CNN segmentation: U-net

Ronneberger: U-net, MICCAI 2015

Jan Kybic

2020

Previous segmentation strategies

- Sliding window
- Predict a class of the central pixel
- Many patches per image $\rightarrow$ not many images needed
- Slow
- Patch size - localization accuracy vs. context

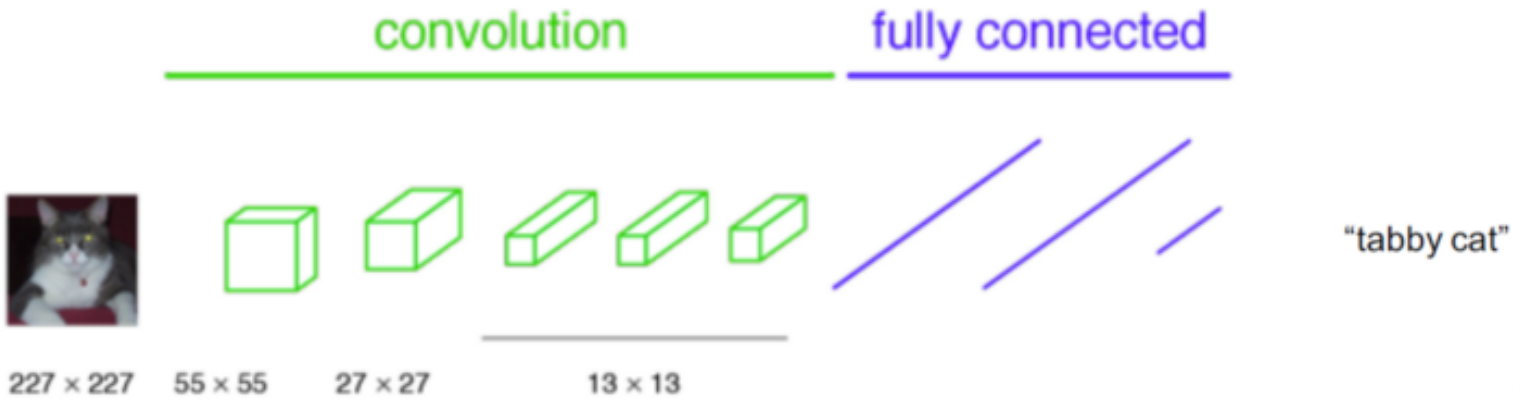
Previous segmentation strategies

- Sliding window
- Predict a class of the central pixel
- Many patches per image $\rightarrow$ not many images needed
- Slow
- Patch size - localization accuracy vs. context

Fully convolutional networks

- no fully connected layers, only convolutional layers
- output - a (smaller) image

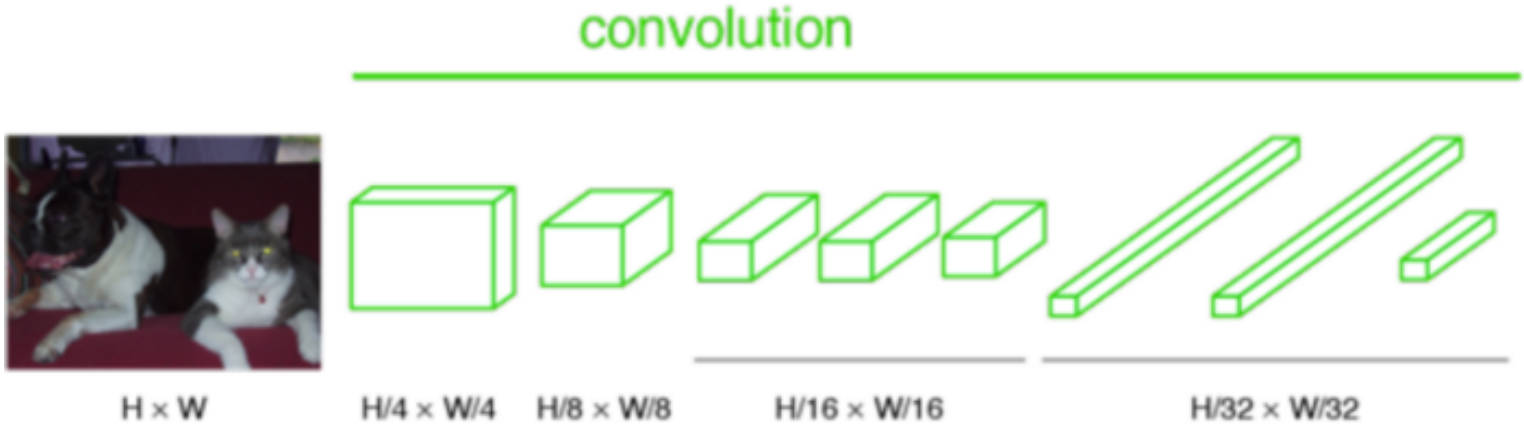
Classification network



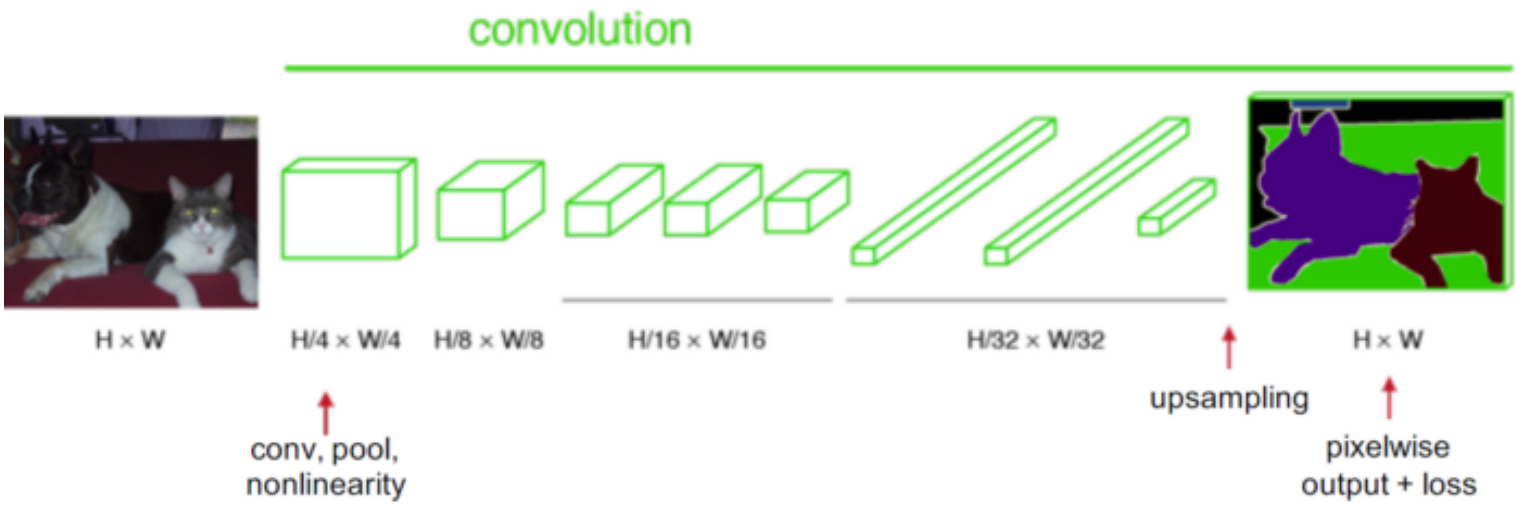
Tsang, towardsdatascience.com

Sik-Ho

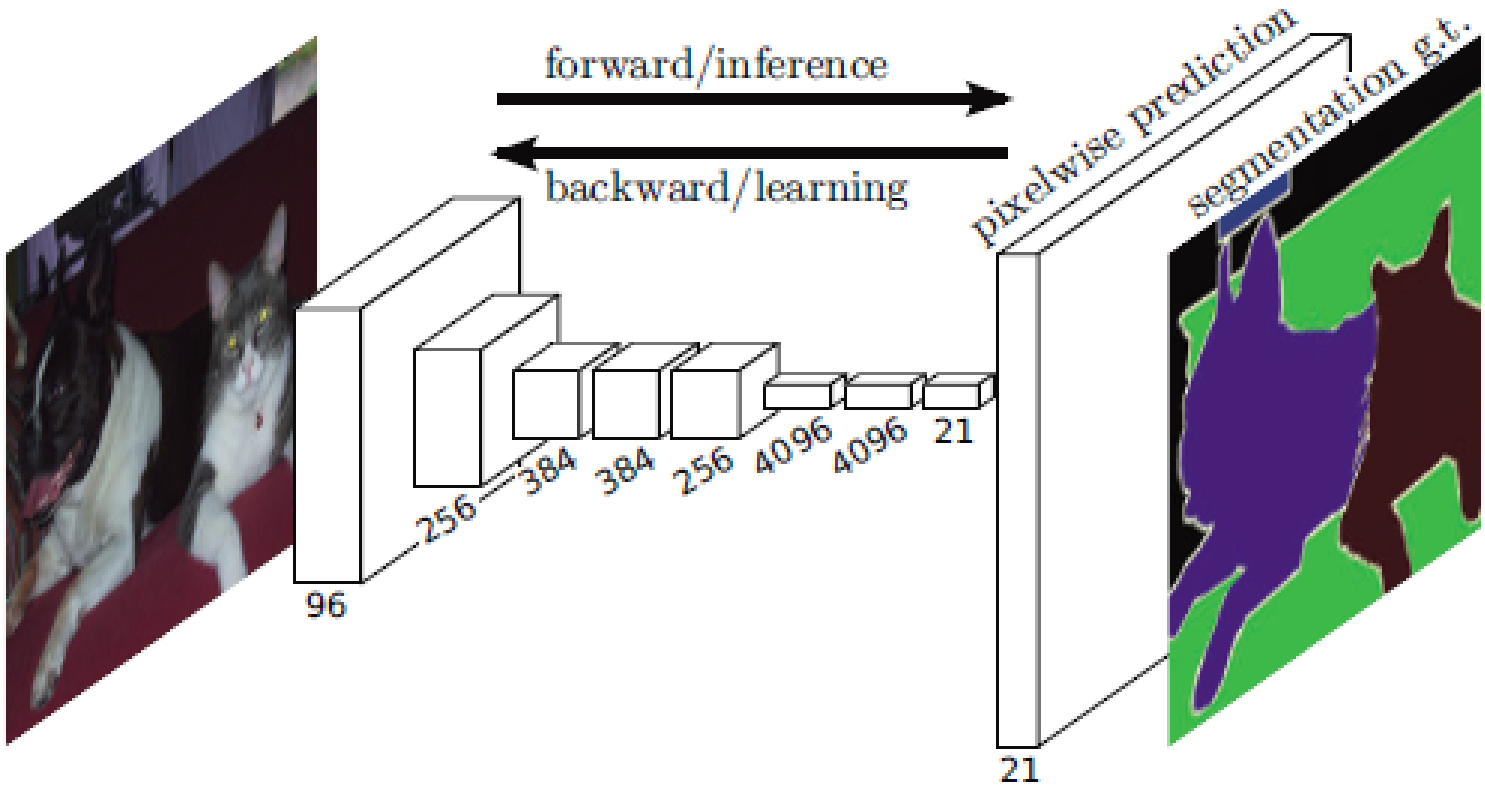
Fully convolutional network



Segmentation

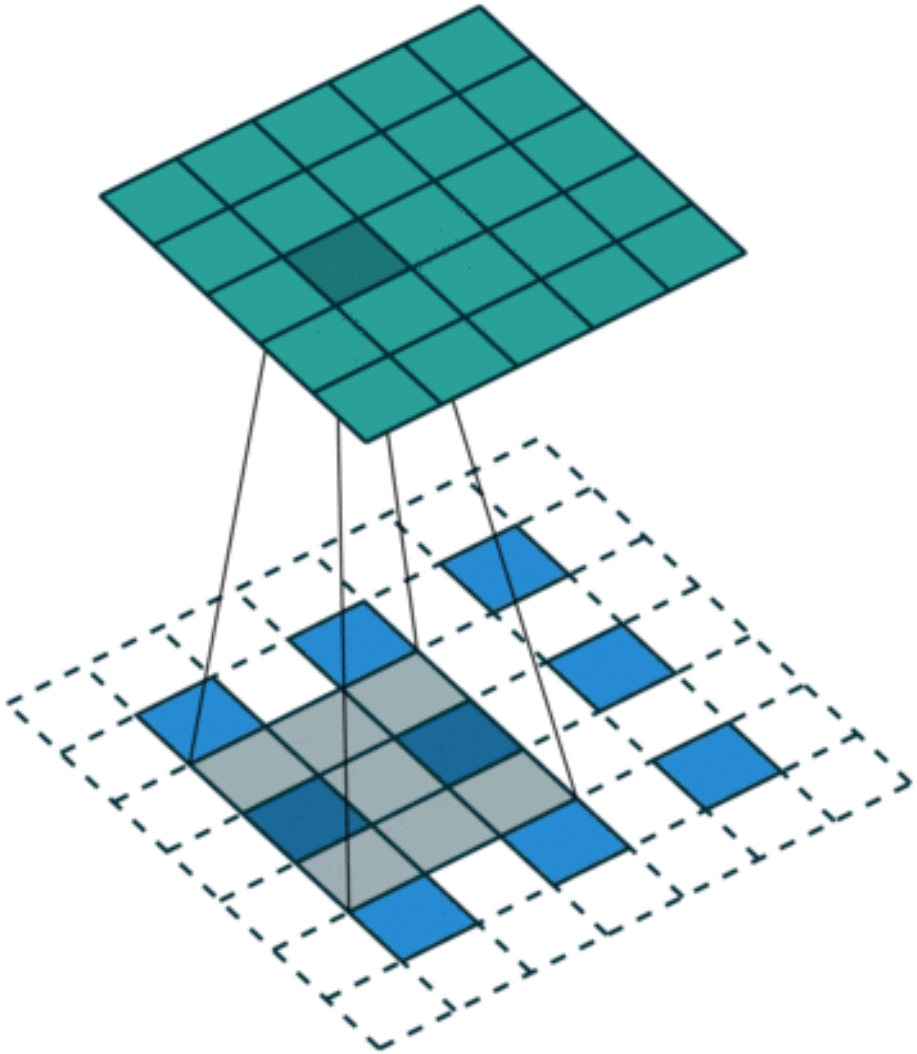


Fully convolutional structure

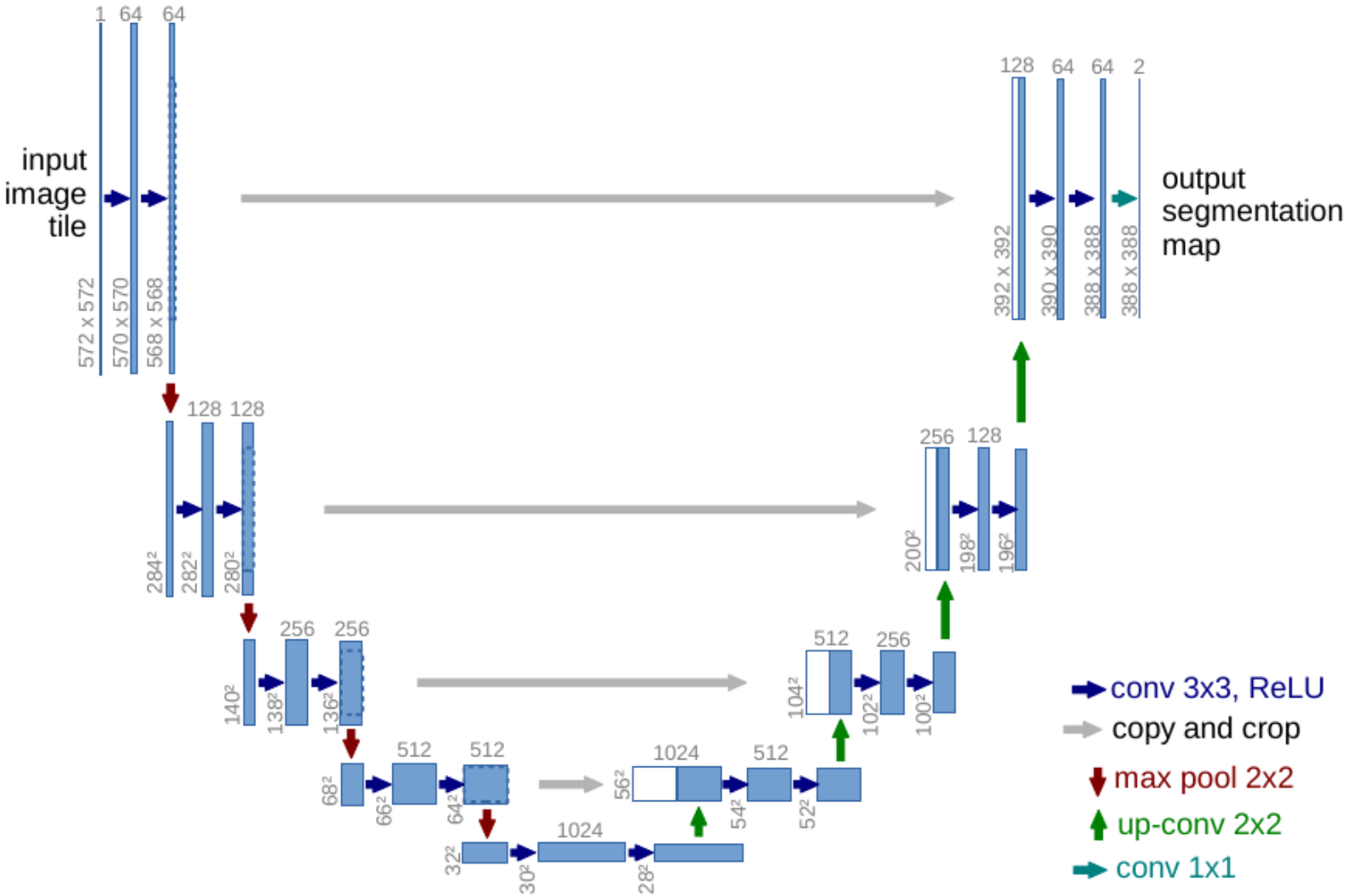


Upconvolution

Transposed convolution, (Deconvolution)



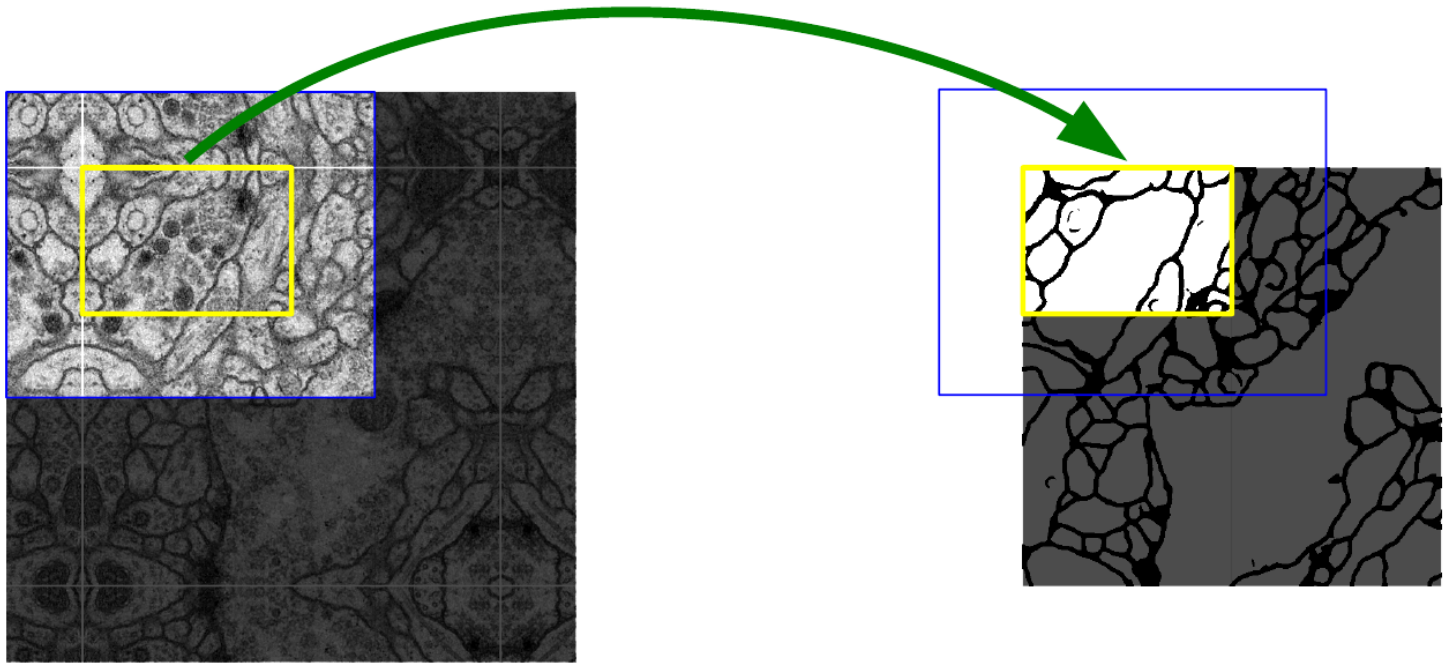
U-net structure



U-net notes

- no fully connected layers
- upsampling to get to the original resolution
- skip connections
- small kernels (3×3)
- only fully valid convolutions, no padding$\rightarrow$cropping
- 2×2max-pooling $\rightarrow$ fixed-size input, even size at all levels
- number of channels increase
- softmax in the final level
- "downsampling" block can be replaced
- extensive augmentation (elastic deformation)
- random weight initialization with $\sigma = \sqrt{2/N}$
- dropout

Tiling strategy



Loss function

- Cross-entropy

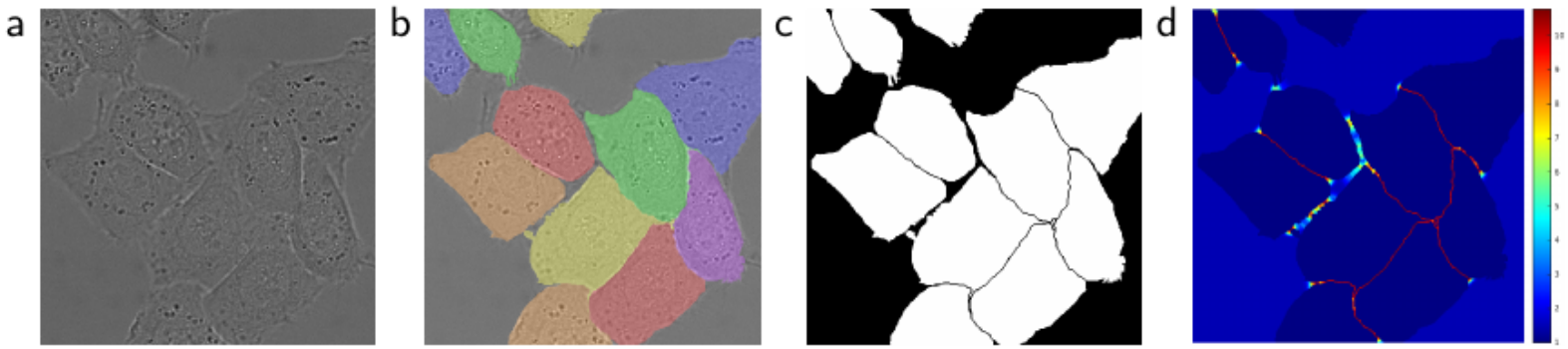
$$E = \sum_{\mathbf{x} \in \Omega} w(\mathbf{x}) \log(p_{\ell(\mathbf{x})}(\mathbf{x}))$$

The separation border is computed using morphological operations. The weight map is then computed as

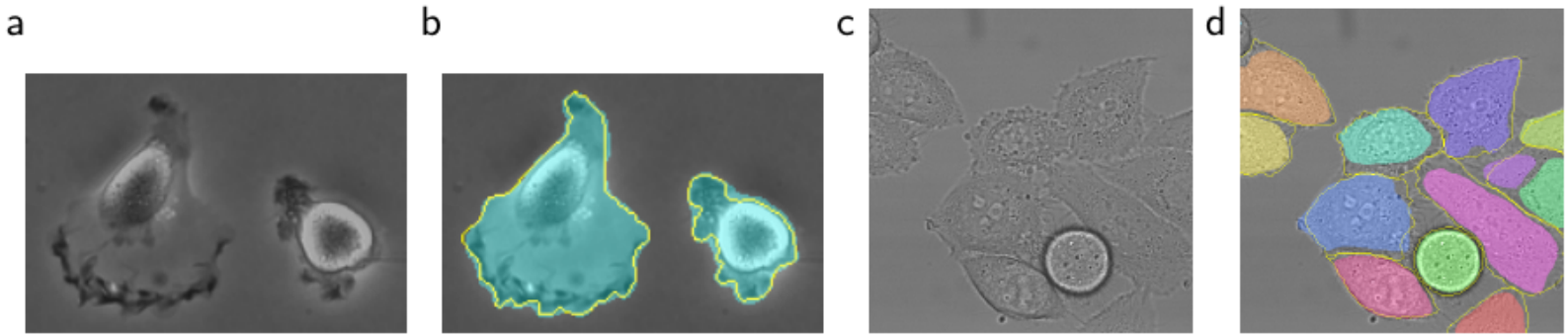
$$w(\mathbf{x}) = w_c(\mathbf{x}) + w_0 \cdot \exp\left(-\frac{(d_1(\mathbf{x}) + d_2(\mathbf{x}))^2}{2\sigma^2}\right) \tag{2}$$

where $w_c : \Omega \to \mathbb{R}$ is the weight map to balance the class frequencies, $d_1 : \Omega \to \mathbb{R}$ denotes the distance to the border of the nearest cell and $d_2 : \Omega \to \mathbb{R}$ the distance to the border of the second nearest cell. In our experiments we set $w_0 = 10$ and $\sigma \approx 5$ pixels.

Border weight example



Example segmentation



Quantitative results

IoU, cell tracking challenge

Name	PhC-U373	DIC-HeLa
IMCB-SG (2014)	0.2669	0.2935
KTH-SE (2014)	0.7953	0.4607
HOUS-US (2014)	0.5323	-
second-best 2015	0.83	0.46
u-net (2015)	**0.9203**	**0.7756**