

# U-Net

## Convolutional Networks for Biomedical Image Segmentation

Unet은 FCN 모델 다음으로 주목받은 segmentation 모델

바이오 의료 이미지 분야에 특화된 FCN 기반 segmentation 모델  
(FCN논문 내용을 참조해서 읽으면 더 효과적)

의료분야에서 Unet 베이스 모델 기반의 segmentation을 애용해 병변 영역 확인

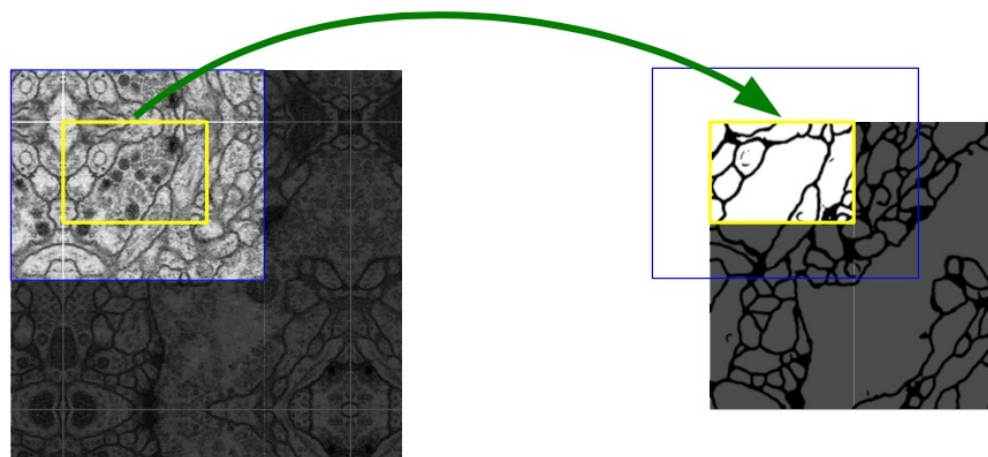
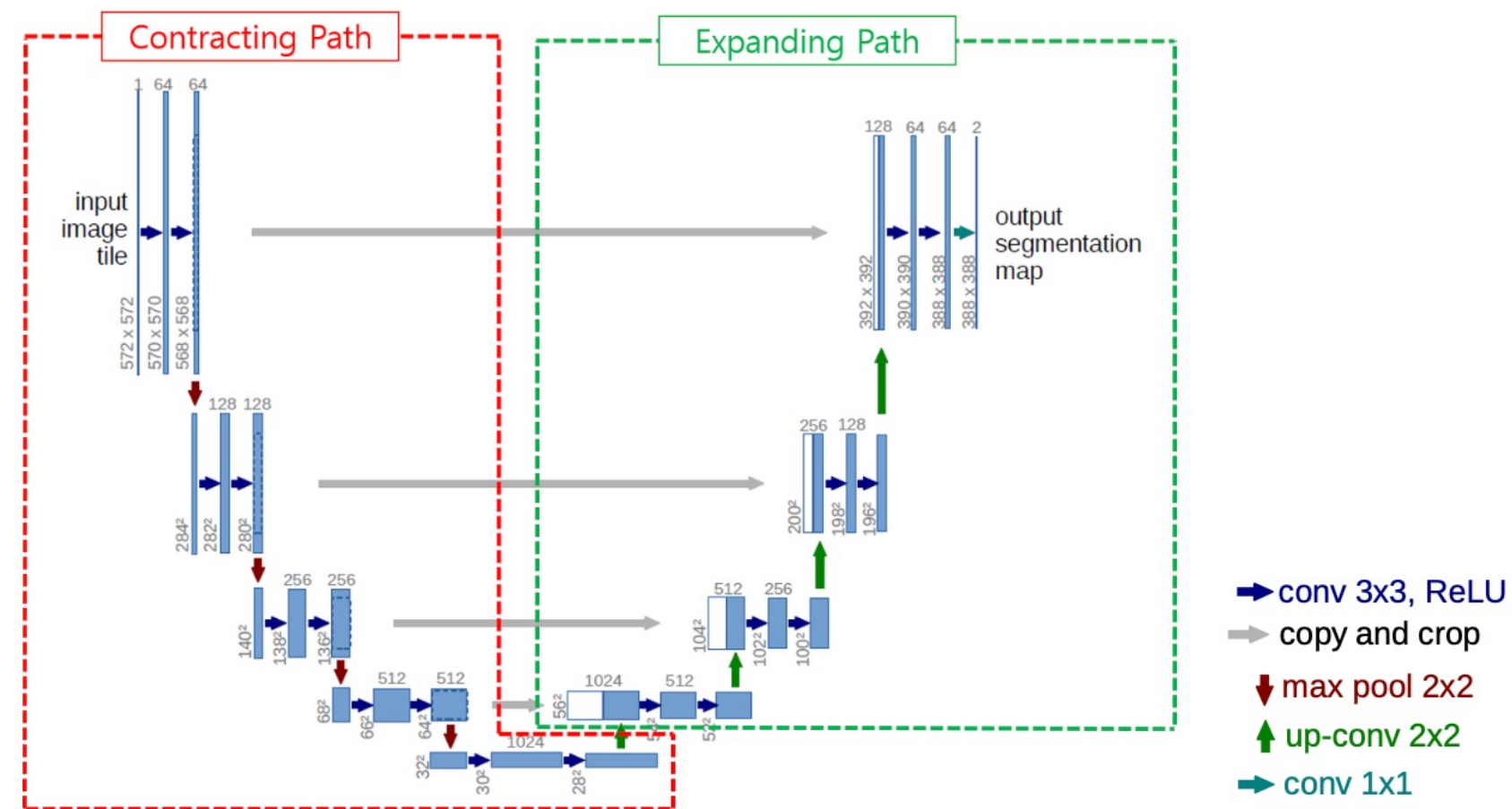
=> so, 의료분야에서 딥러닝 기반 segmentation 모델을 이용하려면 Unet을 잘 파악해야함

# Abstract

**Abstract.** There is large consent that successful training of deep networks requires many thousand annotated training samples. In this paper, we present a network and training strategy that relies on the strong use of data augmentation to use the available annotated samples more efficiently. The architecture consists of a contracting path to capture context and a symmetric expanding path that enables precise localization. We show that such a network can be trained end-to-end from very few images and outperforms the prior best method (a sliding-window convolutional network) on the ISBI challenge for segmentation of neuronal structures in electron microscopic stacks. Using the same network trained on transmitted light microscopy images (phase contrast and DIC) we won the ISBI cell tracking challenge 2015 in these categories by a large margin. Moreover, the network is fast. Segmentation of a 512x512 image takes less than a second on a recent GPU. The full implementation (based on Caffe) and the trained networks are available at <http://lmb.informatik.uni-freiburg.de/people/ronneber/u-net>.

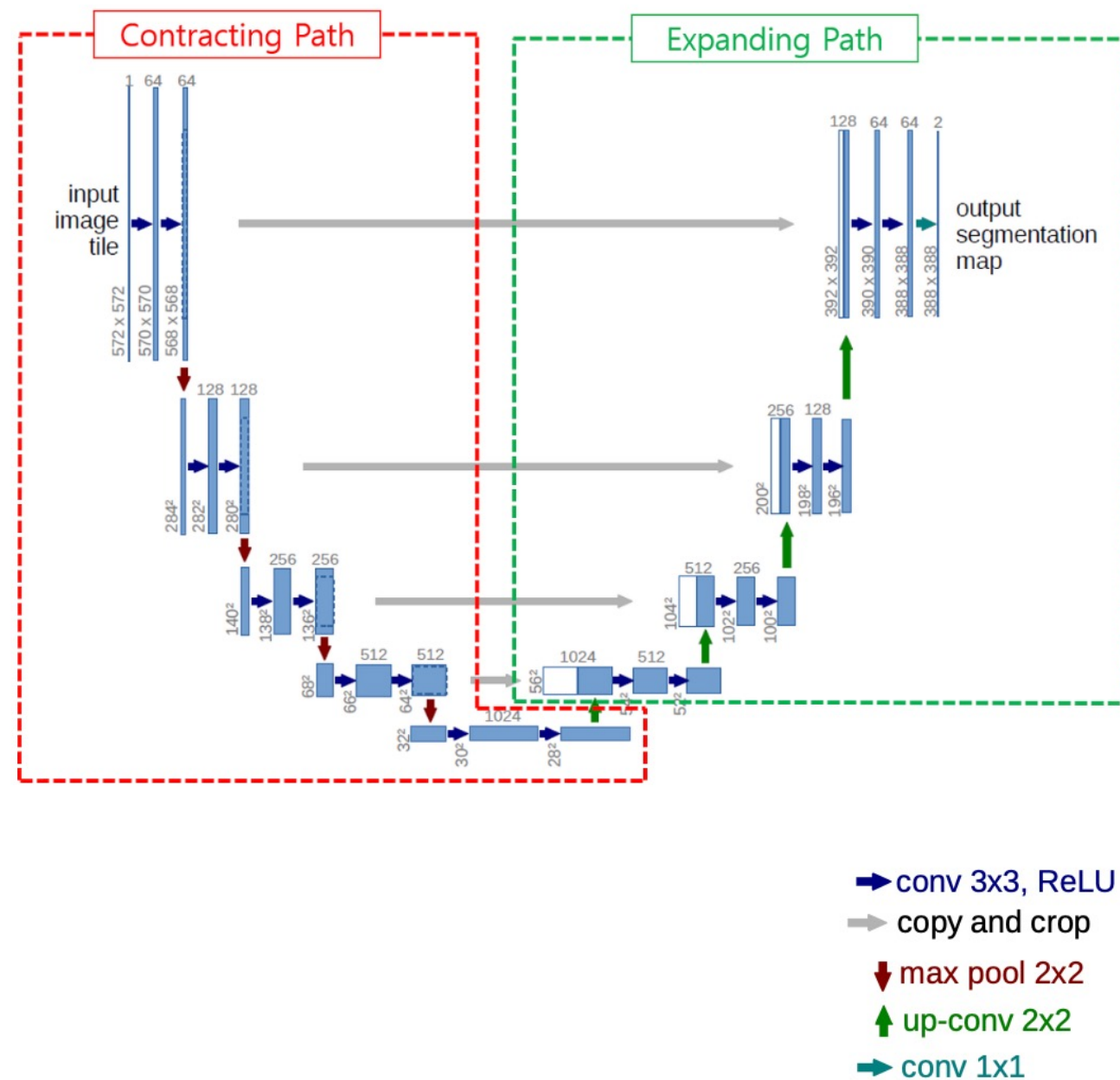
- annotated된 샘플 데이터를 더 효과적으로 사용하기 위해 **강력한 Data Augmentation의 이용**
- 아키텍처 구조 : '**contracting path**'를 통해 context 정보를 구하고, contracting path와 symmetric한 '**expanding path**'를 통해 원본 크기에 맞게 upsampling 하여 segmentation 작업 진행하는 구조
- ISBI cell tracking challenge 2015에 참가해 이전보다 압도적으로 좋은 결과로 우승
- 속도 매우 빠름

# Introduction



- 기존 CNN은 single classification task에 사용되지만, biomedical image processing 분야에서는 한 이미지 내에 모든 pixel을 classification을 하는 segmentation task가 중요
- biomedical tasks를 위한 image data 수가 충분하지 않다는 문제 존재.
  - 데이터 수 문제를 해결하기 위해
    - > (Unet 전) patch 기반의 학습(Sliding-window setup) : Sliding window 기반 segmentation 모델 : ISBI 2012 우승
    - > Sliding-window setup 방식의 두가지 문제점 : 속도 & trade-off (local acc – context 사용)
- 그림1 : expanding path 부분 보면 기존 FCN보다 더 많은 conv filter channel 사용
- 그림2 : over-tile 전략을 이용해 input image resolution에 제한 없이 segmentation 학습 가능해짐

# Network Architecture



## Contracting path

- Two 3×3 Conv filter (stride=1) with zero padding of input feature map
  - feature map size-reduction
  - 2×2 max-pooling
  - feature map size-reduction (to Half)
- => 반복

## Expansive path

- 2x2 De-Conv filter (channel수 절반으로 줄임)
- Crop & Copy and Concatentation
- 1×1 conv filter를 이용해 388×388×2 feature map으로 만듦

# Training

## Loss function

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

- SoftMax

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

- Cross-Entropy

$$\text{Cross Entropy} = \mathcal{C} = \sum_{\mathbf{x} \in \Omega} \log(p_{\ell(\mathbf{x})}(\mathbf{x}))$$

- Weight Map

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

## Weight initialization

- Gaussian distribution initialization (=standard deviation distribution)

- Standard deviation:  $\sqrt{\frac{2}{N}}$

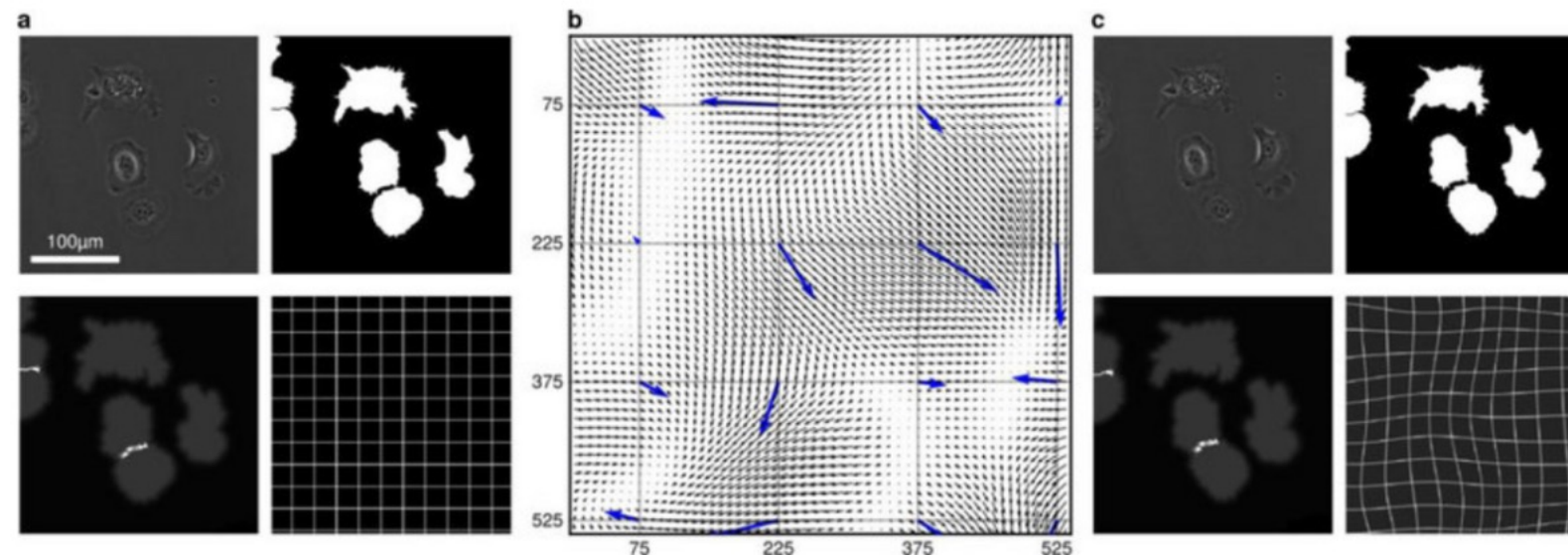
- $N = \text{conv filter size} \times \text{conv filter channel}$



# Training

## Data Augmentation

- Shift : (상하좌우) 이동
- Rotation : 회전
- Gray value : 회색값
- Random elastic deformation : 3x3짜리 random displacement vectors를 사용해서 smooth한 변형이미지 생성



# Experiments

Rank	Group name	Warping Error	Rand Error	Pixel Error
	** human values **	0.000005	0.0021	0.0010
1.	u-net	<b>0.000353</b>	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	<b>0.0582</b>
⋮				
10.	IDSIA-SCI	0.000653	<b>0.0189</b>	0.1027

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)	<b>0.9203</b>	<b>0.7756</b>

Table1. EM segmentation challenge 결과

\*Pixel Error : prediction과 ground truth에 pixel마다 할당된 class들끼리 불일치하는 pixel들의 전체 pixel 기준 비율

\*Warping Error & Rand Error : digital topology error 관련

Table2. 2014, 2015 cell tracking challenge IOU 성능평가 결과

\*PhC-U373 데이터셋

\*DIC-HeLa 데이터셋



# Conclusion

- Unet은 biomedical segmentation application 분야에서 뛰어난 성능을 가지고 있음
- 적은 이미지 셋으로도 합리적인 학습 가능
- NVidia Titan GPU(6GB)를 사용해 짧은시간 동안(10시간) 학습