Advanced Deep Learning 2022
FCNs + U-Nets

Stefan Sommer, Department of Computer Science, UCPH

May 4, 2022

Literature

U-Net: Convolutional Networks for Biomedical Image Segmentation https://arxiv.org/abs/1505.04597

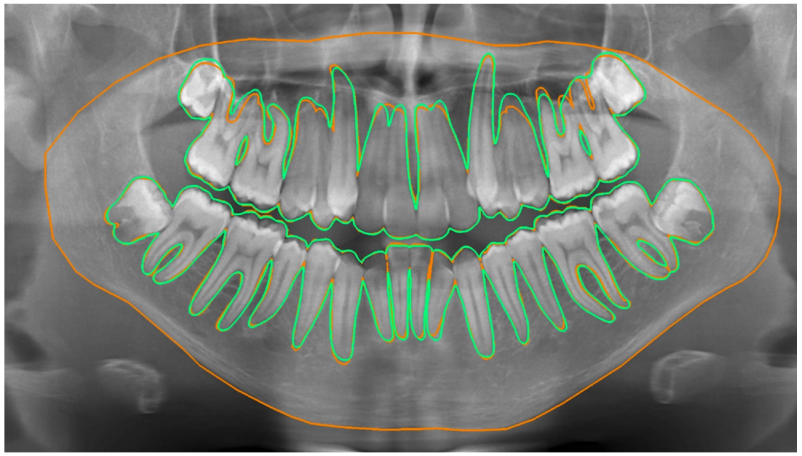
Fully Convolutional Networks for Semantic Segmentation
https://arxiv.org/abs/1411.4038

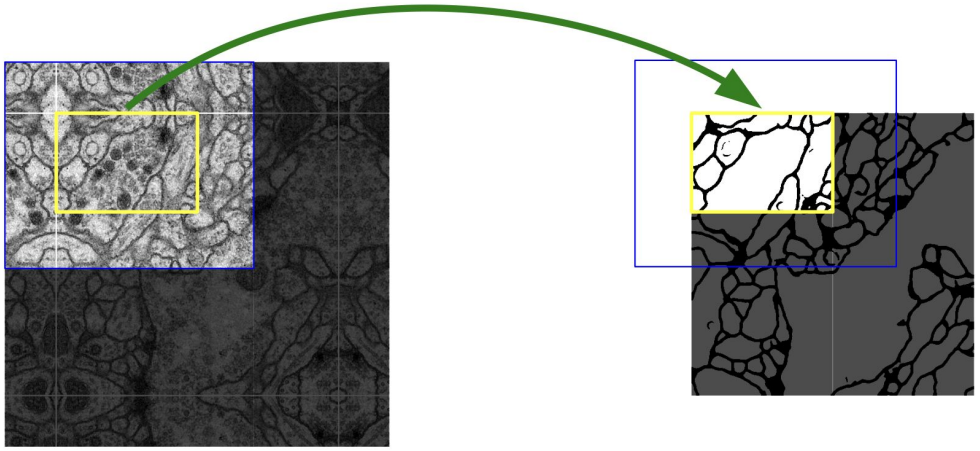
d2l 13.9-13.11:
https://d2l.ai/chapter_computer-vision/semantic-segmentation-and-dataset.html

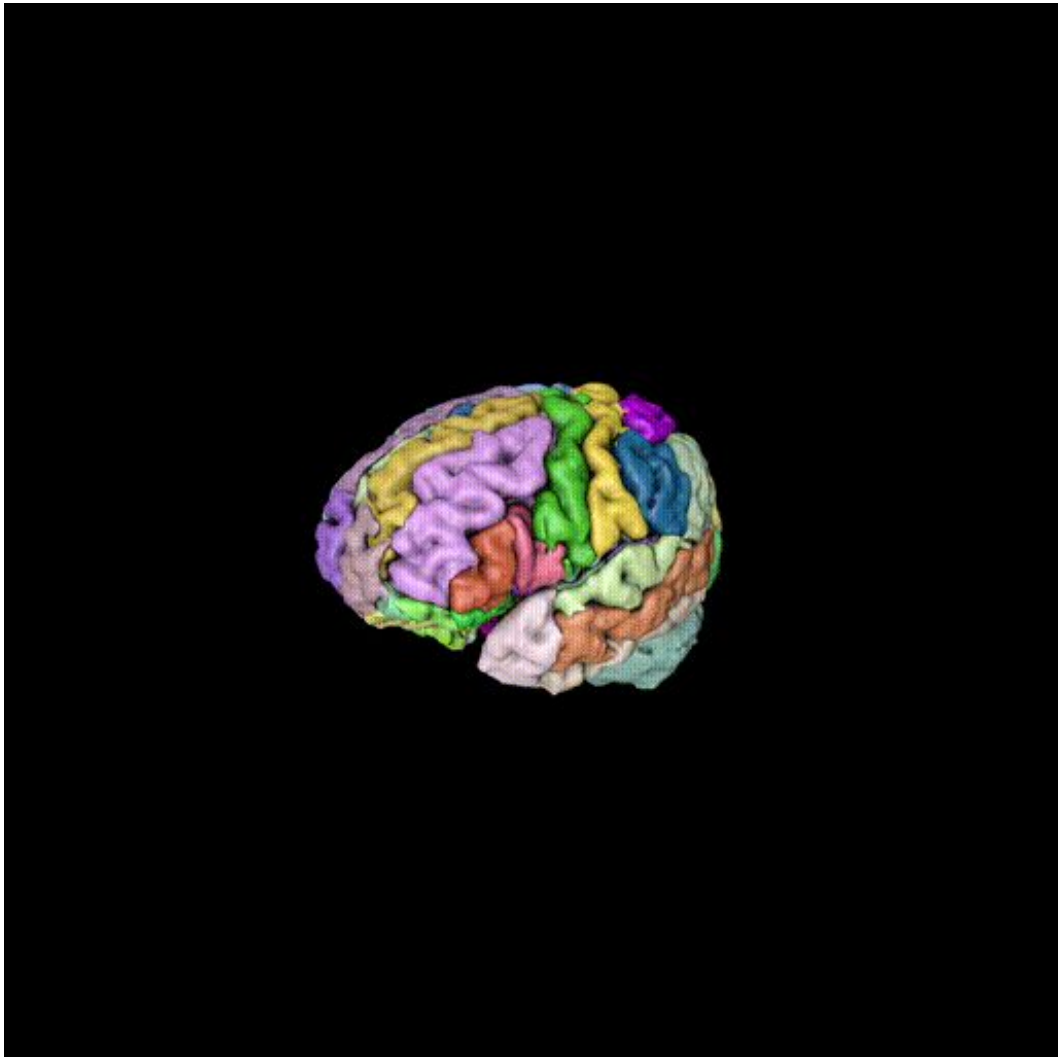
Interactive transpose conv:
https://distill.pub/2016/deconv-checkerboard/

A guide to convolution arithmetic for deep learning
https://arxiv.org/abs/1603.07285

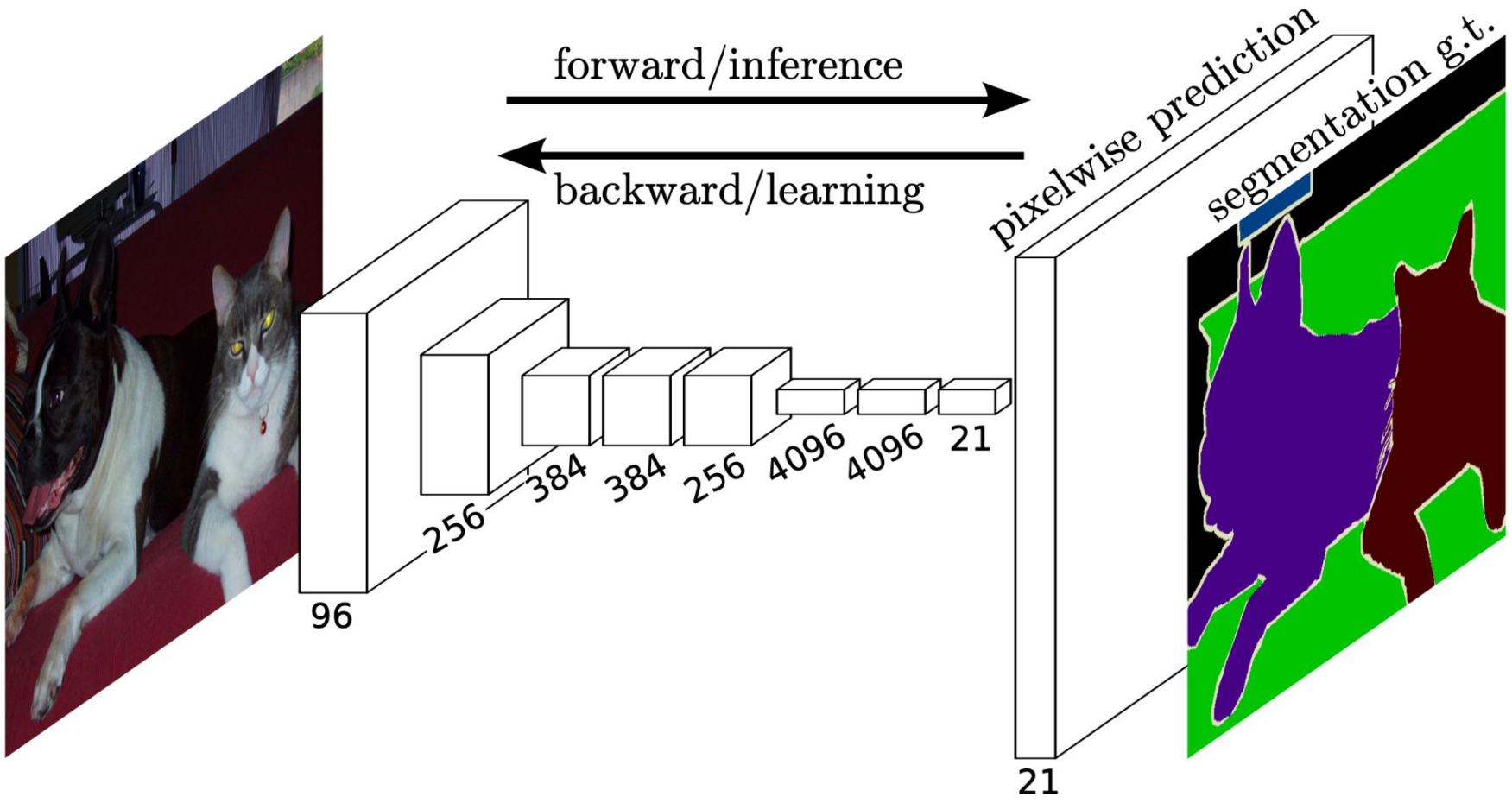
Segmentation



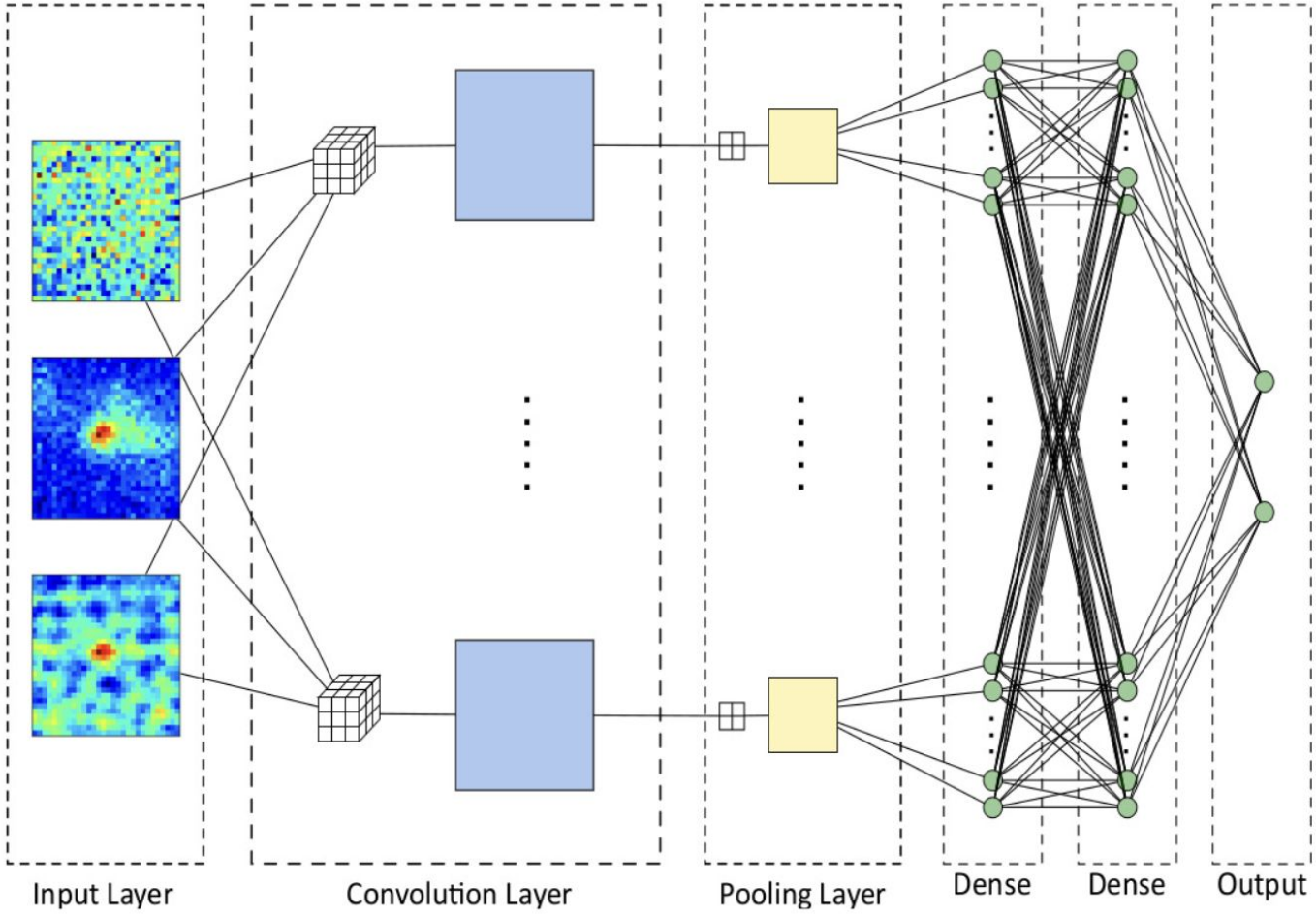




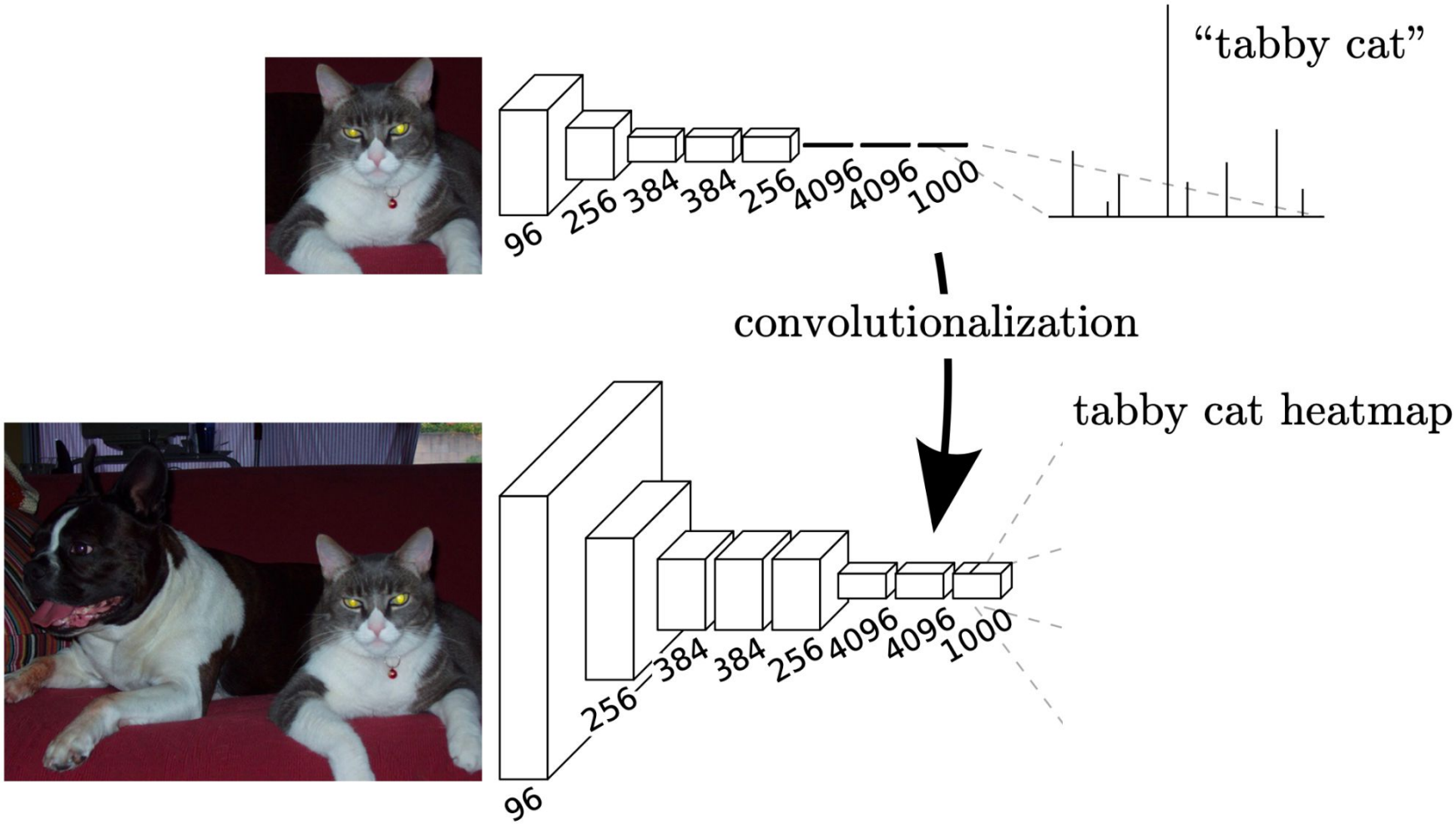
Fully convolutional network (FCN)



Fully connected layer -> conv layer



Fully connected layer -> conv layer

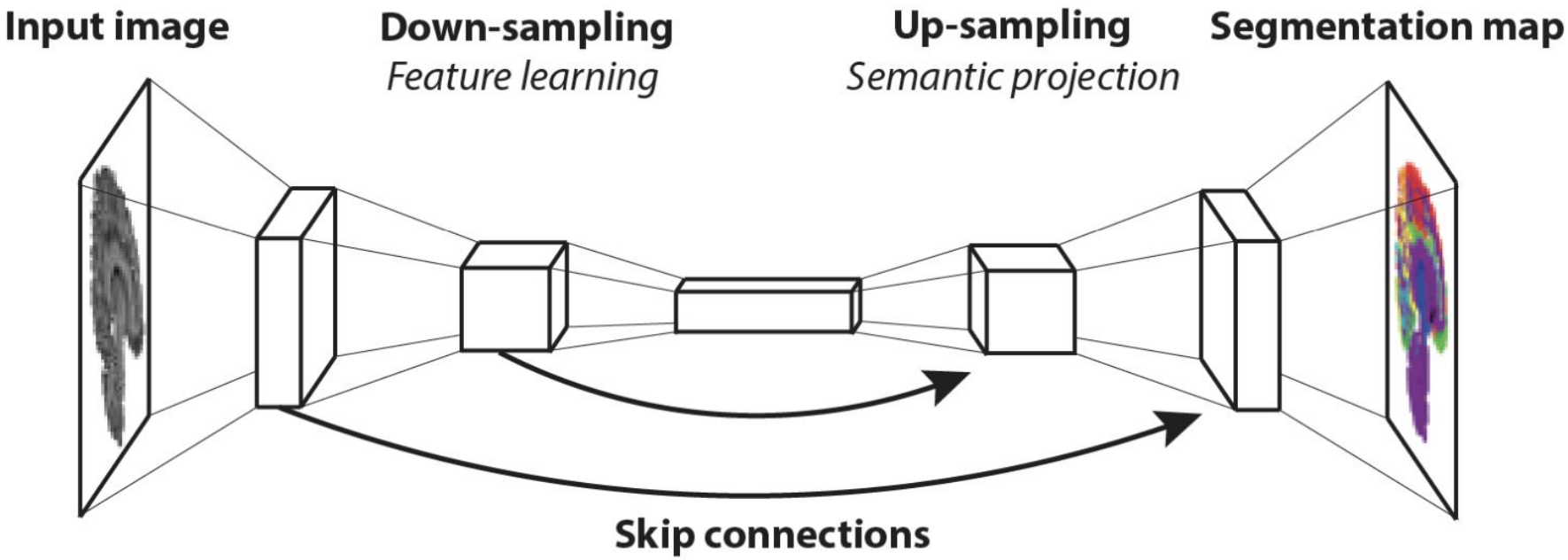


What does ‘fully convolutional’ mean?

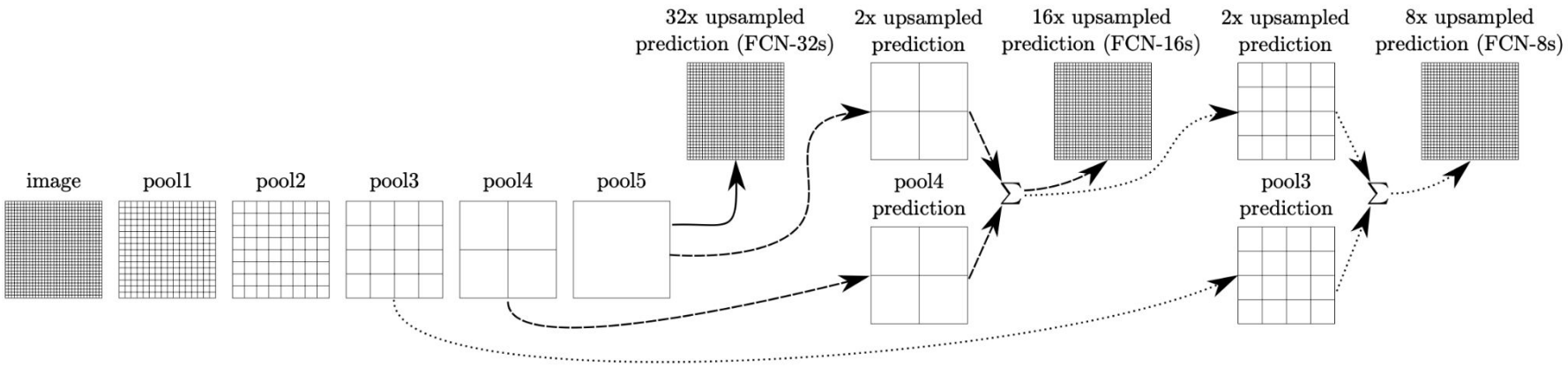
$$\mathbf{y}_{ij} = f_{ks}\left(\{\mathbf{x}_{si+\delta i, sj+\delta j}\}_{0 \le \delta i, \delta j \le k}\right)$$

$$f_{ks} \circ g_{k's'} = (f \circ g)_{k'+(k-1)s', ss'}.$$

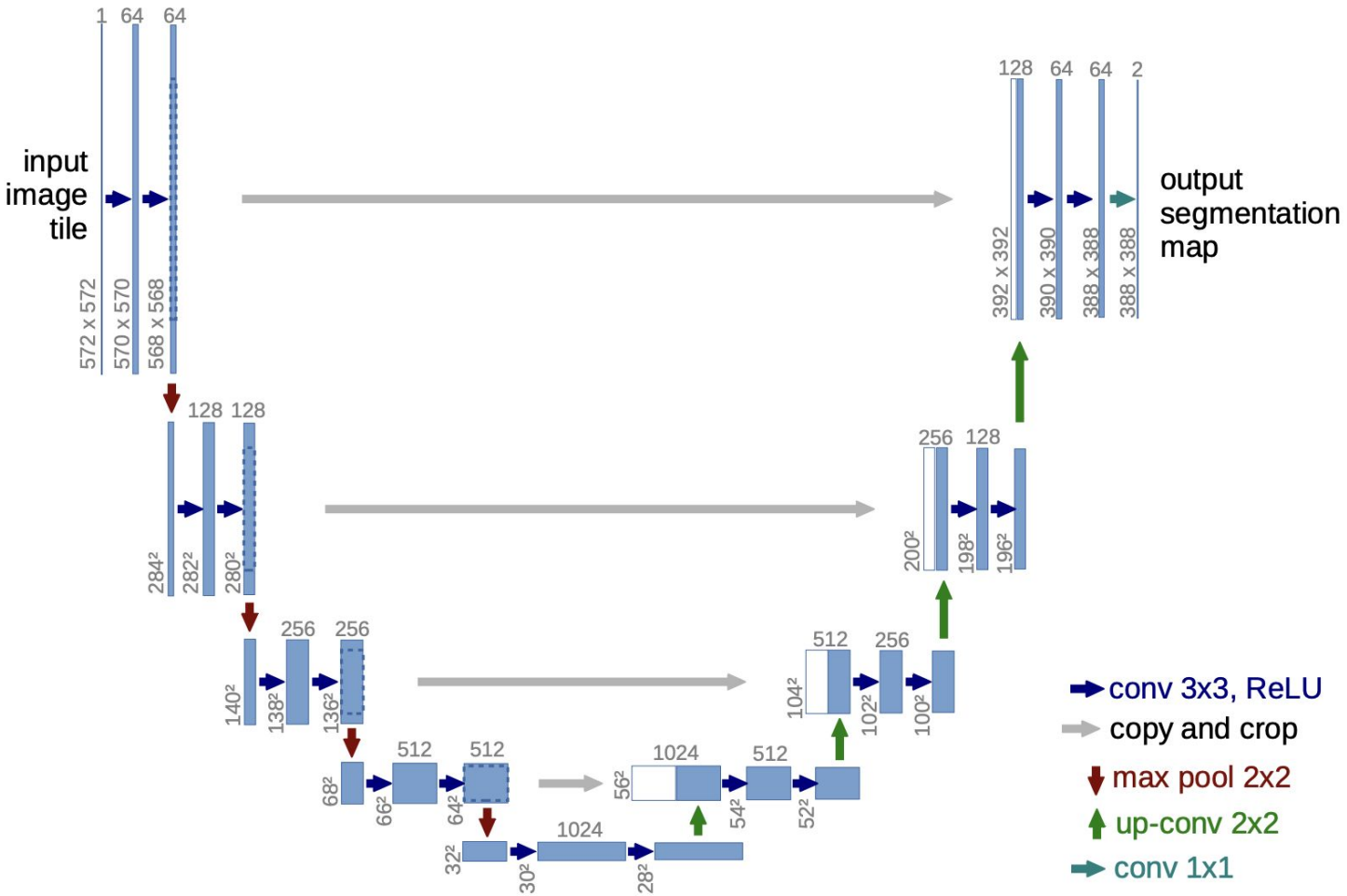
While a general deep net computes a general nonlinear function, a net with only layers of this form computes a nonlinear *filter*, which we call a *deep filter* or *fully convolutional network*.


Input image
Down-sampling
Feature learning
Up-sampling
Semantic projection
Segmentation map
Skip connections

Skip connections in FCNs (addition)



U-Nets


input image tile
output segmentation map
1 64 64
572 x 572
570 x 570
568 x 568
128 128
284²
282²
280²
256 256
140²
138²
136²
512 512
68²
66²
64²
1024
32²
30²
1024
28²
512
56²
512
54²
256
104²
102²
100²
256 128
200²
198²
196²
128 64 64 2
392 x 392
390 x 390
388 x 388
388 x 388
conv 3x3, ReLU
copy and crop
max pool 2x2
up-conv 2x2
conv 1x1