U-Net: Convolutional Networks for Biomedical Image Segmentation

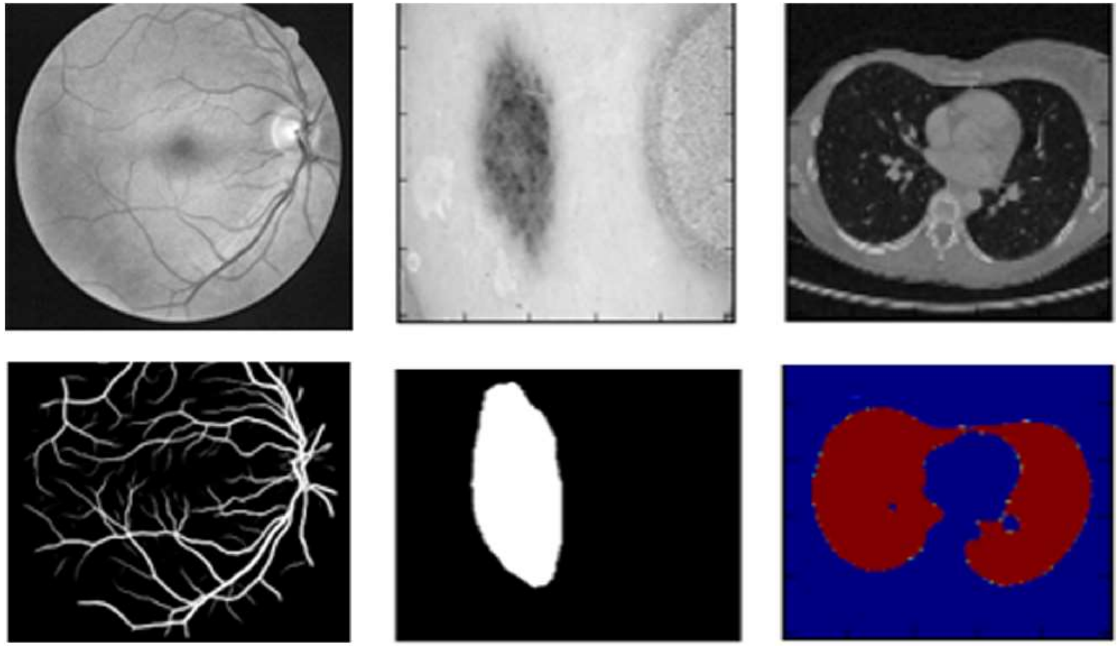
MICCAI 2015

Main problem : Semantic Segmentation

- Divide Image by semantic labels
- Classify objects semantically

Contribution

- Propose better architecture for semantic segmentation
- Can get better performance with very few training data



Overlap-tile strategy

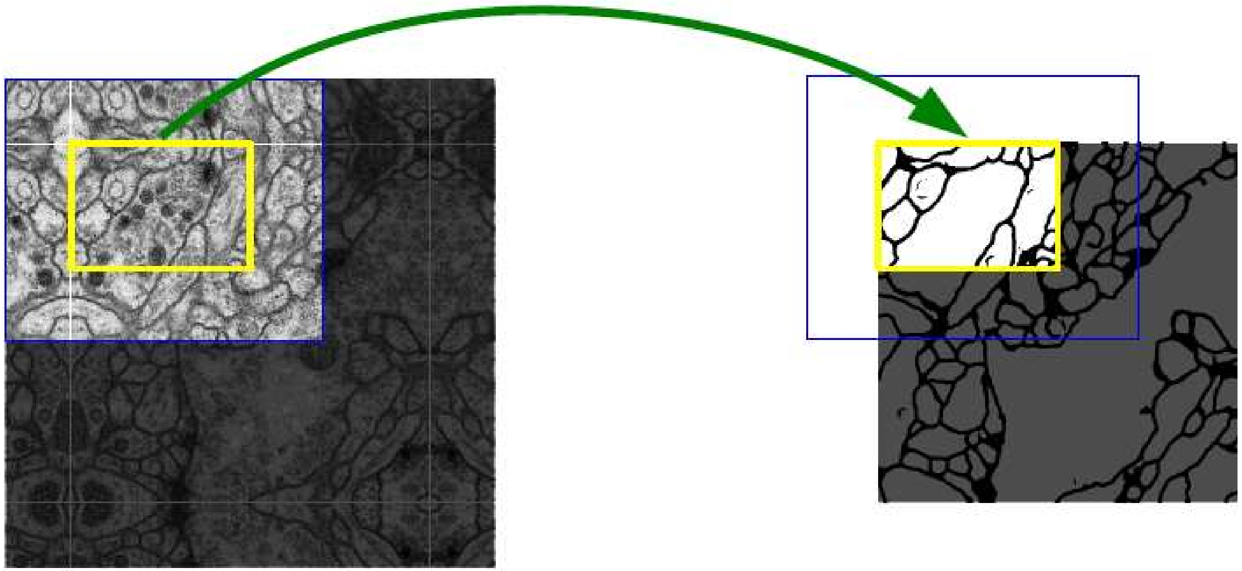
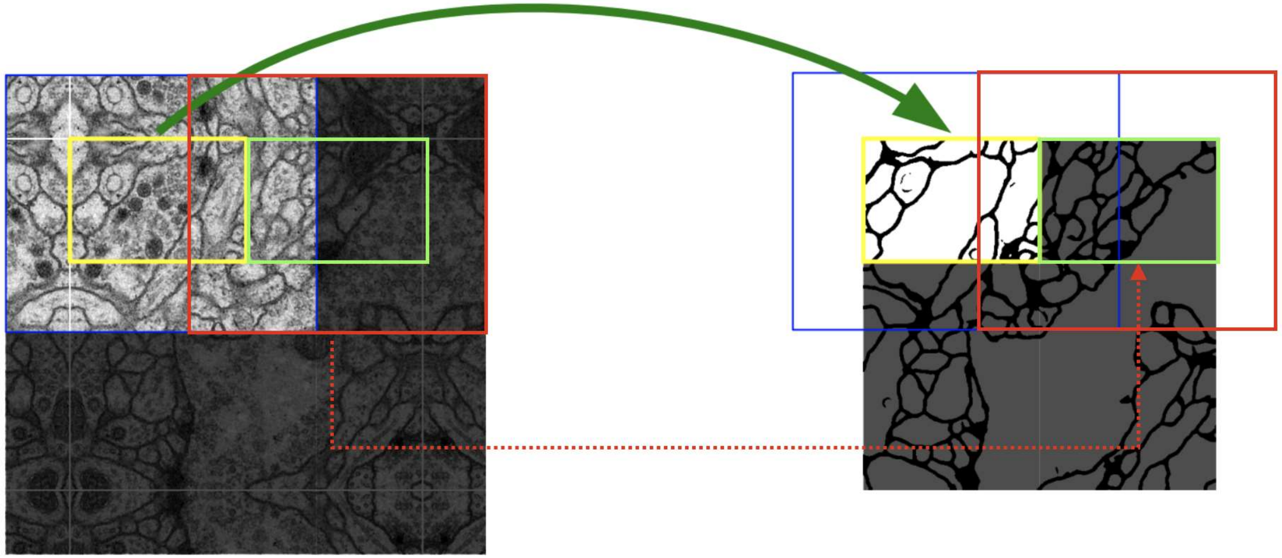


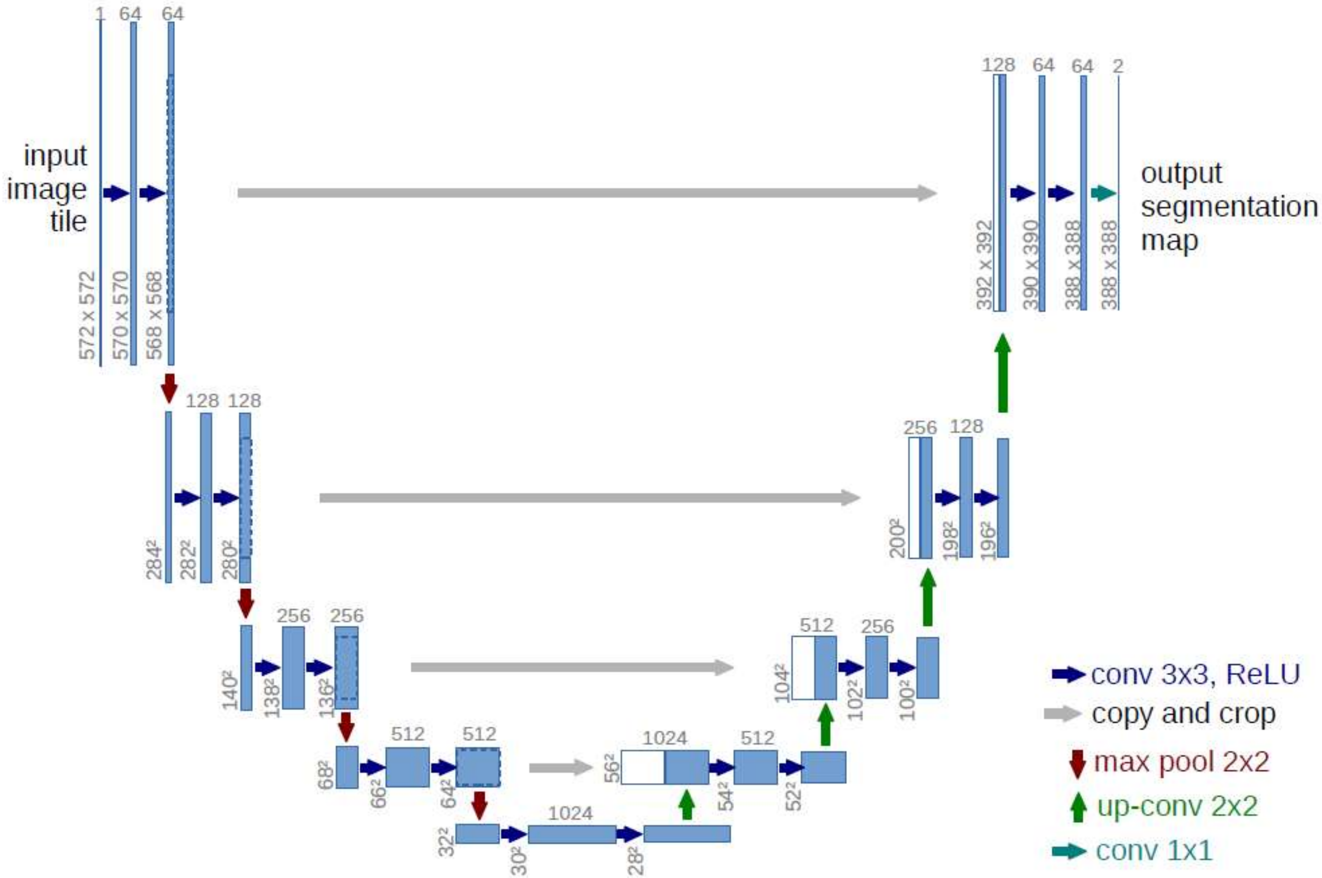
Fig. 2. Overlap-tile strategy for seamless segmentation of arbitrary large images (here segmentation of neuronal structures in EM stacks). Prediction of the segmentation in the yellow area, requires image data within the blue area as input. Missing input data is extrapolated by mirroring

Overlap-tile strategy

- Divide large image
- Ignore overlapped region
- Get seamless segmentation result



Structure



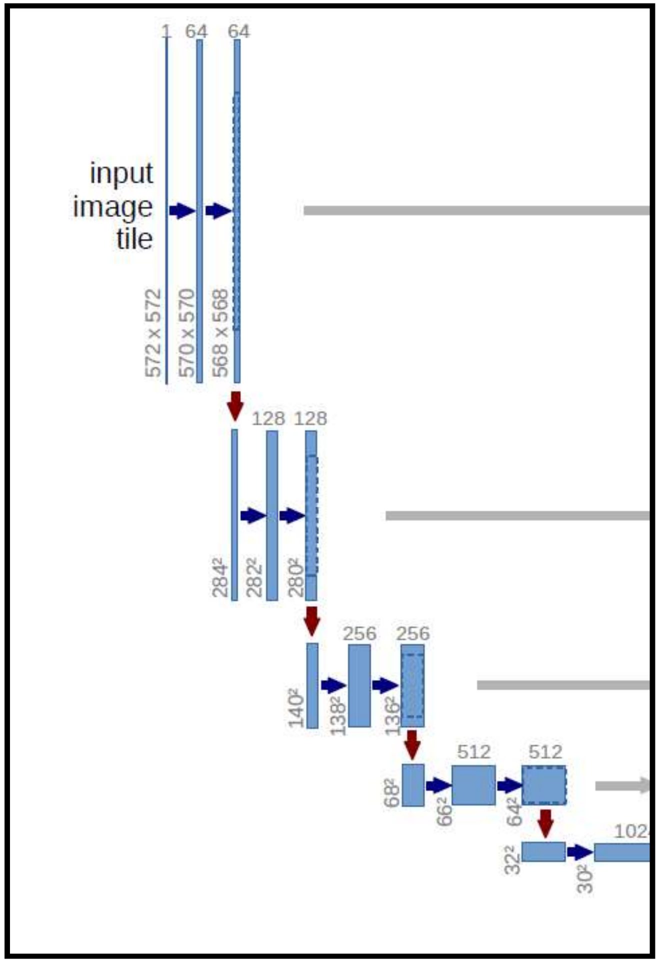
- Encoder – Decoder structure
- Use skip connection
- Better to use context information

Structure : Encoder(Contracting path)

- Encoding image's information and downsample image

Path	Propagation (rightward)	Downsampling (downward)
Module	3x3 CNN (unpadded)	2x2 max pool (stride 2)

In downsampling step, double the number of feature channels!



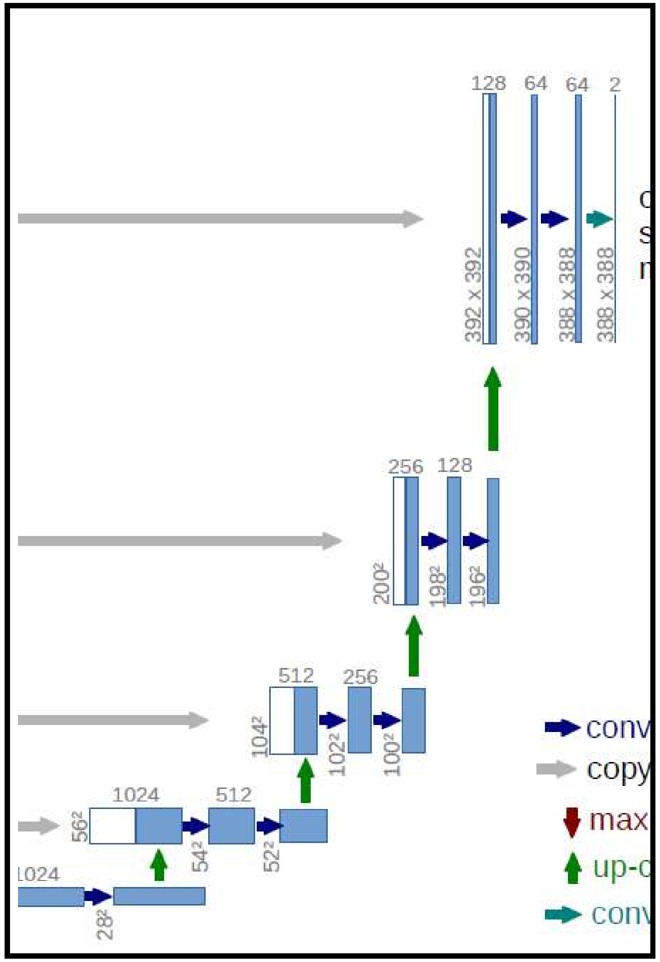
Structure : Decoder(Expansive path)

- Decoding image's information and upsample image

Path	Propagation (rightward)	Upsampling (downward)
Module	3x3 CNN (unpadded)	2x2 upconvolution +concat encoder's feature maps

Cropped feature map of contracting

- Using low level + high level feature of image with multi-scale manner



Training detail : loss function

- Pixel-wise softmax + weight

$$p_k(\mathbf{x}) = \exp(a_k(\mathbf{x})) / \left(\sum_{k'=1}^{K} \exp(a_{k'}(\mathbf{x})) \right)$$

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

Training detail : Weight map

$$w(\mathbf{x}) = \underbrace{w_c(\mathbf{x})}_{\text{Class-related term}} + \underbrace{w_0 \cdot \exp\left(-\frac{(d_1(\mathbf{x}) + d_2(\mathbf{x}))^2}{2\sigma^2}\right)}_{\text{Distance-related term}}$$

Training detail : Weight map

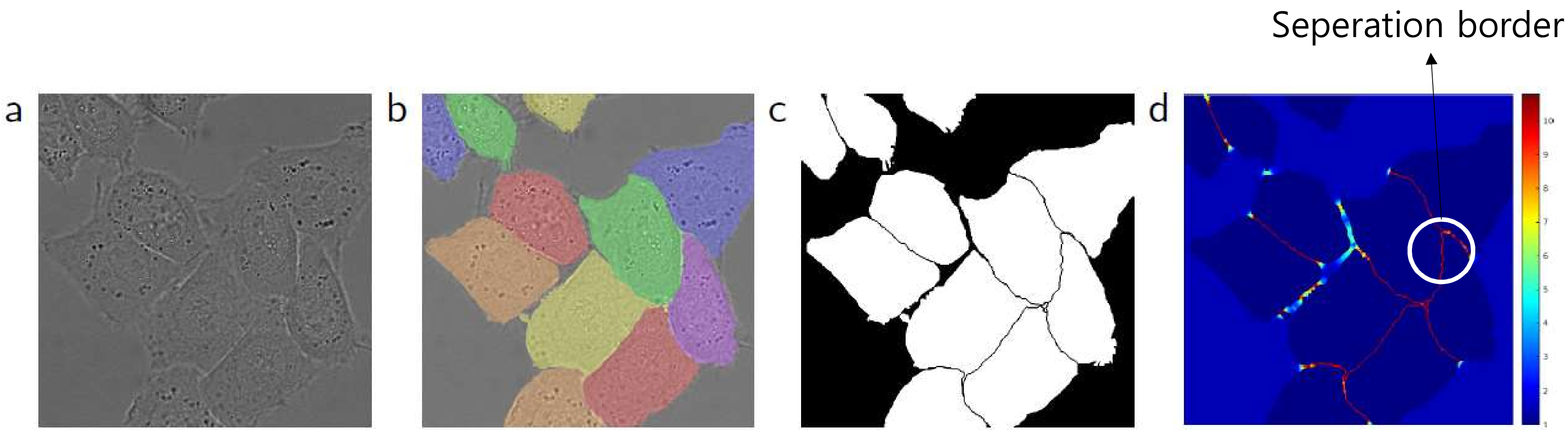


Fig. 3. HeLa cells on glass recorded with DIC (differential interference contrast) microscopy. (**a**) raw image. (**b**) overlay with ground truth segmentation. Different colors indicate different instances of the HeLa cells. (**c**) generated segmentation mask (white: foreground, black: background). (**d**) map with a pixel-wise loss weight to force the network to learn the border pixels.

Training detail : Data augmentation

- Rotation
- Translation
- Elastic deformation

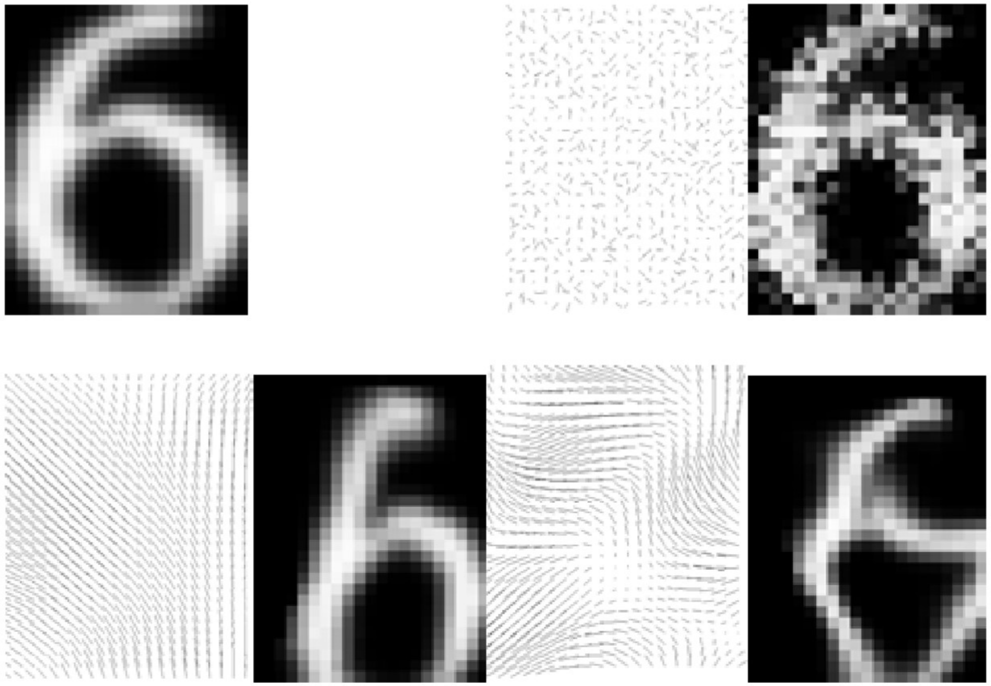
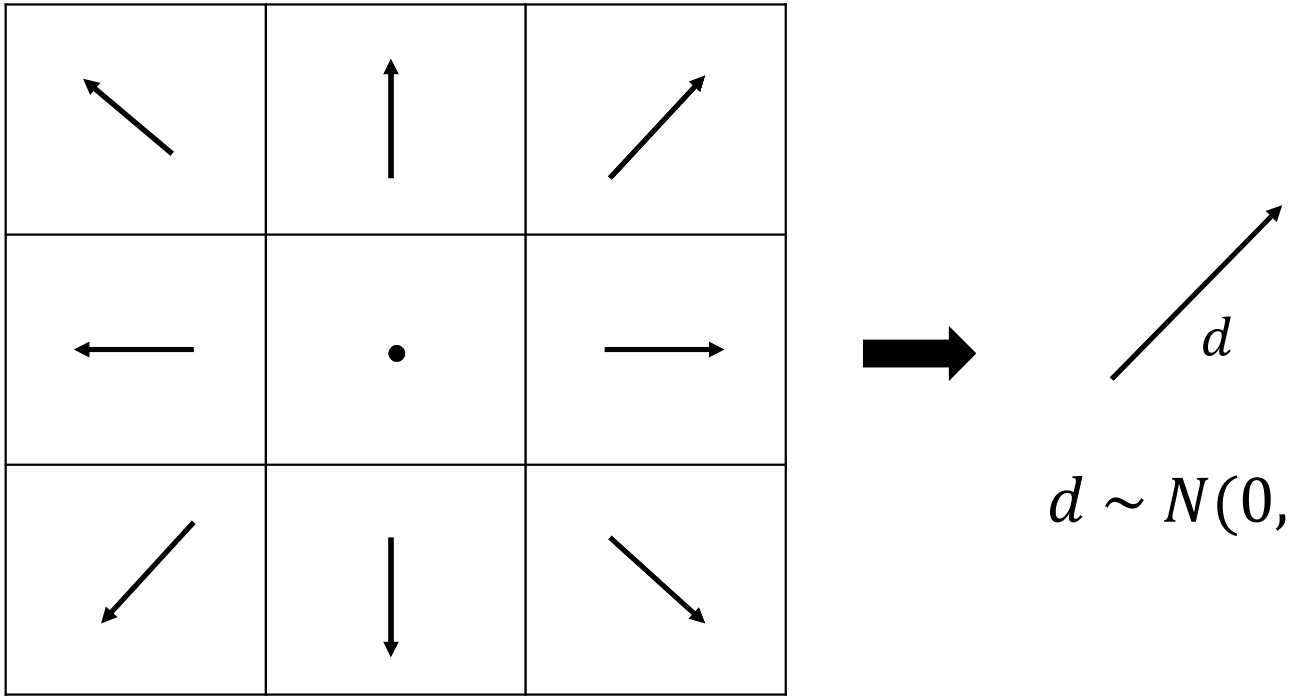


Figure 2. Top left: Original image. Right and bottom: Pairs of displacement fields with various smoothing, and resulting images when displacement fields are applied to the original image.

Training detail : Elastic deformation



$$d \sim N(0,10)$$

Training detail : Elastic deformation

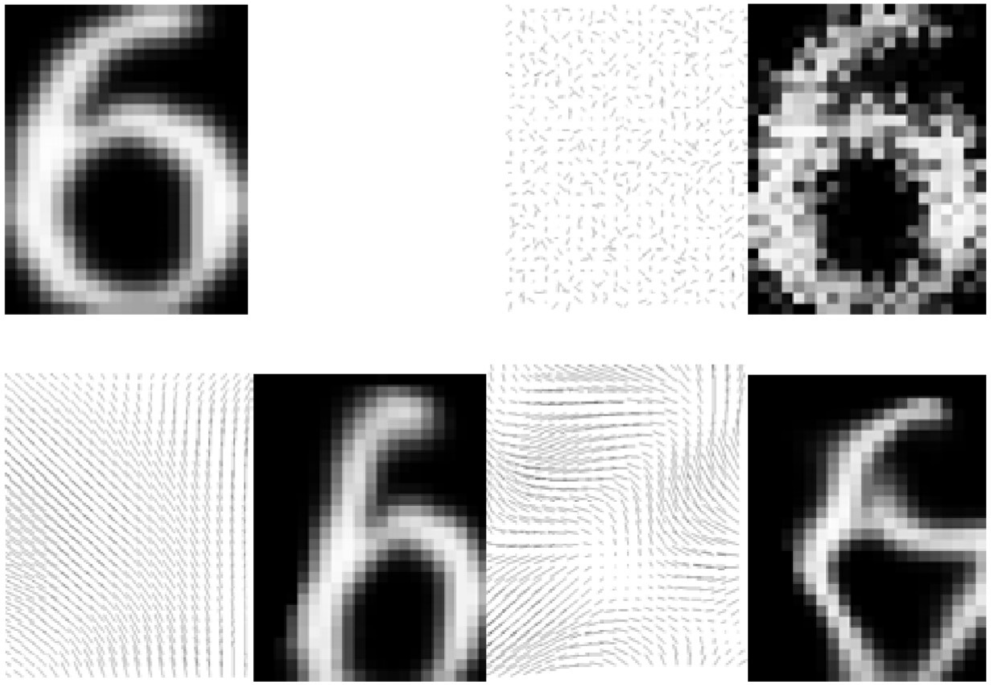


Figure 2. Top left: Original image. Right and bottom: Pairs of displacement fields with various smoothing, and resulting images when displacement fields are applied to the original image.

Result

- Achieved SOTA performance in semantic segmentation about medical image

- Propose brilliant architecture and augmentation method

Table 1. Ranking on the EM segmentation challenge [14] (march 6th, 2015), sorted by warping error.

Rank	Group name	Warping Error	Rand Error	Pixel Error
	** human values **	0.000005	0.0021	0.0010
1.	u-net	**0.000353**	0.0382	0.0611
2.	DIVE-SCI	0.000355	0.0305	0.0584
3.	IDSIA [1]	0.000420	0.0504	0.0613
4.	DIVE	0.000430	0.0545	**0.0582**
⋮				
10.	IDSIA-SCI	0.000653	**0.0189**	0.1027

Table 2. Segmentation results (IOU) on the ISBI cell tracking challenge 2015.

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**

Result

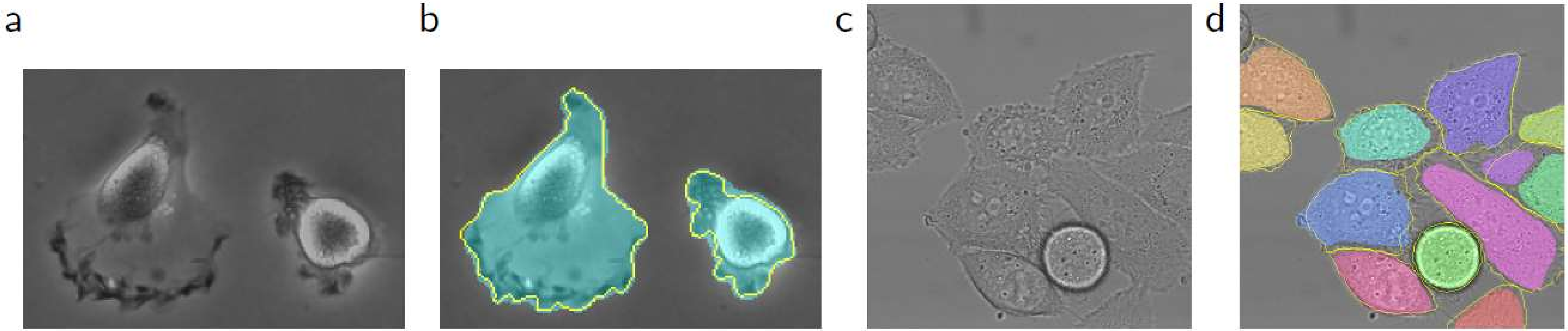
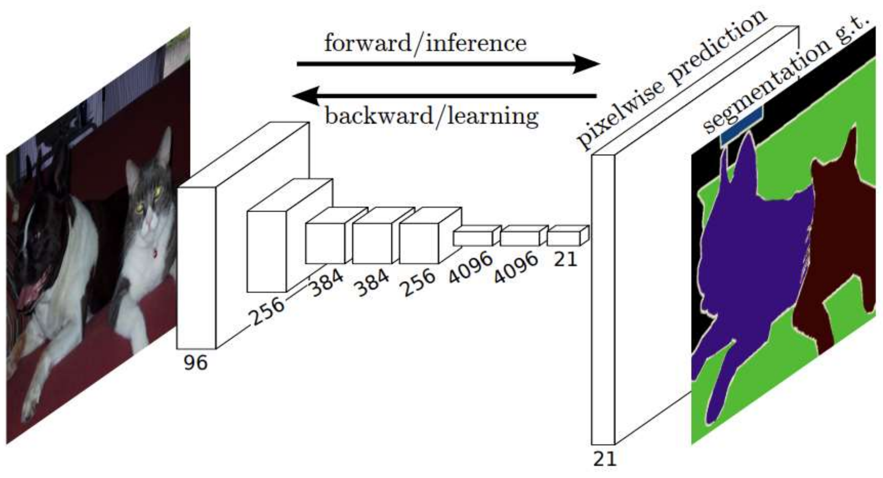


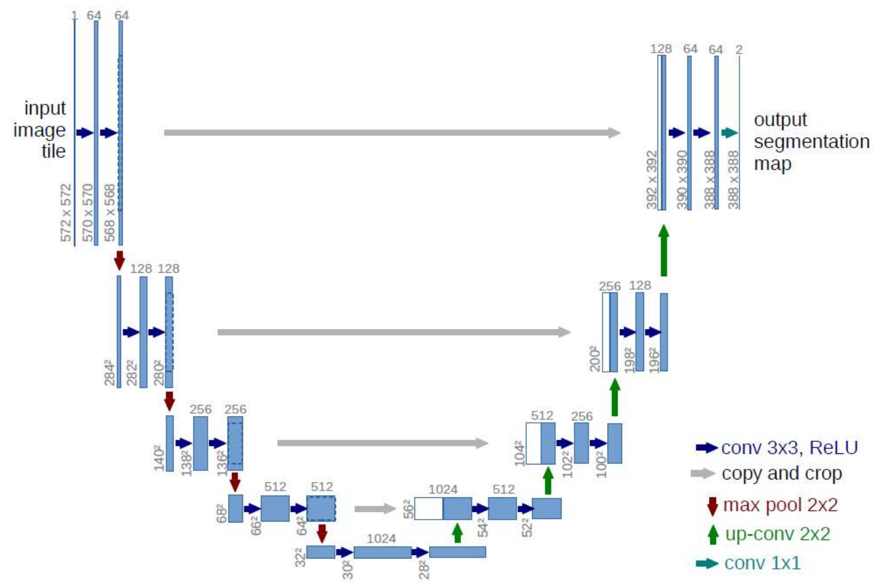
Fig. 4. Result on the ISBI cell tracking challenge. (**a**) part of an input image of the "PhC-U373" data set. (**b**) Segmentation result (cyan mask) with manual ground truth (yellow border) (**c**) input image of the "DIC-HeLa" data set. (**d**) Segmentation result (random colored masks) with manual ground truth (yellow border).

Conclusion & Meaning

- Propose brilliant architecture for biomedical image semantic segmentation
- Propose new efficient & good structure for deep neural networks for semantic segmentation



Fully-Convlutional Network(FCN)



U-Net