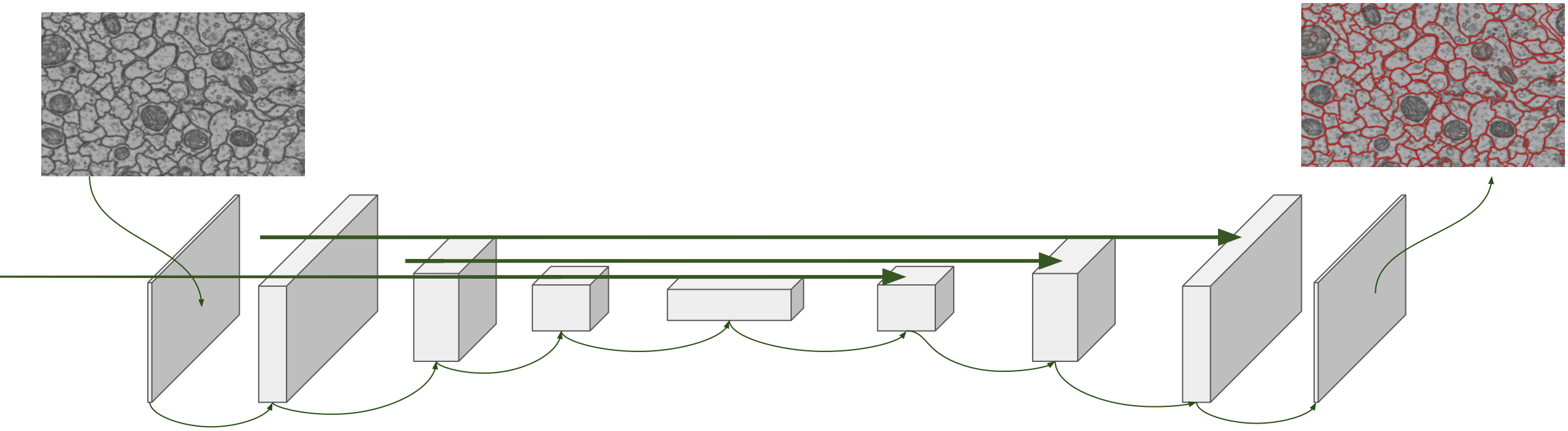


image-to-image transforms

$$\mathbb{R}^{H \times W}$$

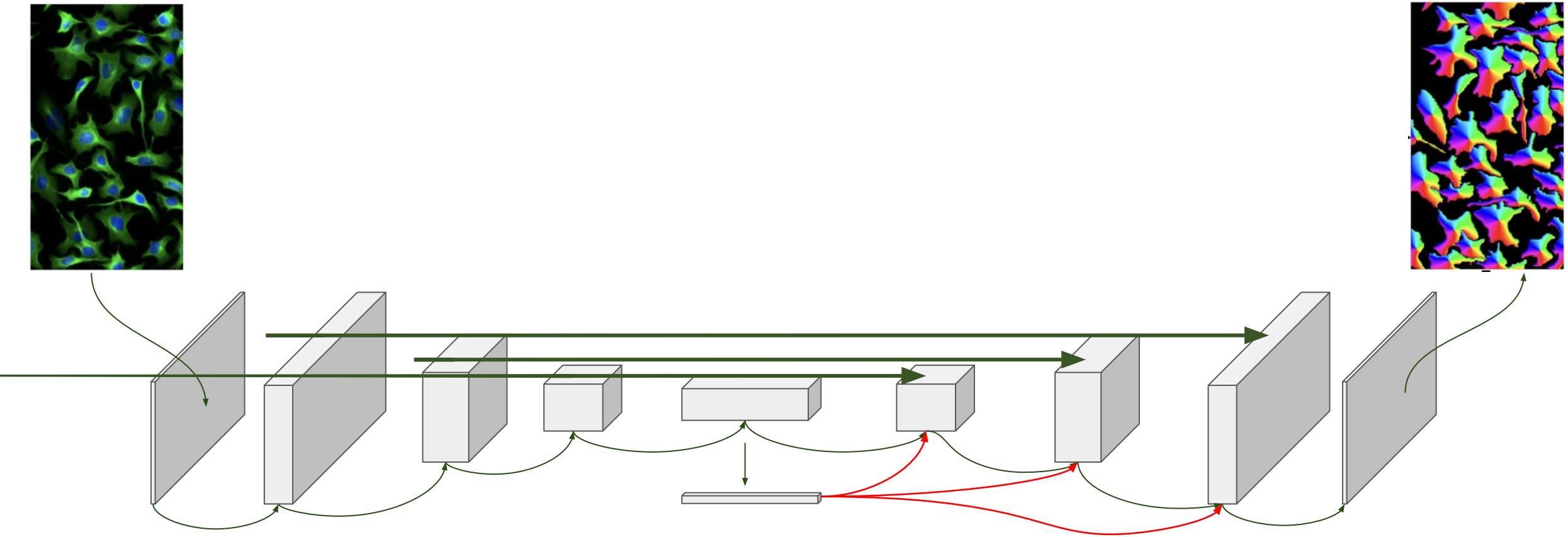


U-Net in bioimage analysis



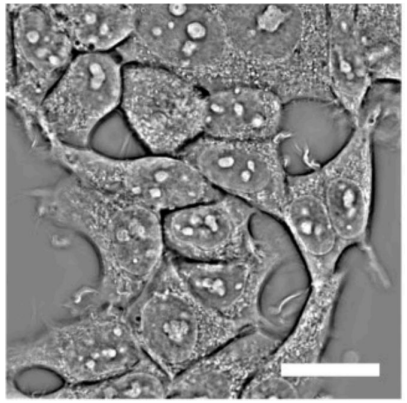
[Ronneberger et al., 2015]

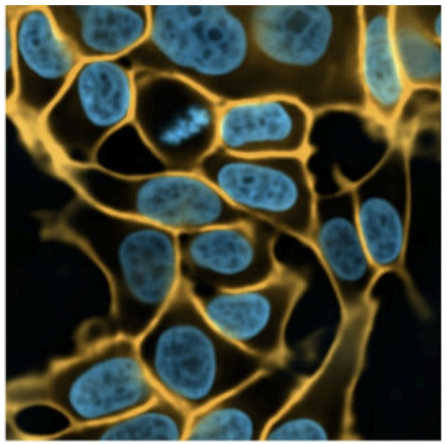
U-Net in bioimage analysis

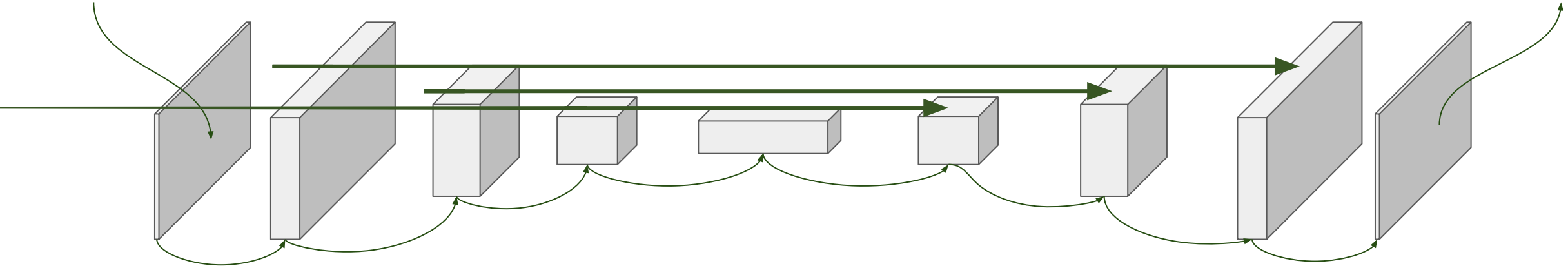


[Stringer et al., 2021]

U-Net in bioimage analysis

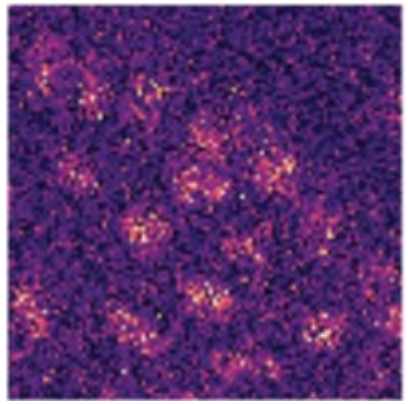


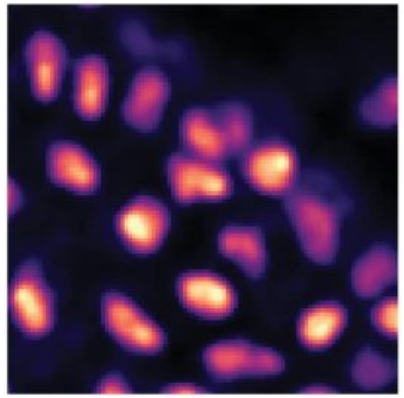


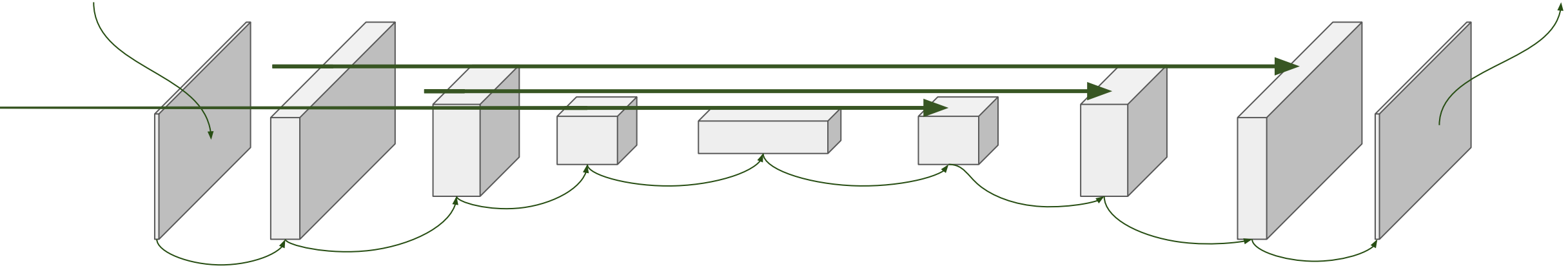


[Liu et al., 2024]

U-Net in bioimage analysis







[Weigert et al., 2018]

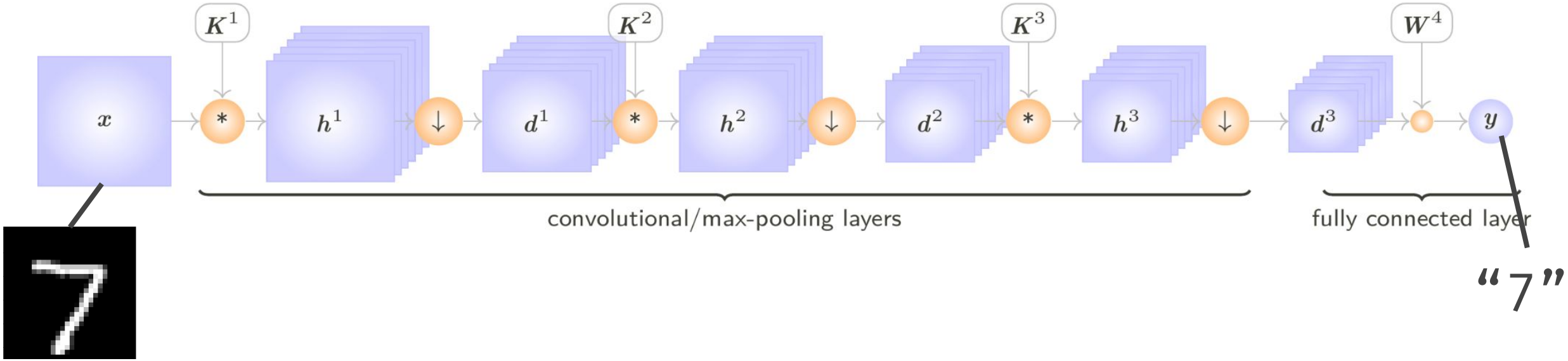
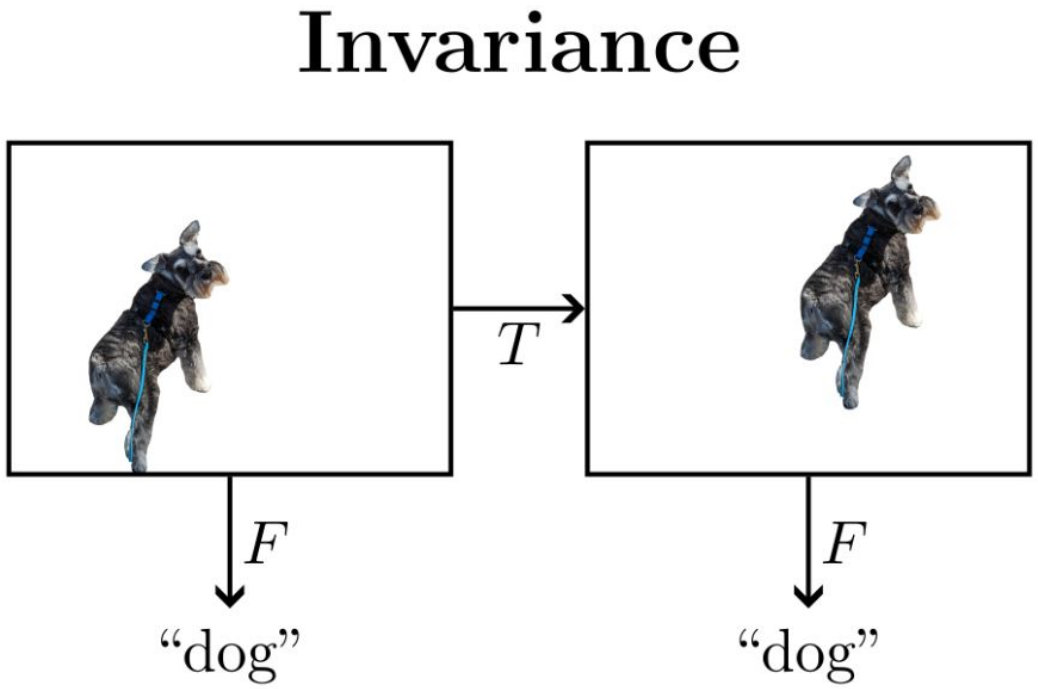
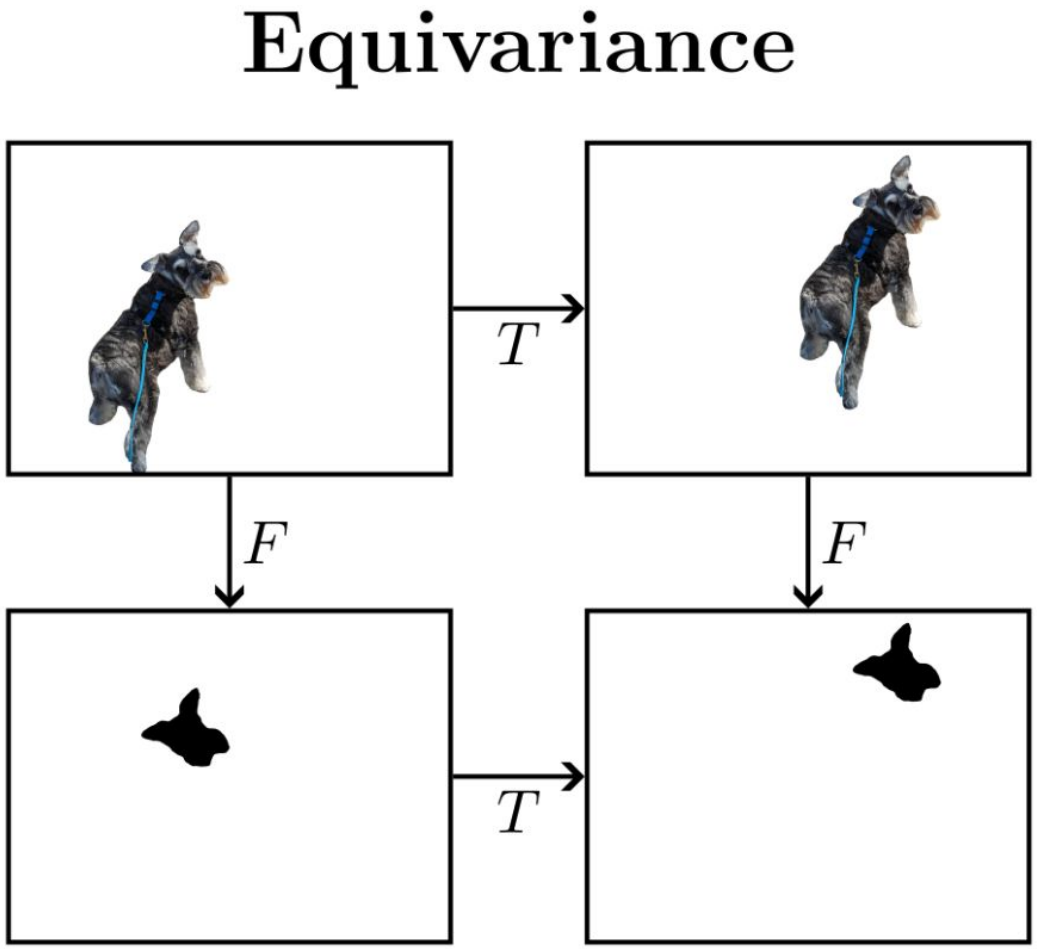


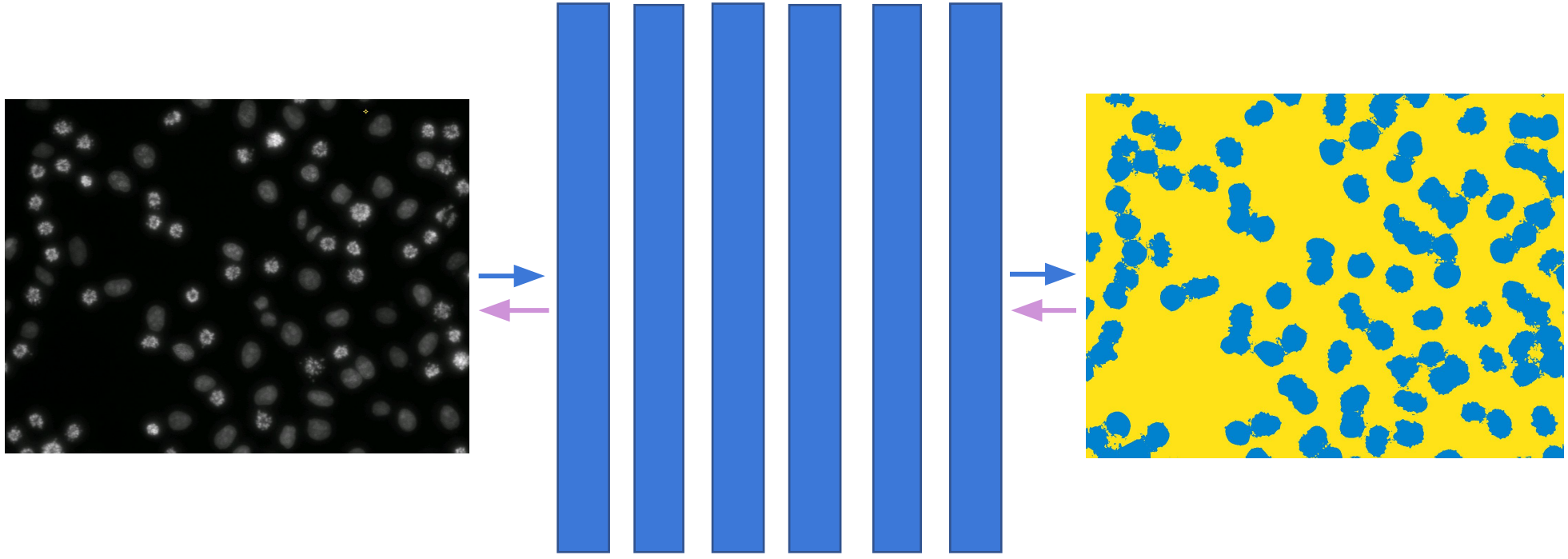
image-to-image: what’s missing?

shift equivariance

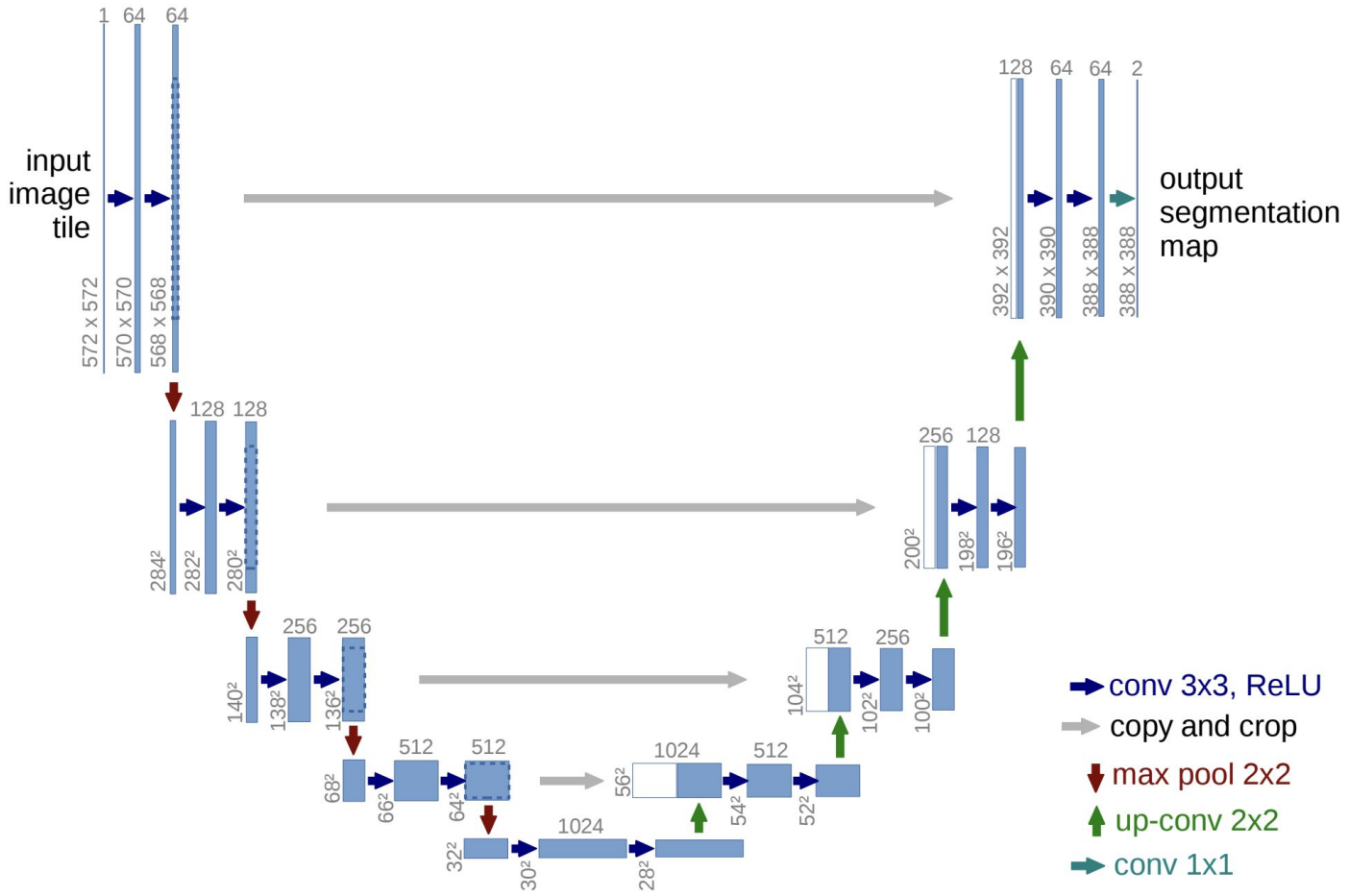




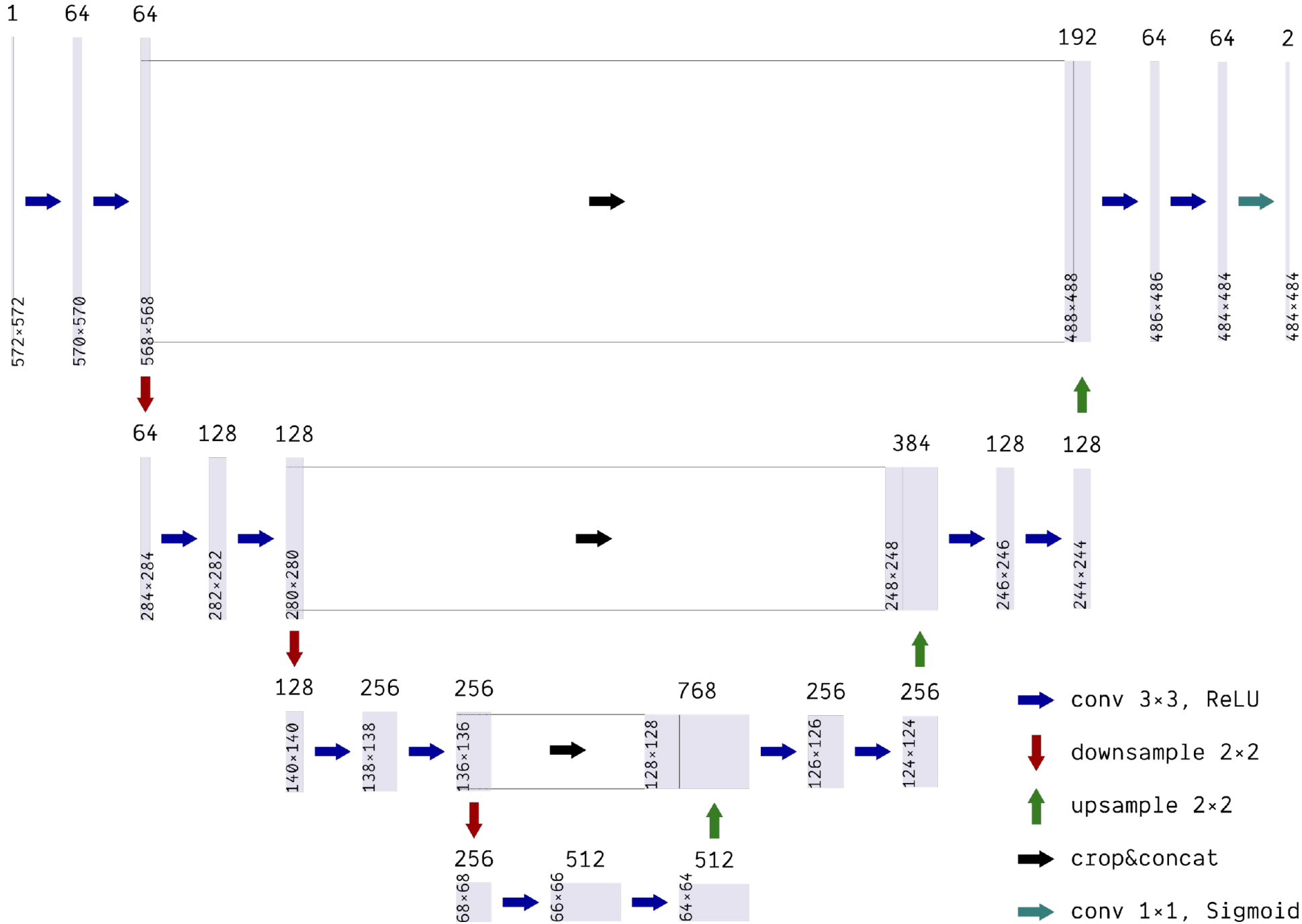
stacking convolutions



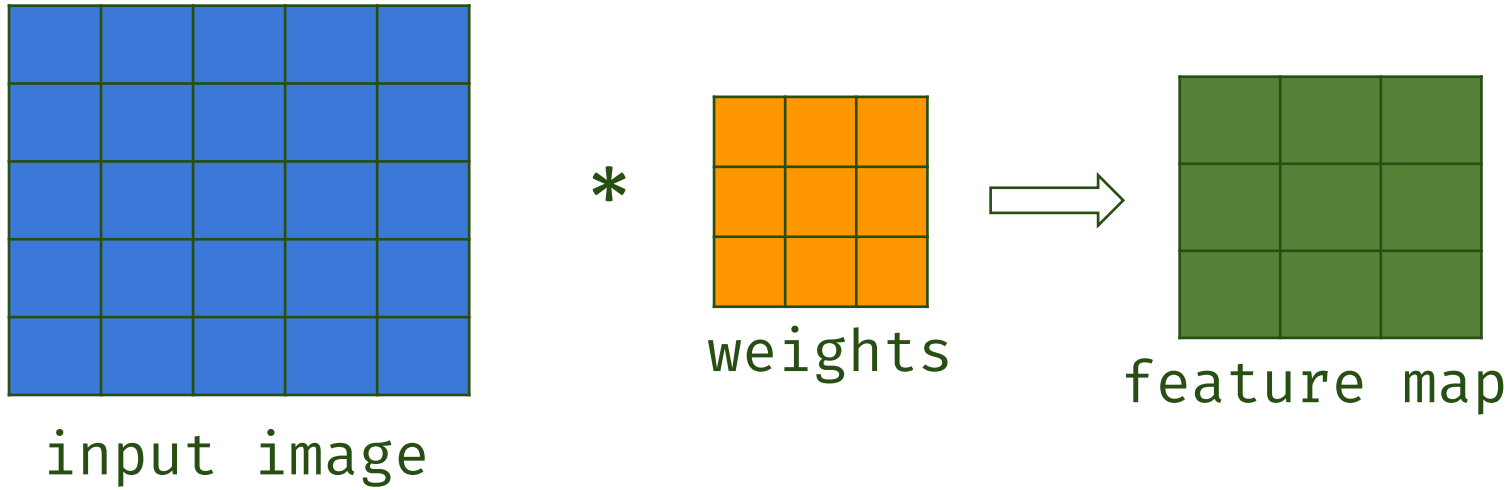
the U-Net



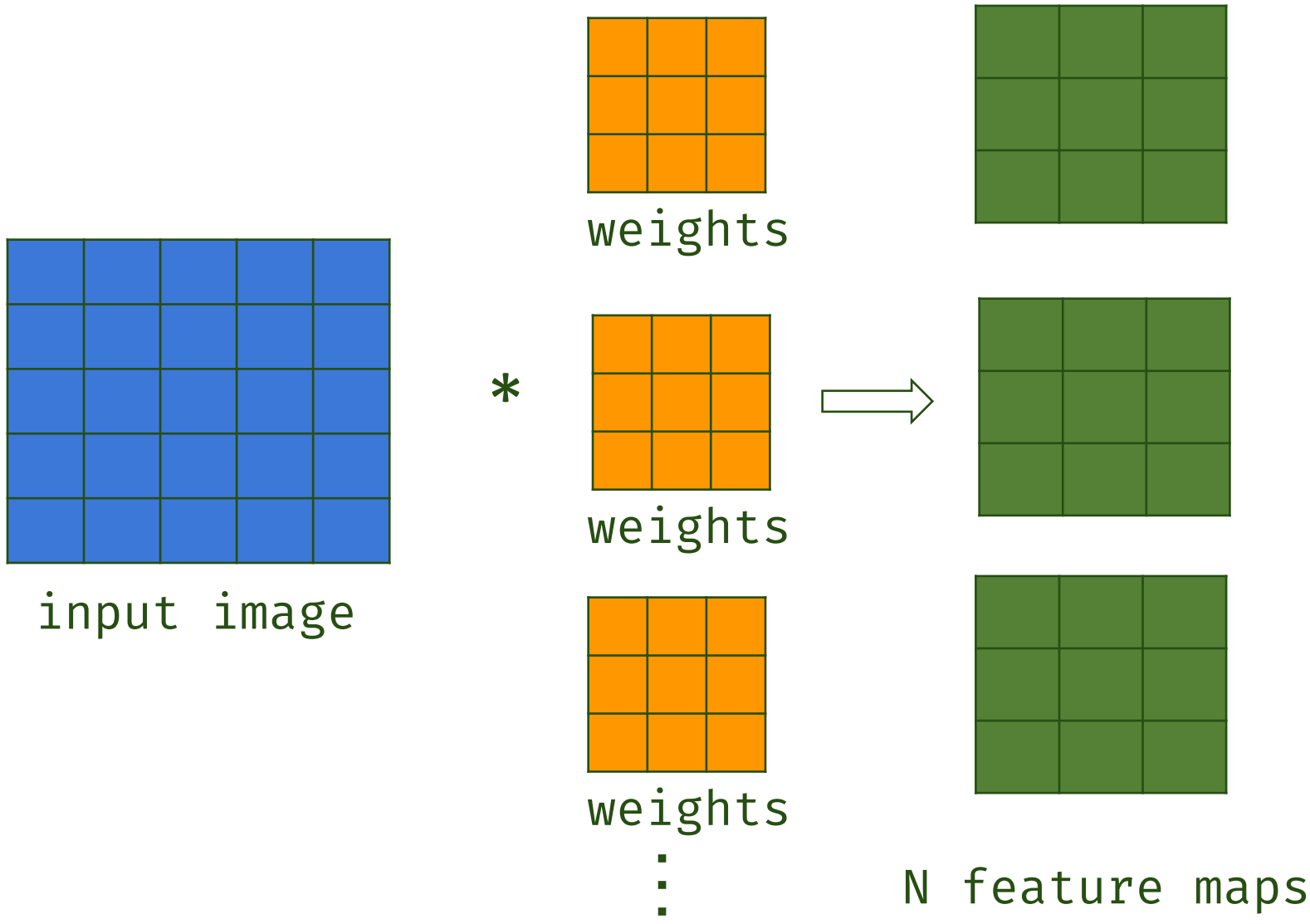
[Ronneberger et al., 2015]


1
64
64
192
64
64
2
572×572
570×570
568×568
488×488
486×486
484×484
484×484
64
128
128
384
128
128
284×284
282×282
280×280
248×248
246×246
244×244
128
256
256
768
256
256
140×140
138×138
136×136
128×128
126×126
124×124
256
512
512
68×68
66×66
64×64
conv 3×3, ReLU
downsample 2×2
upsample 2×2
crop&concat
conv 1×1, Sigmoid

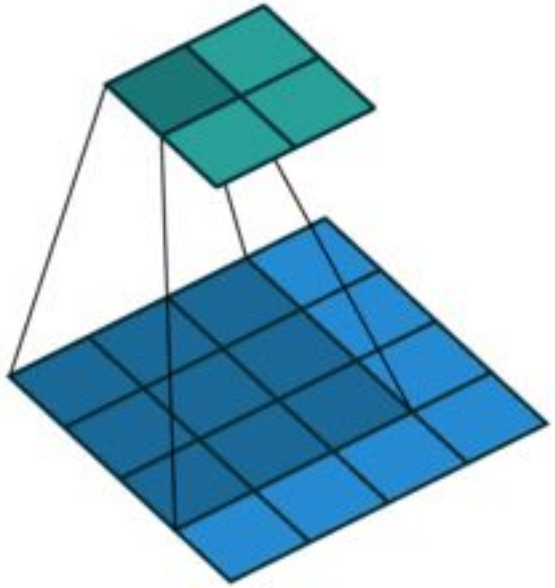
convolutional layer



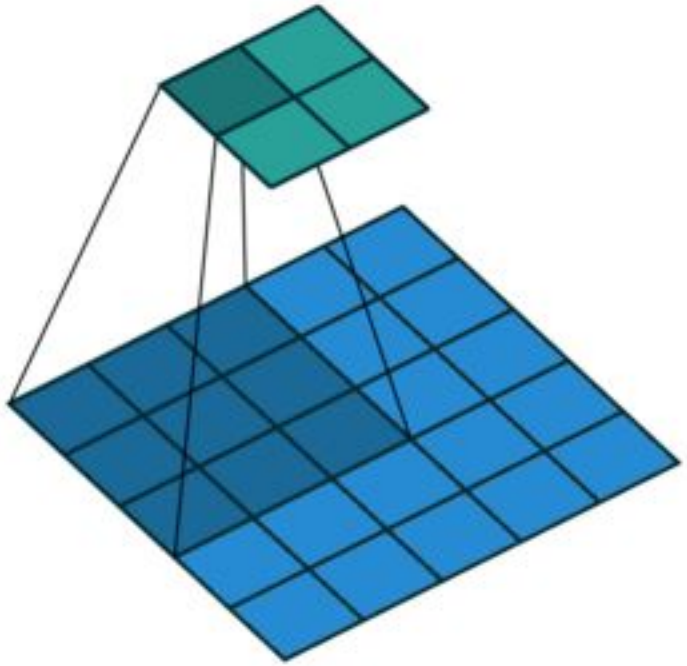
convolutional layer



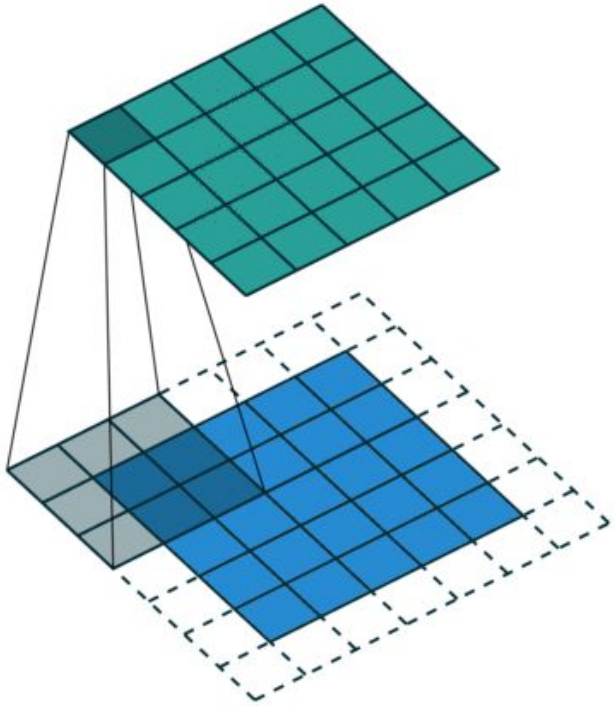
hyperparameters of convolution



padding = 0,
stride = 1



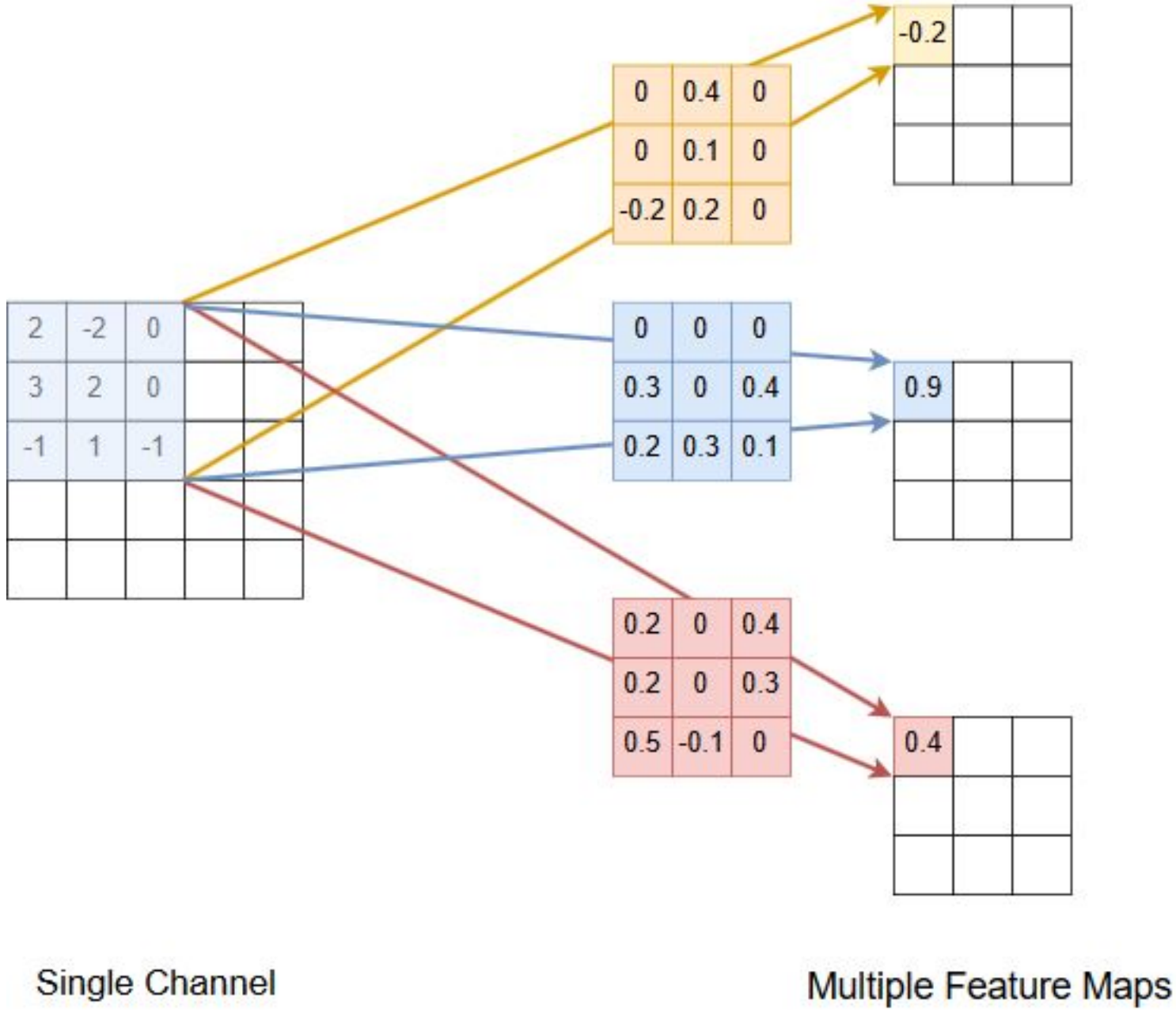
padding = 0,
stride = 2



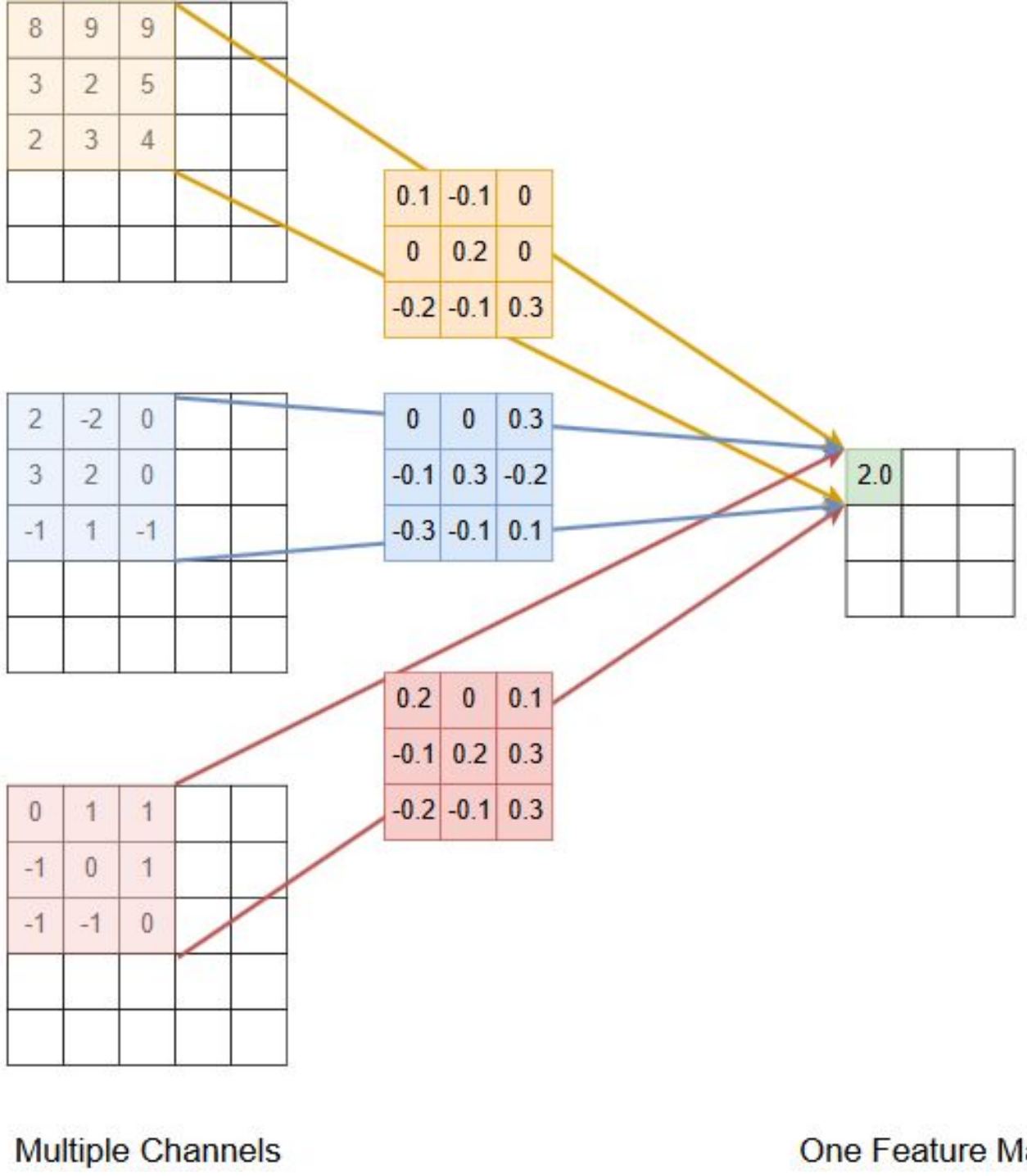
padding = 1,
stride = 1

https://hannibunny.github.io/mlbook/neuralnetworks/convolutionDemos.html

multi-channel output

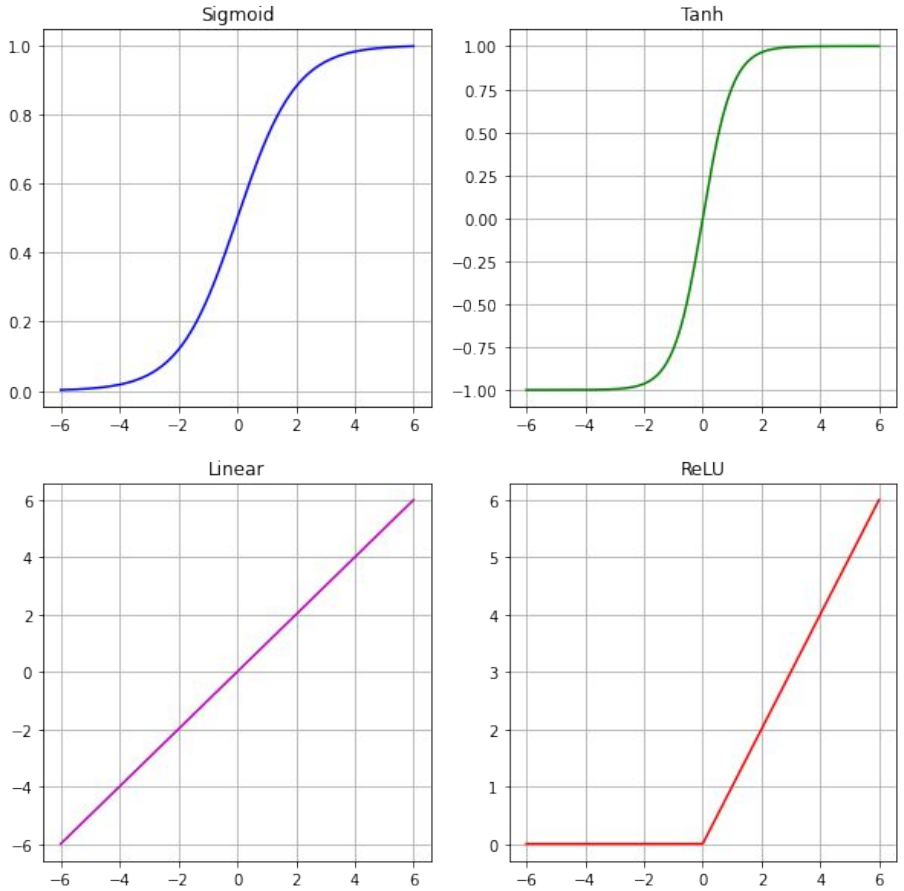


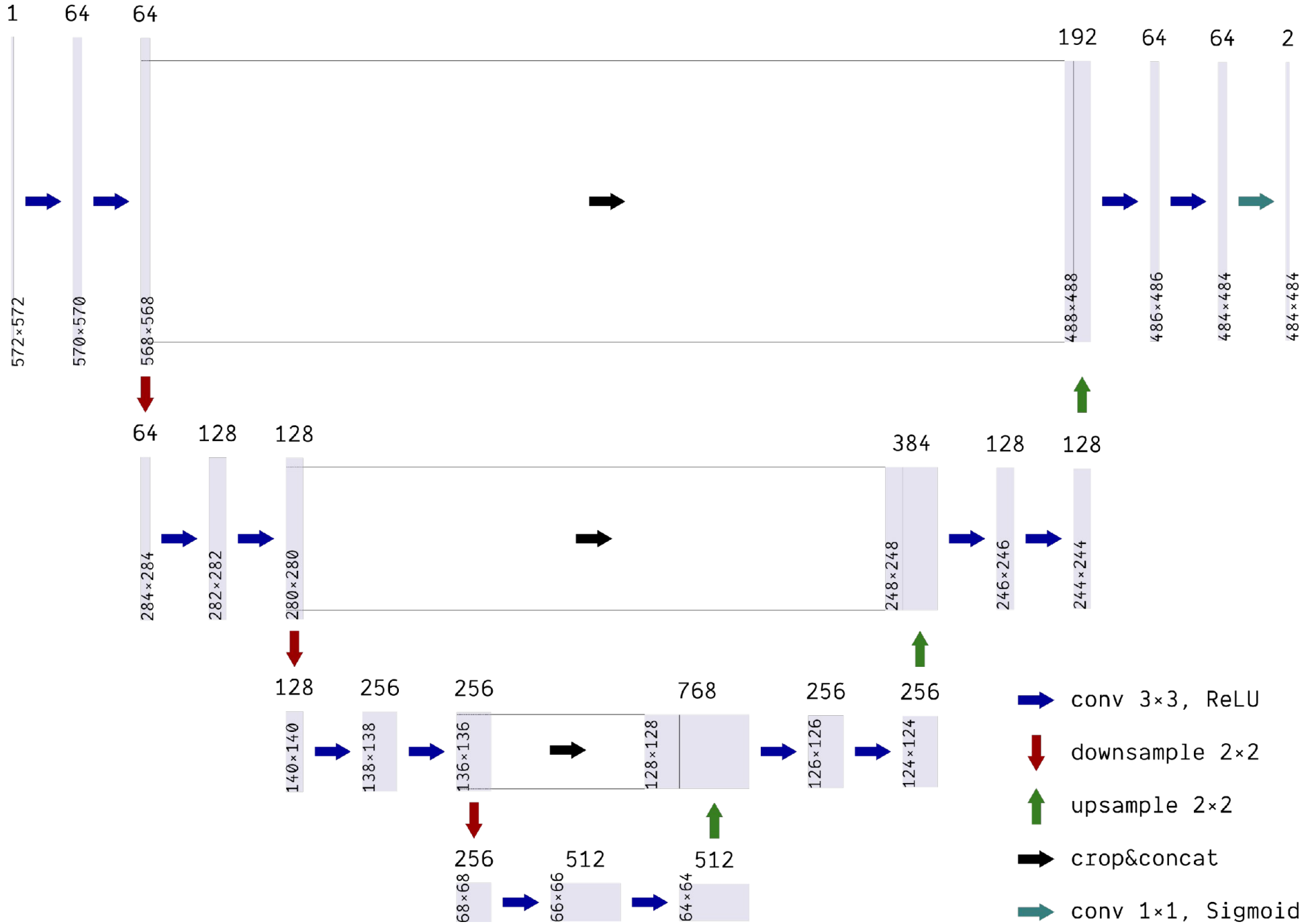
multi-channel input



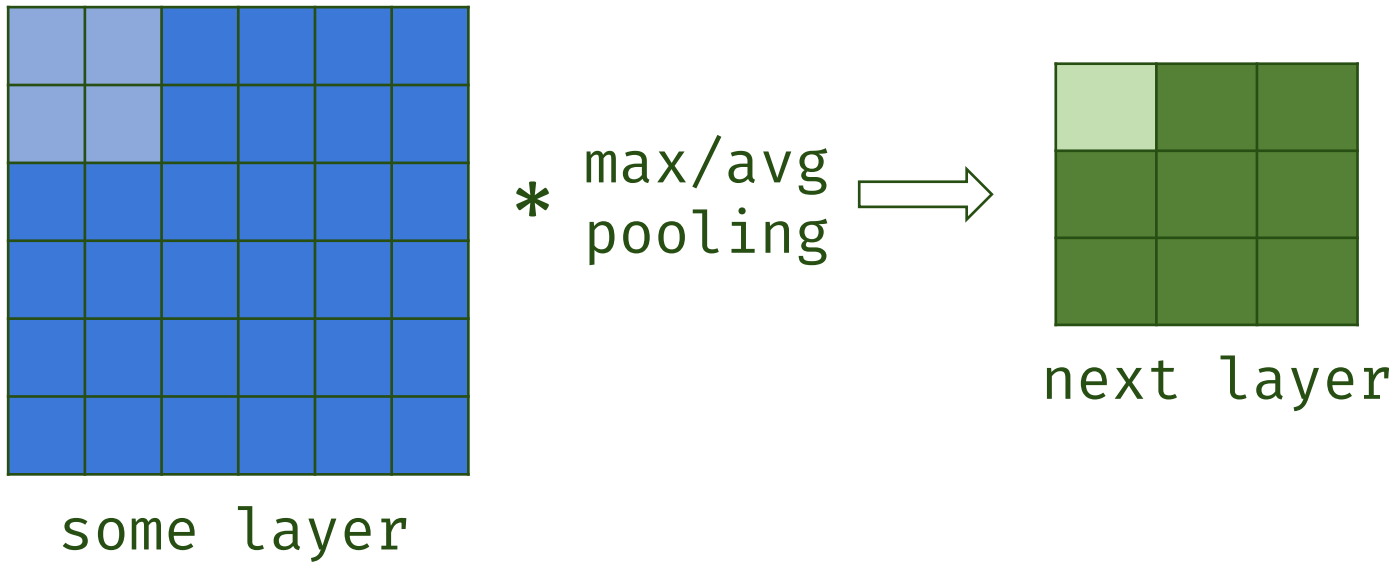
activation functions

update feature-maps element-wise

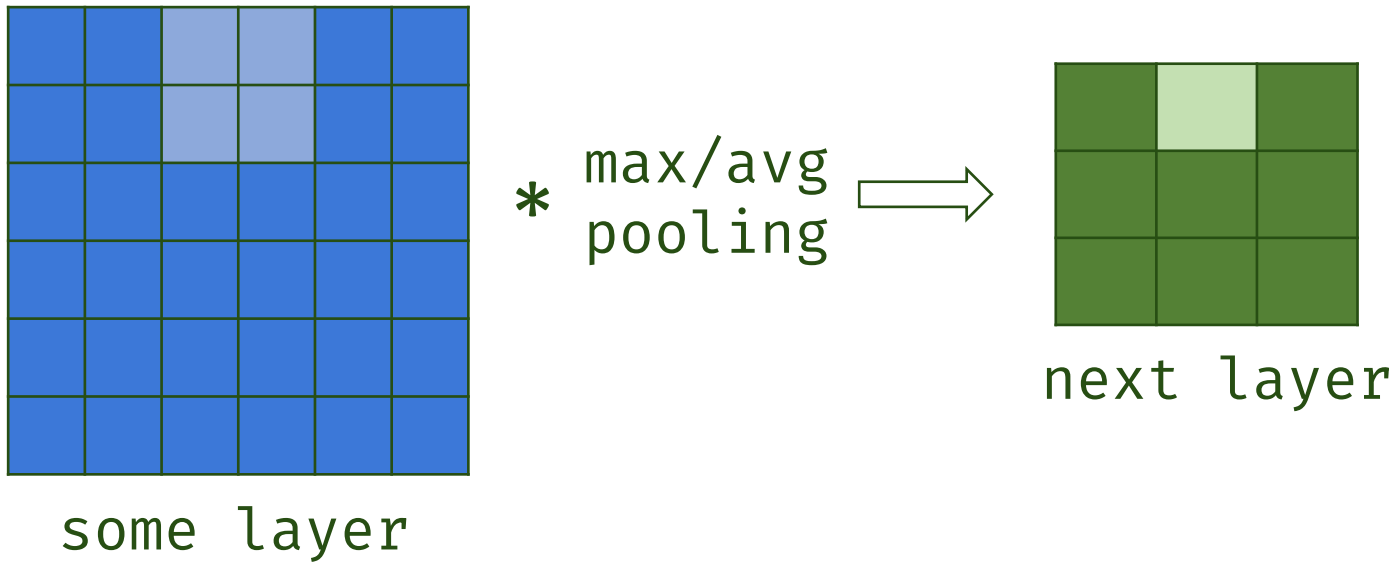



1
64
64
192
64
64
2
572×572
570×570
568×568
488×488
486×486
484×484
484×484
64
128
128
384
128
128
284×284
282×282
280×280
248×248
246×246
244×244
128
256
256
768
256
256
140×140
138×138
136×136
128×128
126×126
124×124
256
512
512
68×68
66×66
64×64
conv 3×3, ReLU
downsample 2×2
upsample 2×2
crop&concat
conv 1×1, Sigmoid

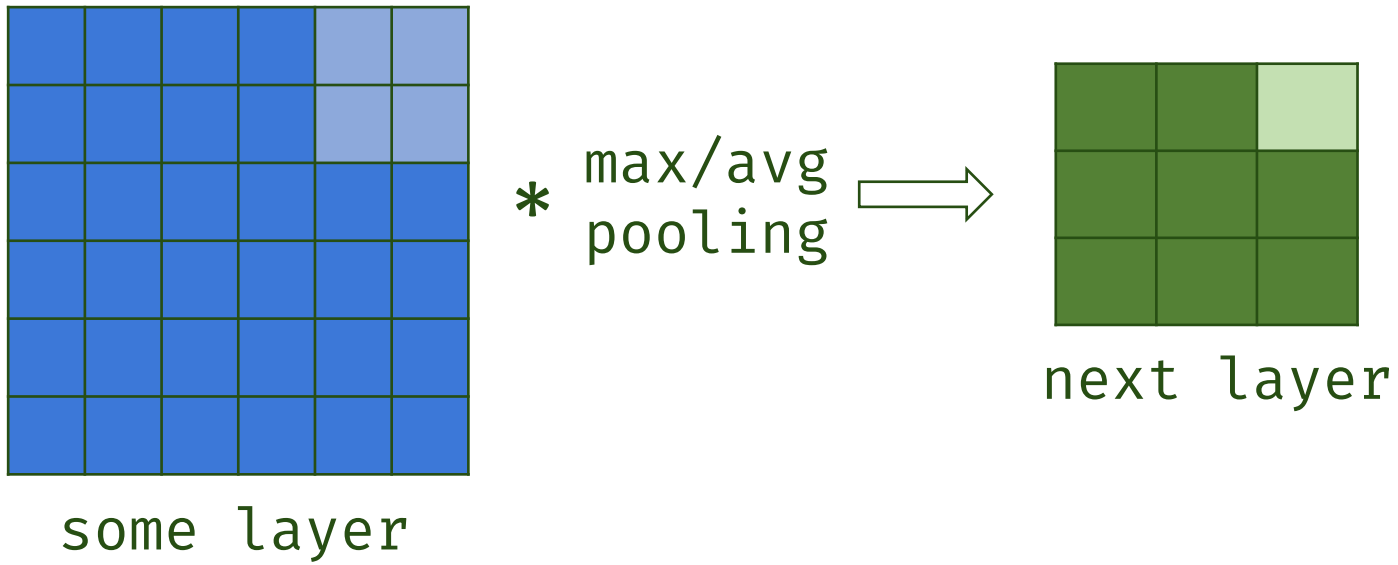
downsampling



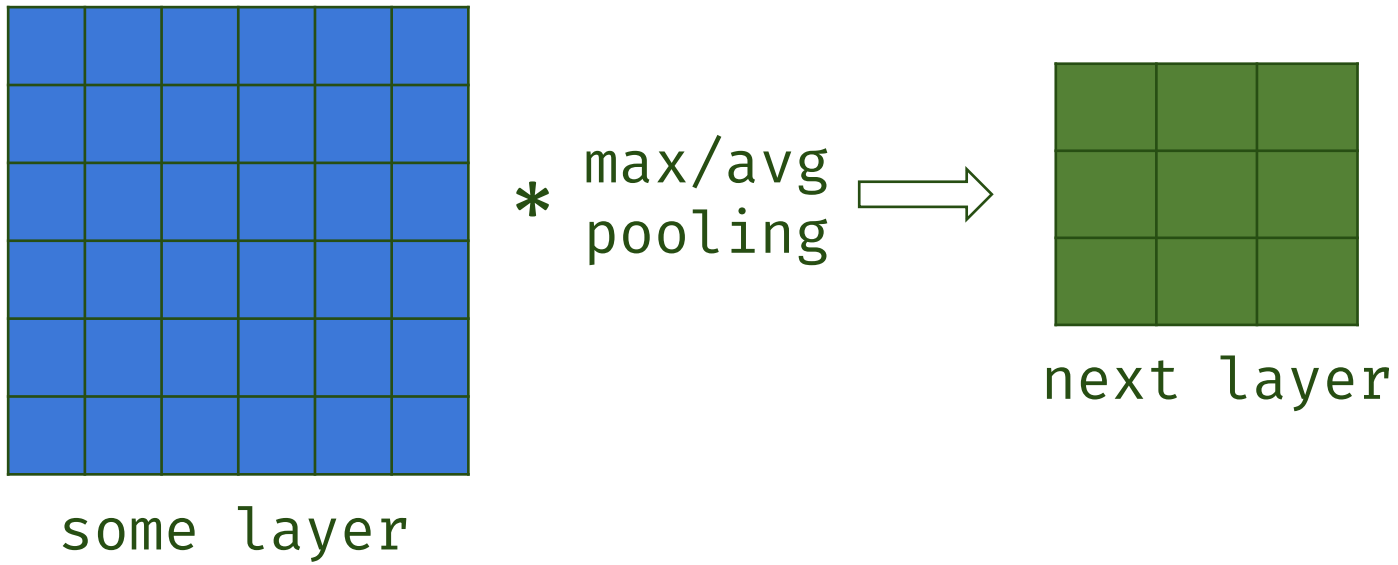
downsampling

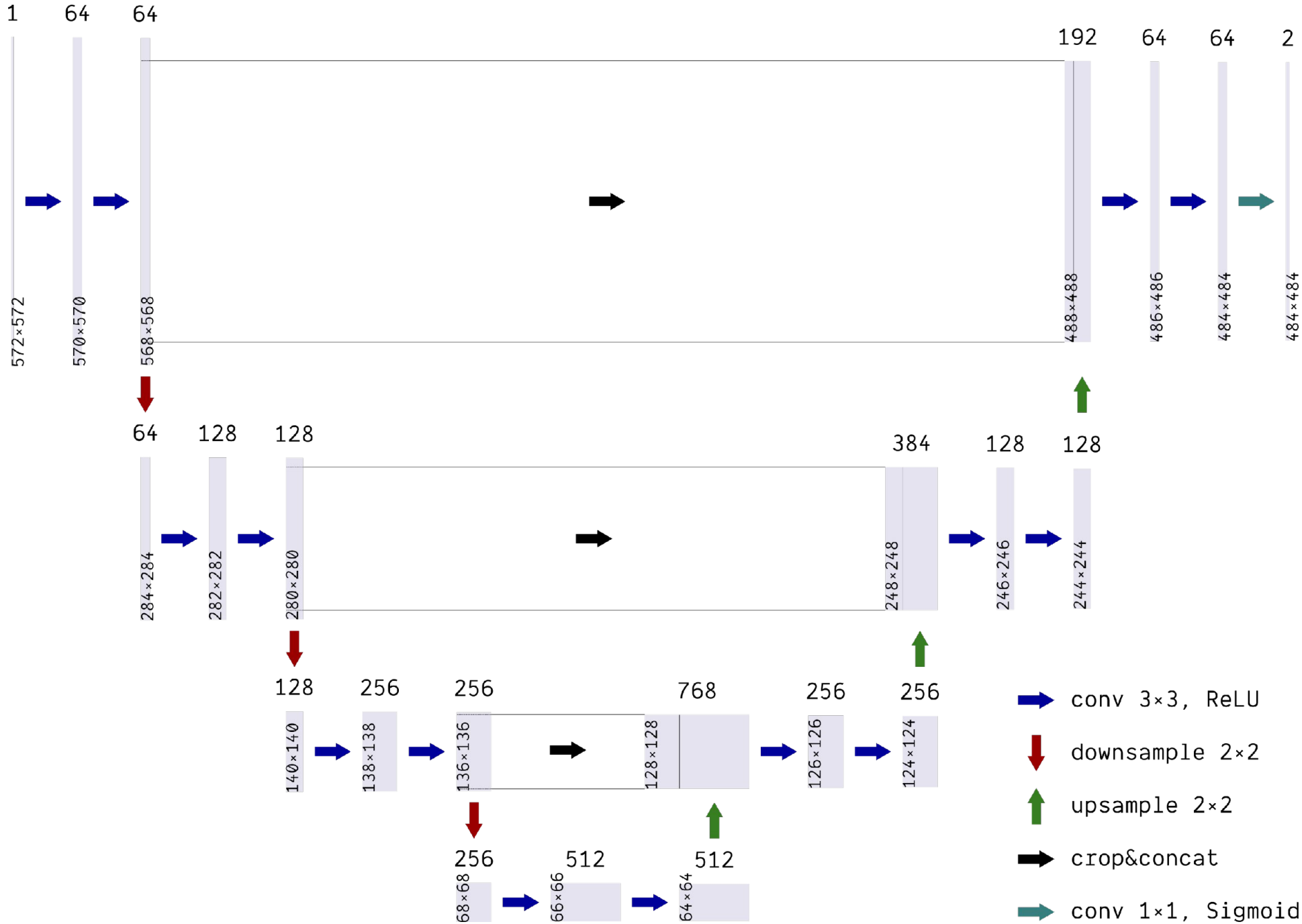


downsampling



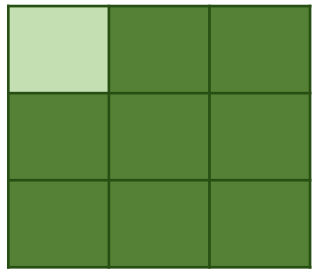
downsampling




1
64
64
192
64
64
2
572×572
570×570
568×568
488×488
486×486
484×484
484×484
64
128
128
384
128
128
284×284
282×282
280×280
248×248
246×246
244×244
128
256
256
768
256
256
140×140
138×138
136×136
128×128
126×126
124×124
256
512
512
68×68
66×66
64×64
conv 3×3, ReLU
downsample 2×2
upsample 2×2
crop&concat
conv 1×1, Sigmoid

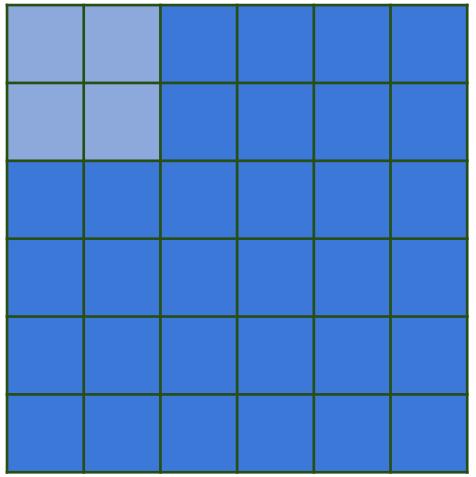
upsampling

upsampling



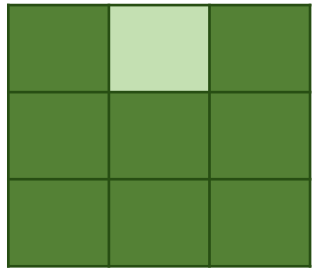
some layer

* nearest-neighbor/
bilinear



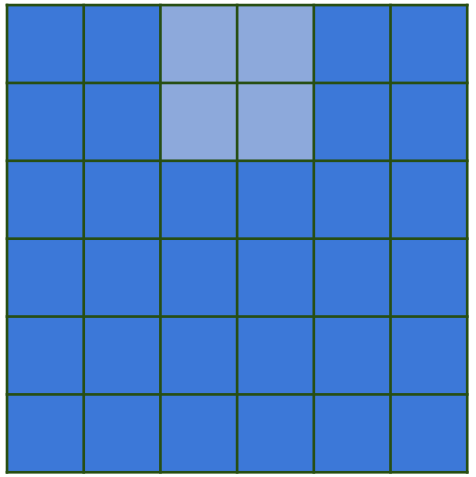
next layer

upsampling



some layer

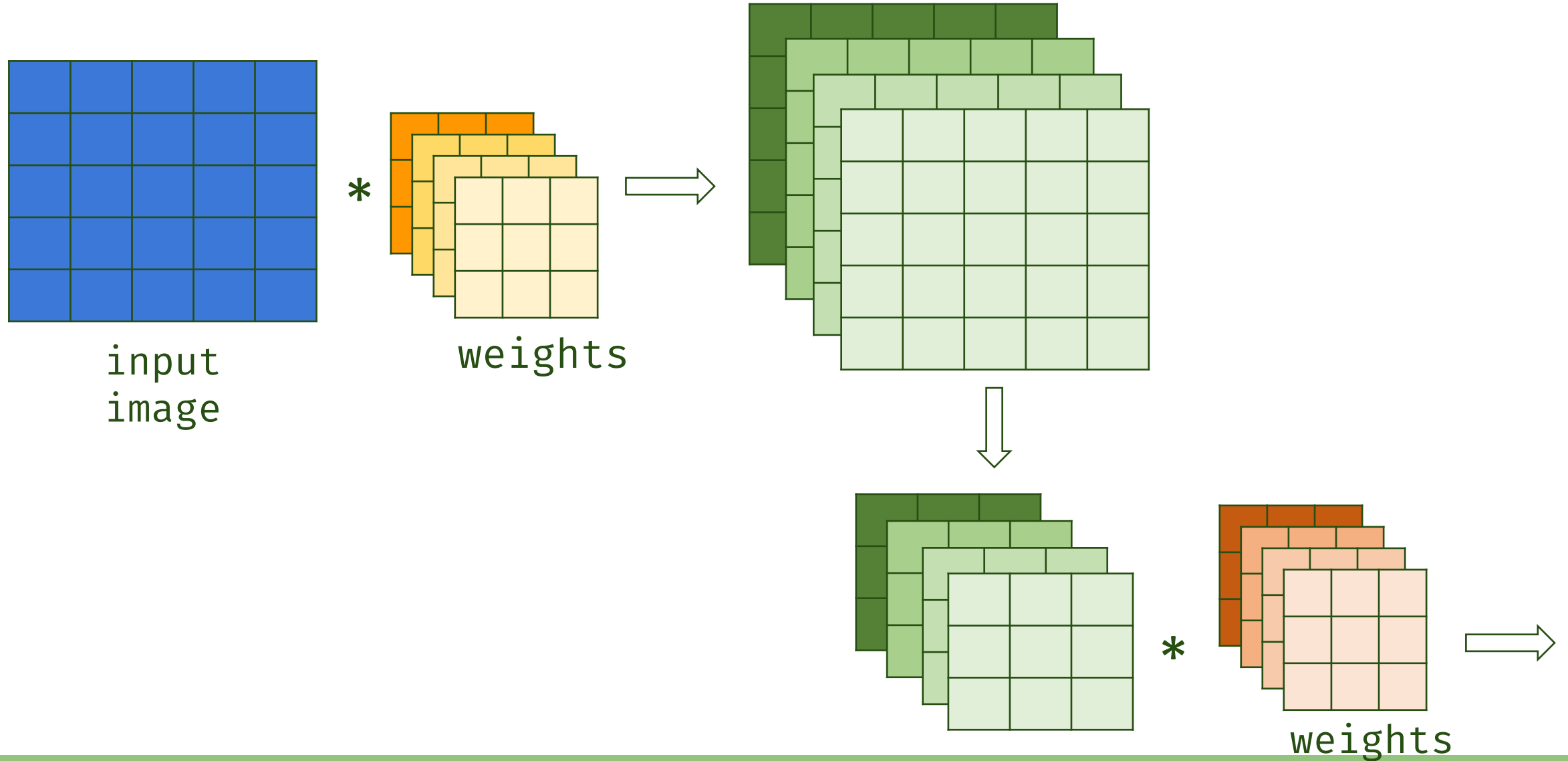
* nearest-neighbor/
bilinear



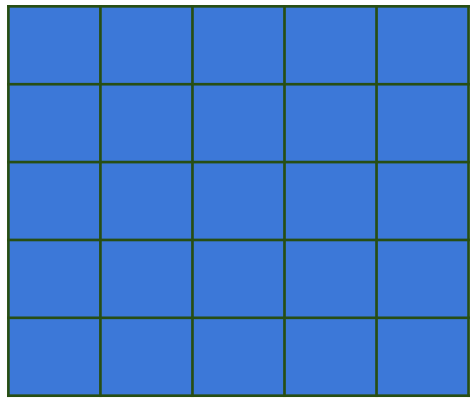
next layer

upsampling

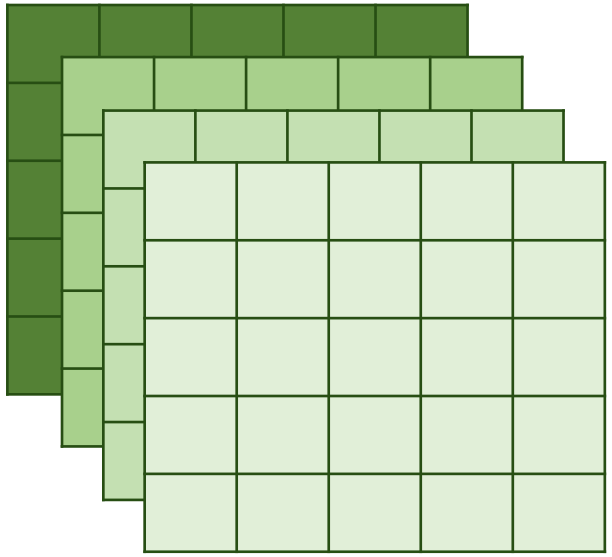
building an encoder



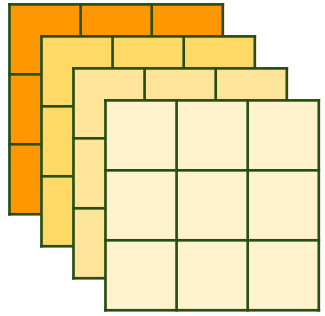
building a decoder



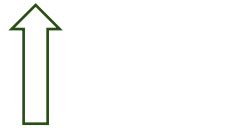
output
image

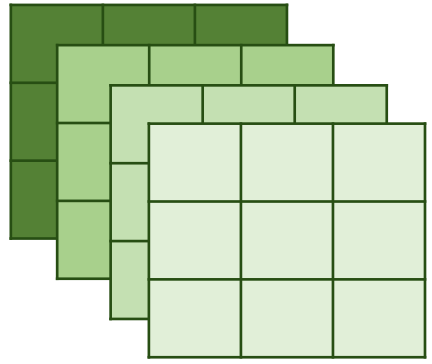


*

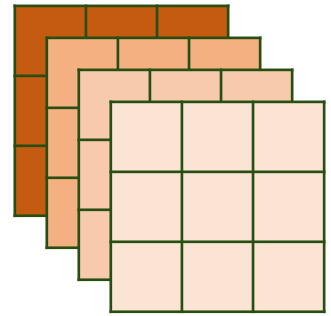


weights

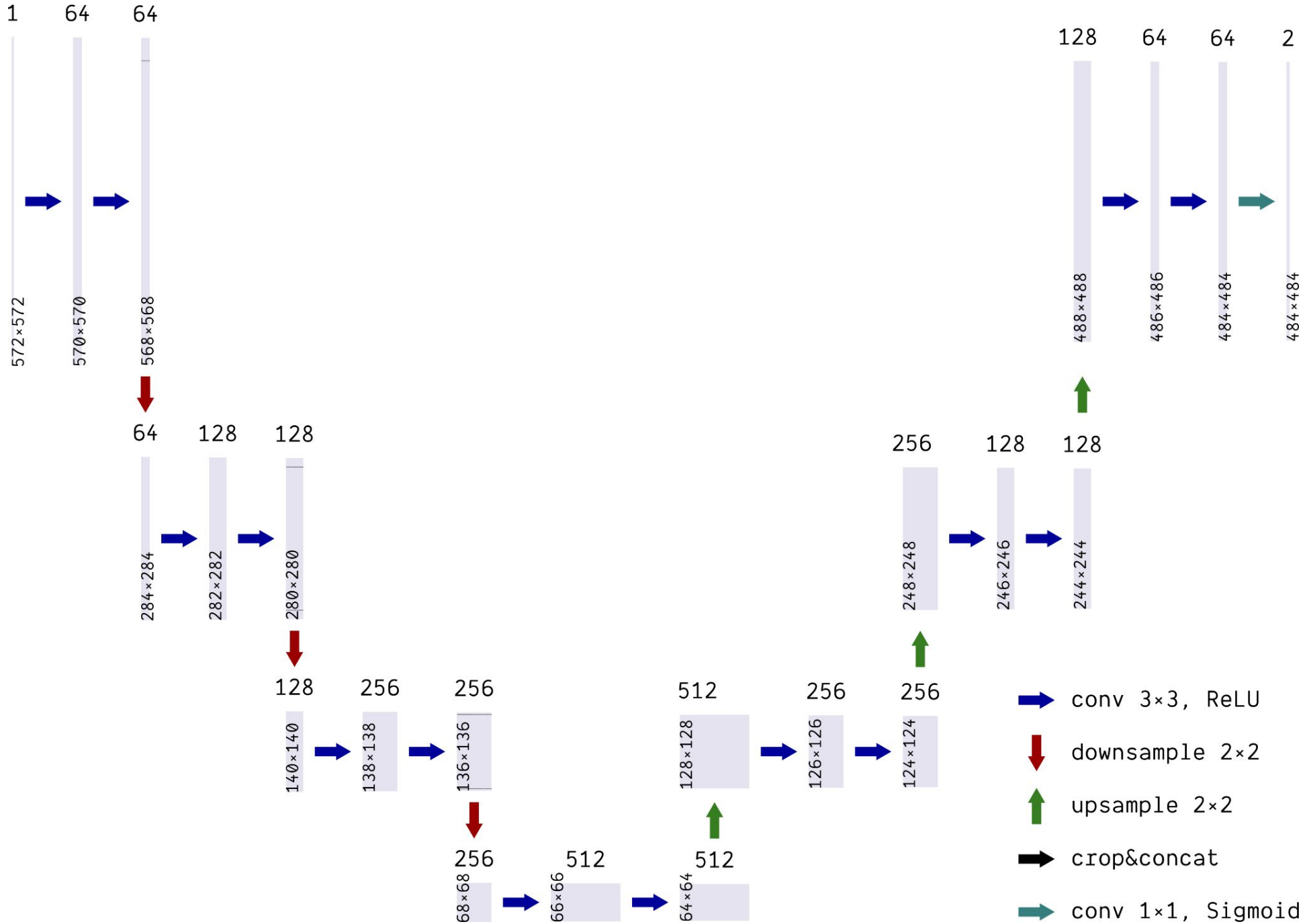


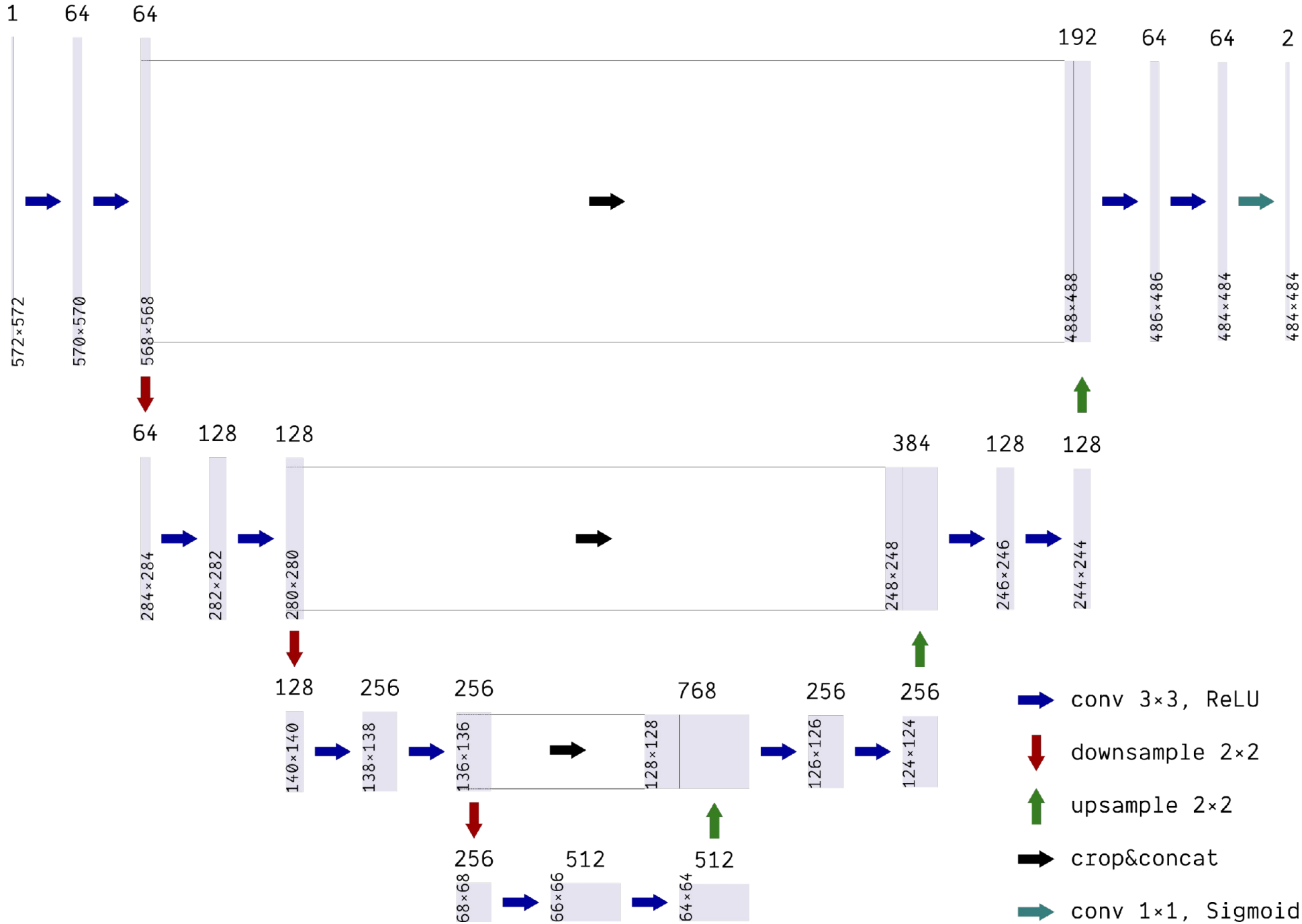


*

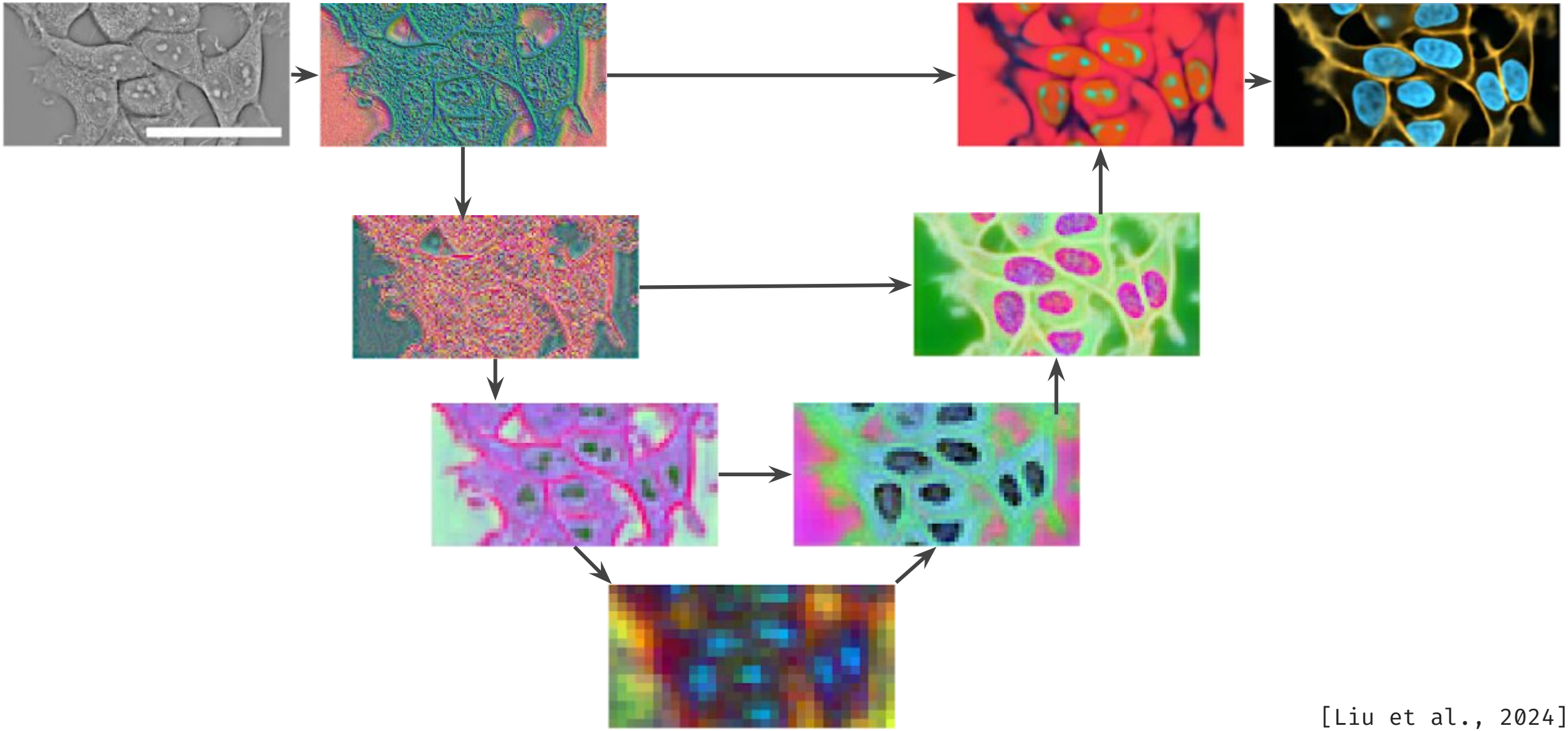


weights


1
64
64
572×572
570×570
568×568
64
128
128
284×284
282×282
280×280
128
256
256
140×140
138×138
136×136
256
512
512
68×68
66×66
64×64
512
256
256
128×128
126×126
124×124
256
128
128
248×248
246×246
244×244
128
64
64
2
488×488
486×486
484×484
484×484
conv 3×3, ReLU
downsample 2×2
upsample 2×2
crop&concat
conv 1×1, Sigmoid


1
64
64
192
64
64
2
572×572
570×570
568×568
488×488
486×486
484×484
484×484
64
128
128
384
128
128
284×284
282×282
280×280
248×248
246×246
244×244
128
256
256
768
256
256
140×140
138×138
136×136
128×128
126×126
124×124
256
512
512
68×68
66×66
64×64
conv 3×3, ReLU
downsample 2×2
upsample 2×2
crop&concat
conv 1×1, Sigmoid

multi-scale representation

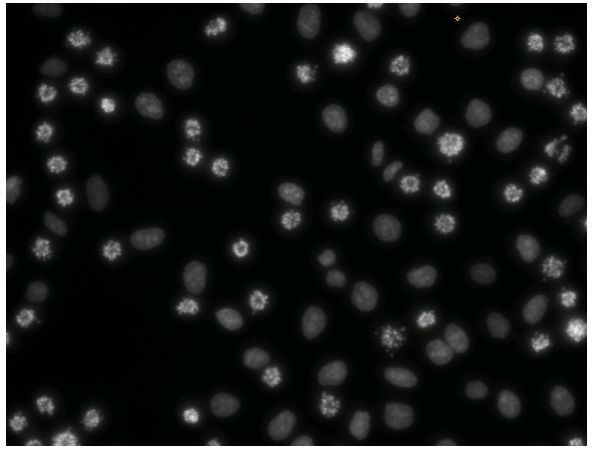


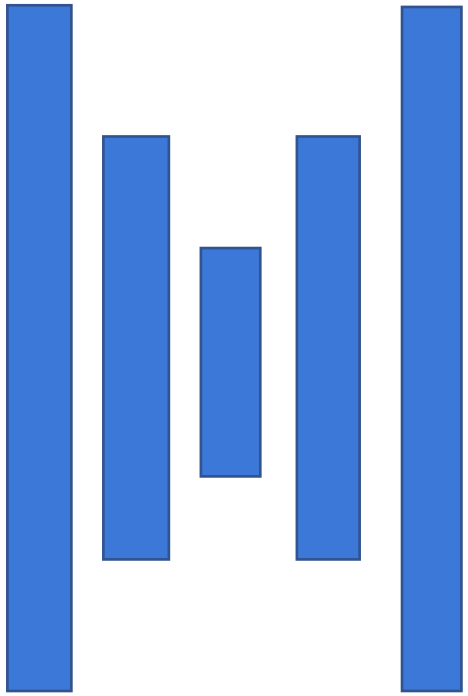
[Liu et al., 2024]

training

Forward pass

Backpropagation





Groundtruth

Some other image

Loss for training

Trying to become some other image

Current output: transformed image

Metric for validation

Go build your own U-Net!