

Image Segmentation

2024-04-03

이번 차시는 Image Segmentation을 수행하는 FCN과 U-Net을 알아보시다.

소속 : CV 2조

성명 : 박석우

목차

1	Computer Vision 이란?	3
	CV(Computer Vision)이란 무엇인지 알아보고 image segmentation에 대해 간략히 알아봅니다.		
2	FCN: Fully Convolutional Networks for Semantic Segmenation	5
	FCN의 구조와 techniques에 대해 알아봅니다.		
3	U-net: Convolutional Networks for Biomedical Image Segmentation	16
	U-Net의 구조와 techniques에 대해 알아봅니다.		

U-Net: Convolutional Networks for Biomedical Image Segmentation

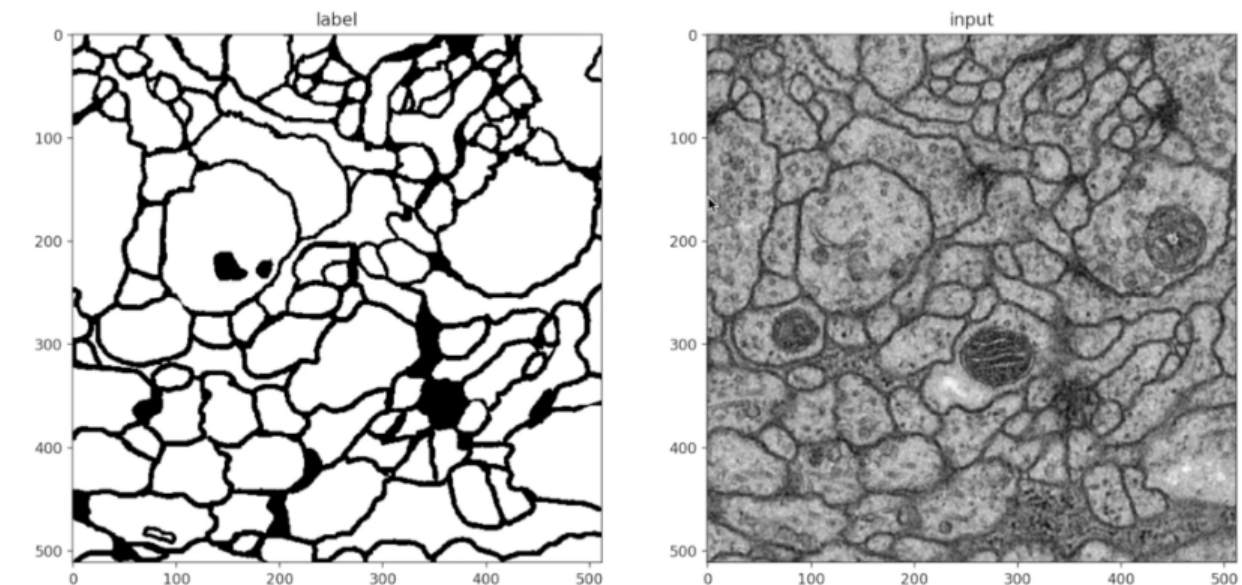
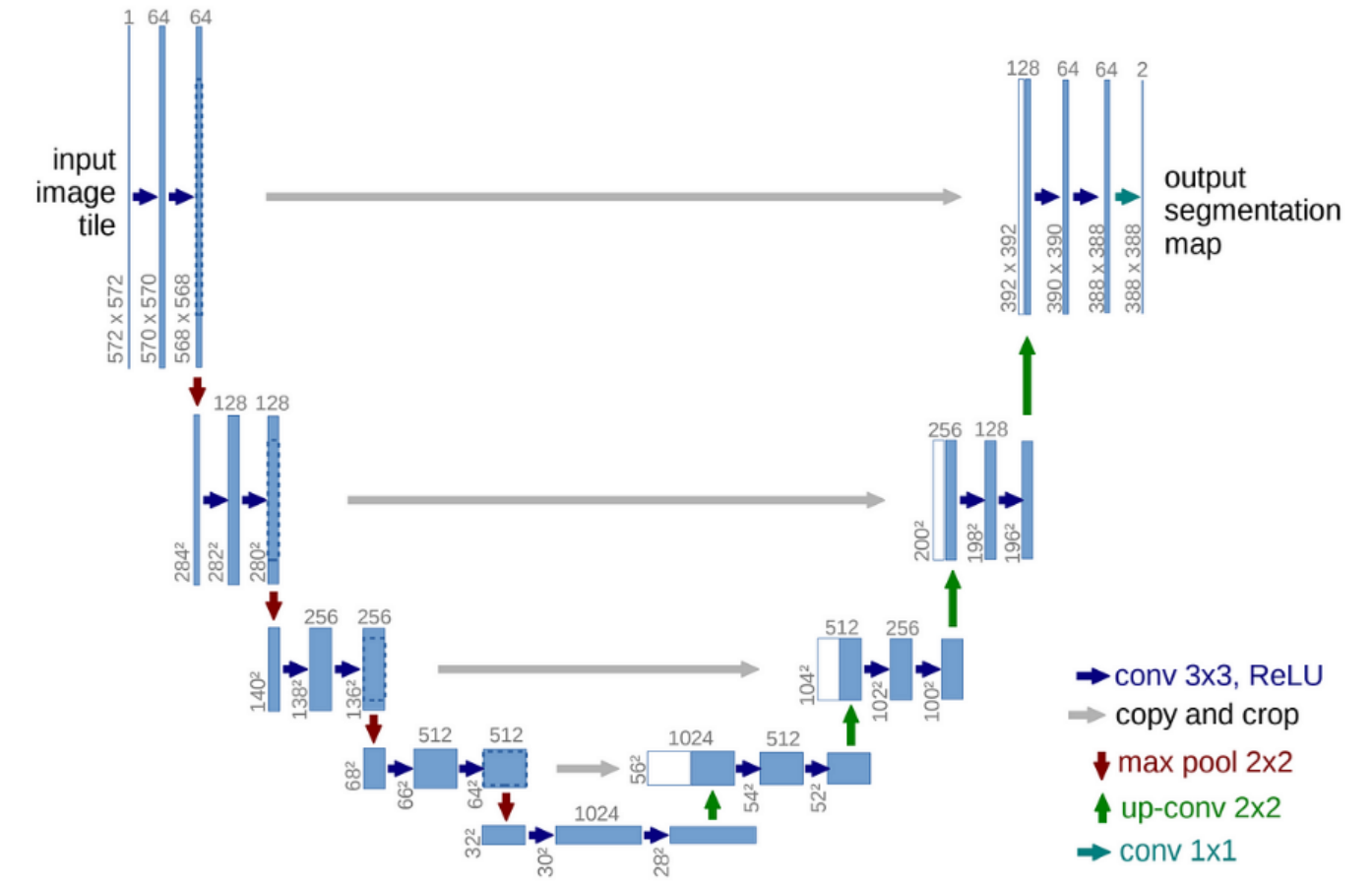
Abstract

U-Net: Convolutional Networks for Biomedical Image Segmentation

Olaf Ronneberger, Philipp Fischer, and Thomas Brox

Computer Science Department and BIOS Centre for Biological Signalling Studies,
University of Freiburg, Germany
ronneber@informatik.uni-freiburg.de,
WWW home page: <http://lmb.informatik.uni-freiburg.de/>

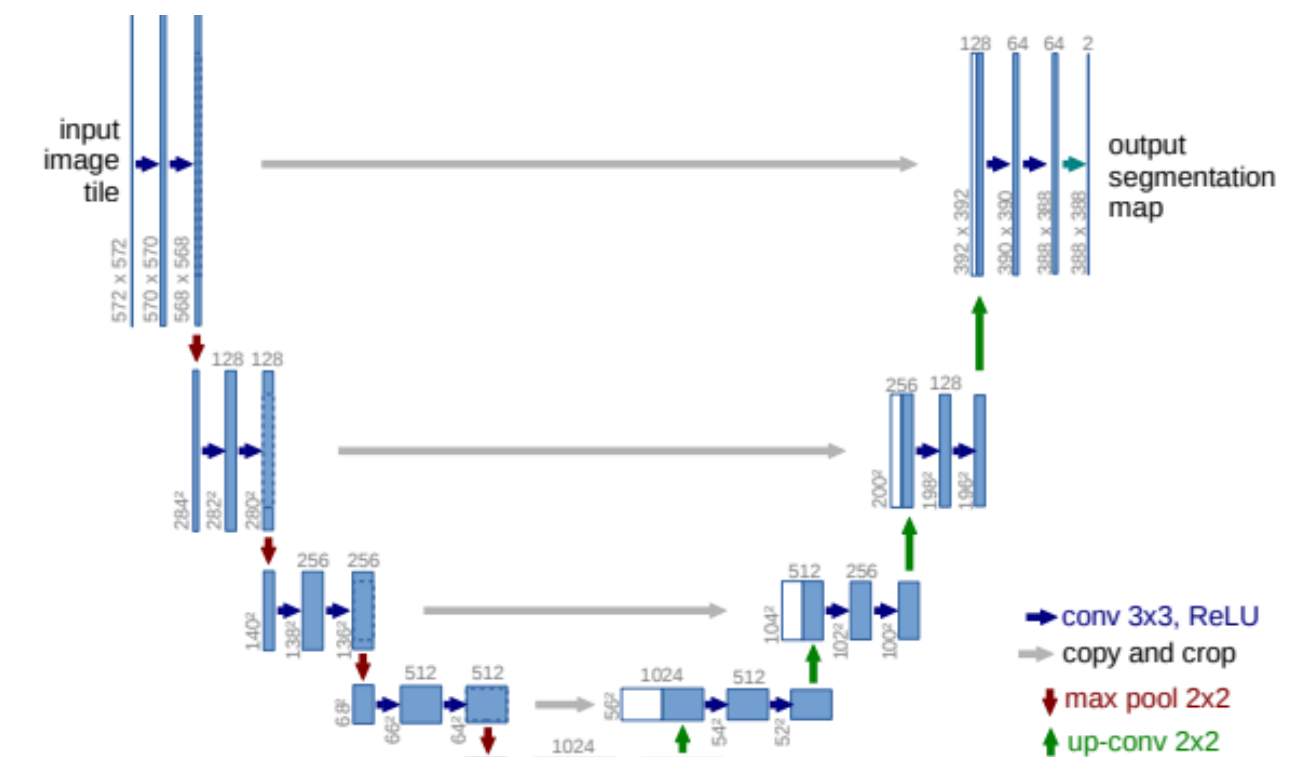
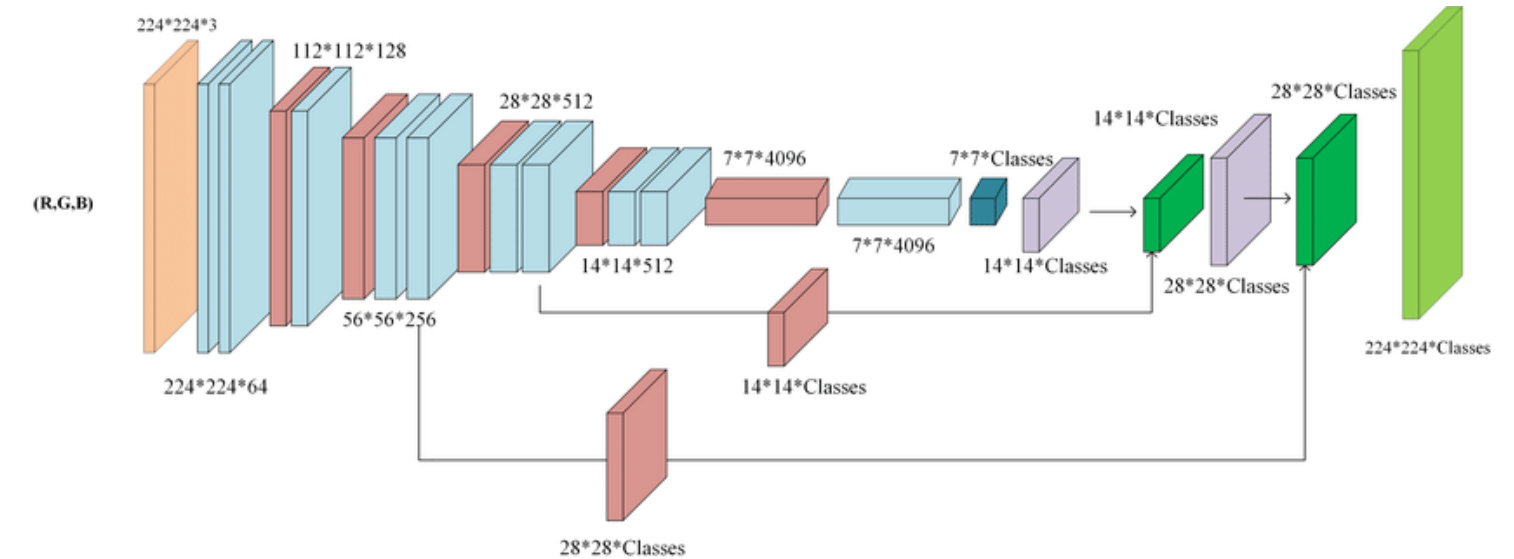
- 우리는 강력한 data augmentation을 활용한 모델을 소개합니다.
- 구조는 context를 capture하는 contracting path(수축 경로)와 정확한 localization이 가능한 대칭적인 expanding path(확장 경로)로 이루어져 있습니다.
- 다양한 biomedical image dataset(주로 현미경으로 얻은 이미지, 개수가 매우 적다는 것이 특징임)에서 end-to-end 방식으로 이전 방법을 능가하는 U-net을 소개합니다.



U-Net: Convolutional Networks for Biomedical Image Segmentation

1. Introduction

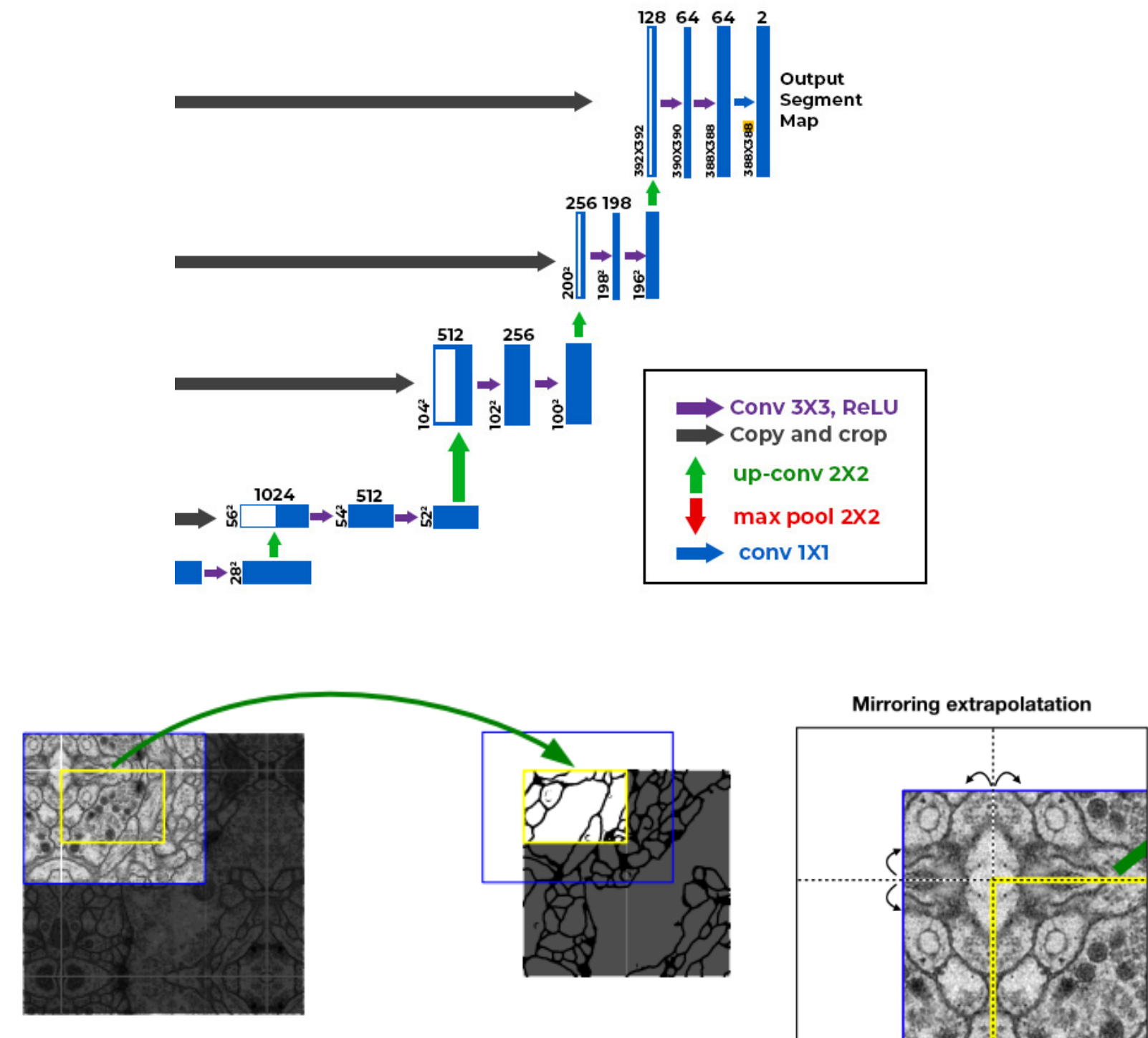
- 우리는 FCN을 확장하고 수정하여 매우 작은 개수의 training images와 더 정확한 segmentation이 가능하게 하는 구조를 만들었습니다.
- Main idea는
 - expanding path에 연속적인 layers를 보충하여 output의 해상도를 증가시킬 때 정확도를 높혔고
 - skip architecture처럼 high resolution features를 upsampled outputs과 결합했습니다.



U-Net: Convolutional Networks for Biomedical Image Segmentation

1. Introduction

- 또한 upsampling 과정에서 우리의 중요한 수정은 upsampling part도 많은 수의 channels를 가졌다는 점인데 결과적으로 contracting path와 expanding path가 대칭을 이루게 되어 u-shape의 구조를 갖게 되었습니다.
- Padding 없이 convolution 연산을 수행하기 때문에 초기 이미지에 비해 segmentation이 작게 나오는 결과가 발생합니다. 따라서 segmentation을 원하는 input image보다 더 큰 입력을 필요로 하고 이는 overlap-title strategy 해결하였습니다.
- 노란색 영역을 segmentation하기 위해 파란색 크기의 input을 받아야 하는데 이미지의 비어있는 부분을 반대방향의 data로부터 예측하는 mirroring extrapolation을 사용하였습니다. 가능한 이유는?



U-Net: Convolutional Networks for Biomedical Image Segmentation

1. Introduction

- 또한 해결해야 하는 tasks가 매우 작은 training data를 가지고 있기 때문에 data augmentation을 수행해야 했는데 training images에 elastic deformation을 수행하였습니다.

+ Data augmentation

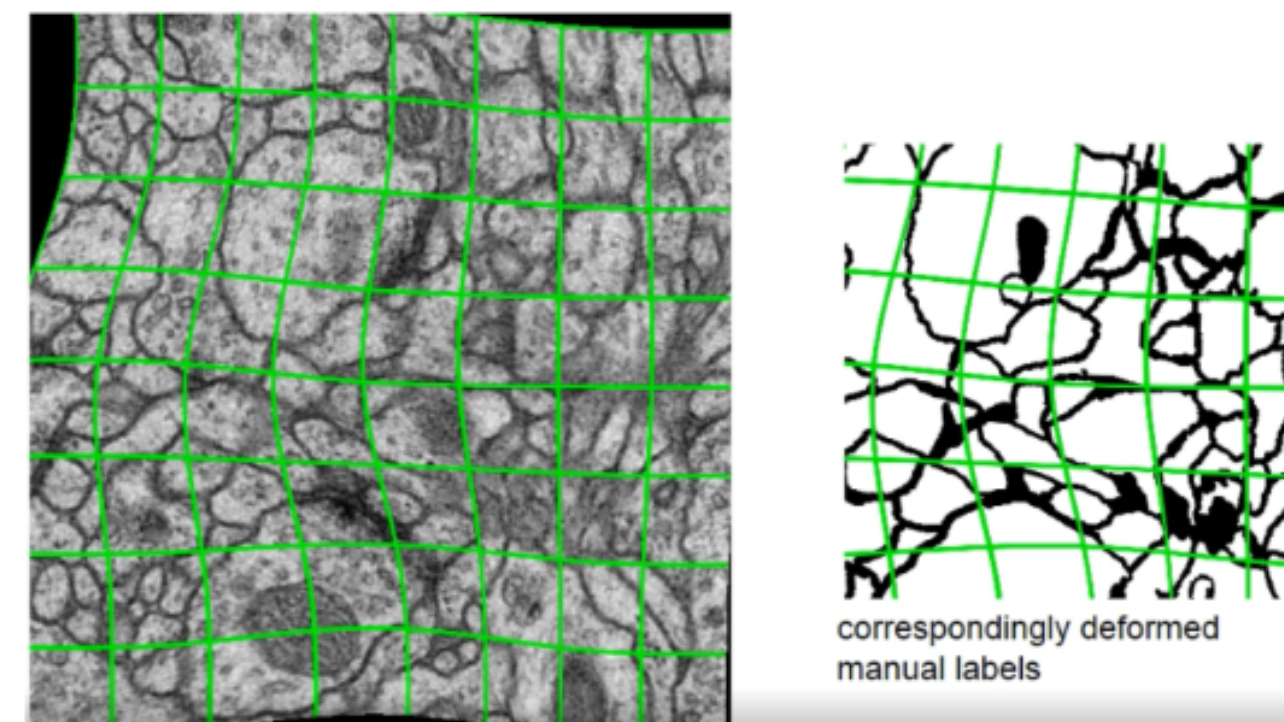
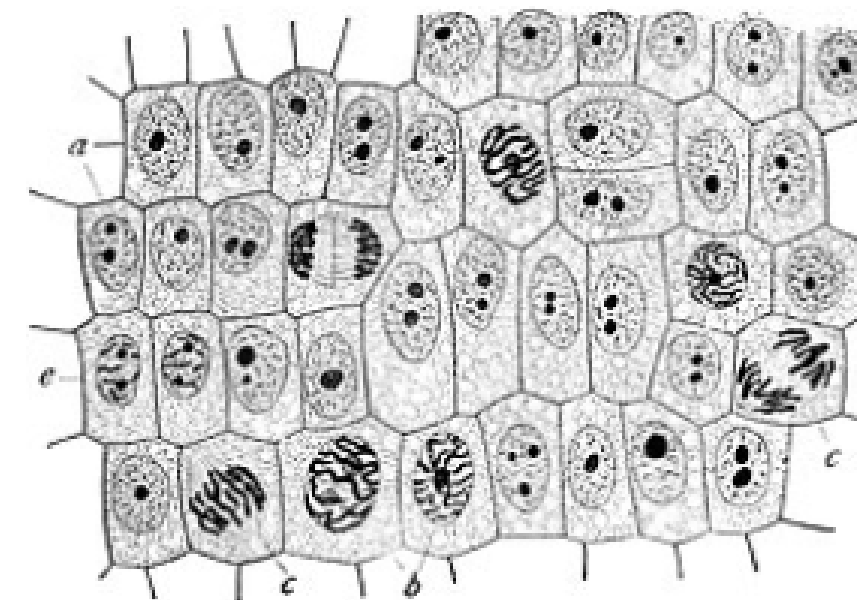
shift: 이미지를 일정 거리만큼 옮김

rotate: 이미지를 일정 각도로 회전시킴

gray value: 컬러 이미지를 그레이 스케일 이미지로 변환함.

elastic deformation: 이미지에 무작위 변형을 시킴 → 의료 영상 분야에서 유용하게 쓰임

- 이를 통해 deformations에 대한 invariance를 학습하게 되었습니다.



U-Net: Convolutional Networks for Biomedical Image Segmentation

1. Introduction

- 또 다른 tasks에선 touching objects of the same class를 분류해야 했는데(instance classification) 이는 touching cells를 가르는 background labels의 loss에 weight를 부여하는 방식으로 해결했습니다.
- 이렇게 결과적으로 만들어진 U-Net은 biomedical image segmentation에서 좋은 성능을 보이고 있습니다.
- + pathwise training(biomedical image는 크기가 다양한데 대체로 큰편)

정리해보면, U-Net은

FCN을 기반으로 upsampling 부분에 layers와 channels를 보충해 대칭적인 u-shaped model을 만들었고

경계 이미지와 patchwise segmenation을 잘 수행하기 위해 mirroring extrapolation을 이용하는 overlap-tile strategy를 사용하고

부족한 데이터를 늘리기 위해 elastic deformation을 통한 data augmentation을 수행하여 biomedical image segmentation에서 좋은 성능을 보였다.

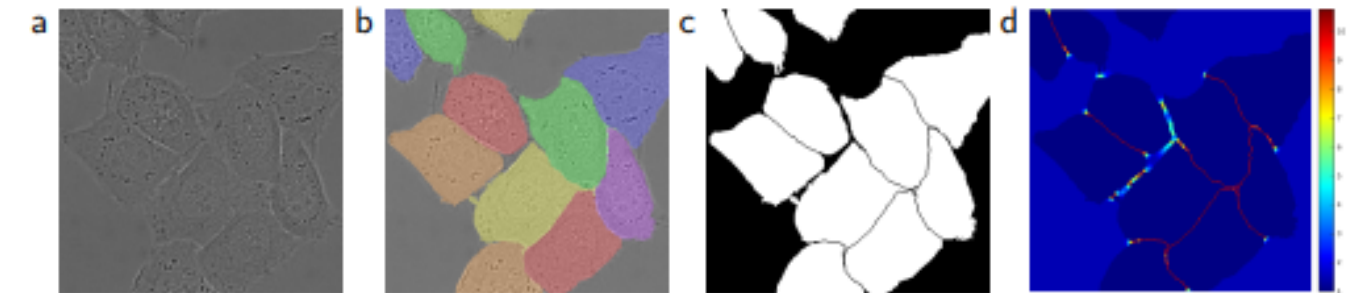
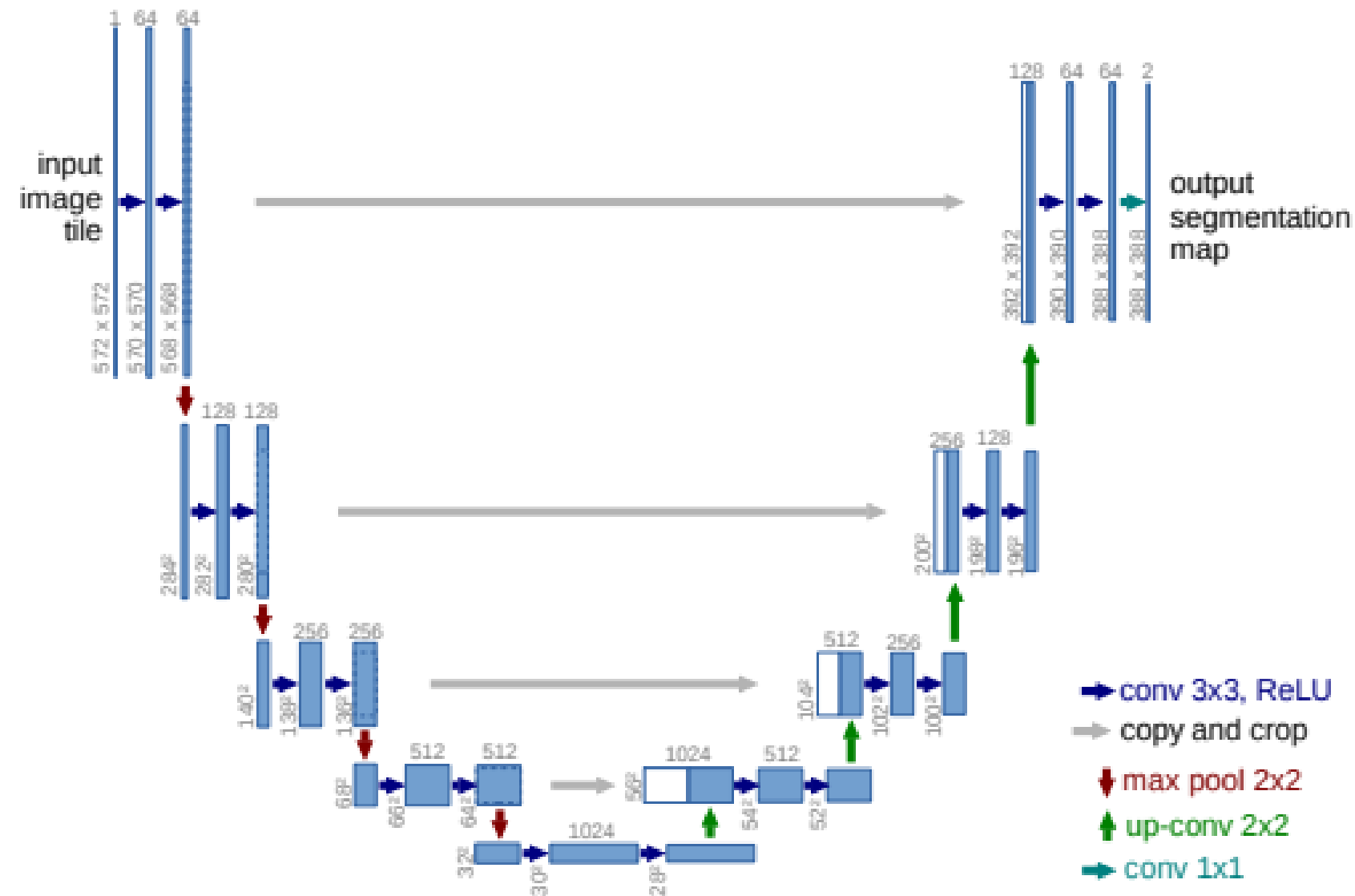


Fig. 3. HeLa cells on glass recorded with DIC (differential interference contrast) microscopy. (a) raw image. (b) overlay with ground truth segmentation. Different colors indicate different instances of the HeLa cells. (c) generated segmentation mask (white: foreground, black: background). (d) map with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



U-Net: Convolutional Networks for Biomedical Image Segmentation

2. Network Architecture



- input tile size가 짝수여야지 원활하게 2*2 max pooling이 진행됨.
- 마지막 1*1 convolution이 64-component feature vector를 class에 맞게 mapping 해준다.

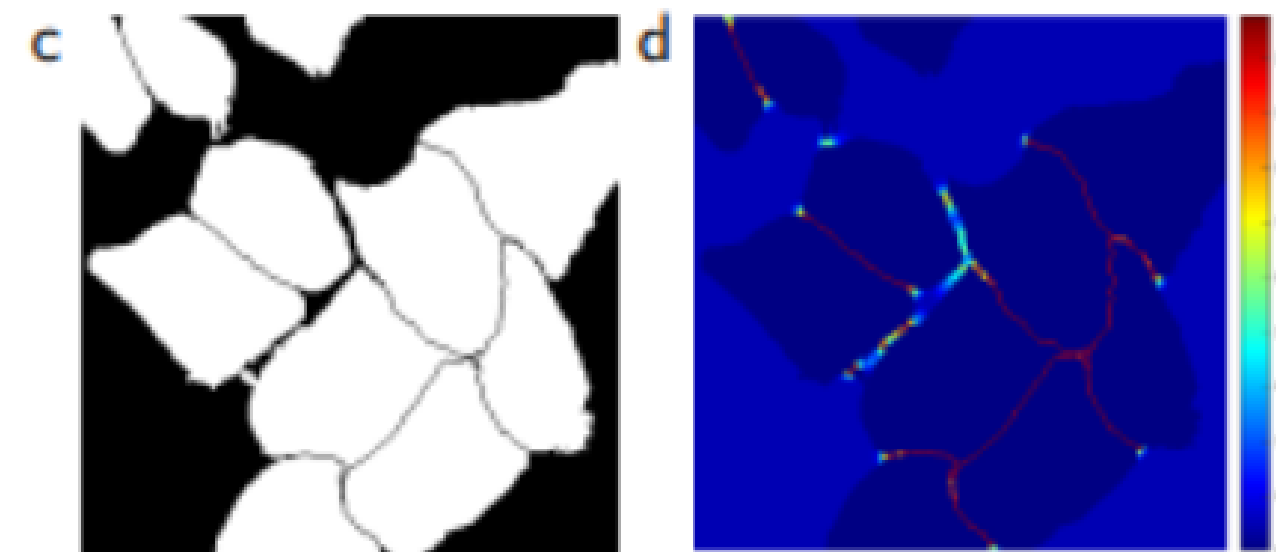
U-Net: Convolutional Networks for Biomedical Image Segmentation

3. Training

- overhead를 최소화 하고 GPU memory의 사용을 최대화 하기 위해
input tiles은 크게하고 batch size를 줄여야 해 batch size를 1로 잡음
- momentum을 0.99인 큰 값으로 잡아 current optimization step에 이전 training samples의 대부분이 반영되도록 하였음(이전 단계의 gradient를 현재 단계에 반영하는 방식)
- 손실함수는 soft-max와 cross-entropy를 결합한 방식
- weight map을 사전에 계산해두는데 touching cells의 경계에 빨간색의 더 높은 가중치를 부여함.(frequencies를 고려한 느낌)
- weight는 class의 frequency를 고려한 weight에 가까운 cell의 경계까지의 거리 2개에 반비례하도록(즉, 경계에 있는 셀일 수록 높은 가중치 부여) 설계하였다.

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

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



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

U-Net: Convolutional Networks for Biomedical Image Segmentation

4. Experiments

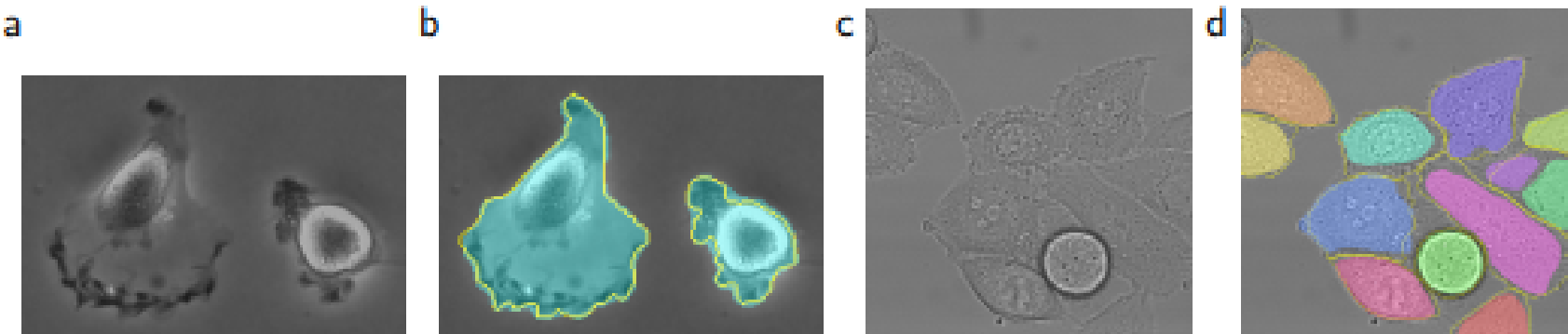
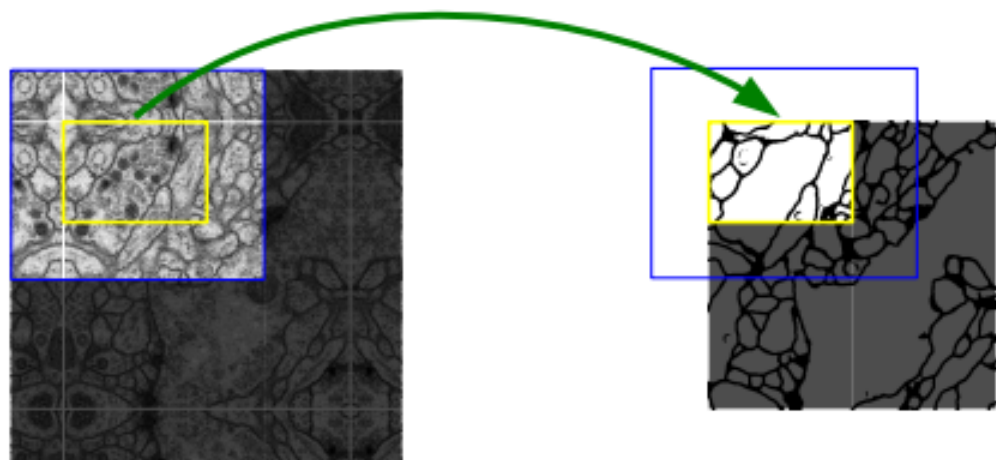


Table 1. Ranking on the EM segmentation challenge [14] (march 6th, 2015), sorted by warping error.

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

U-Net: Convolutional Networks for Biomedical Image Segmentation

5. Conclusion

- U-Net은 다양한 biomedical segmentation applications에서 좋은 performance를 보여준다.
- 사용한 주요 techniques
 - u-shaped architecture(layers, channels)
 - overlap-tile strategy & mirroring extrapolation
 - data augmentation(by. elastic deformation)

FCN: Fully Convolutional Networks for Semantic Segmentation

문제

- FCN의 1*1 convolutional layers의 마지막 층의 kernel의 갯수가 24개였다.
- 이는 데이터의 클래스가 몇개로 label됨을 의미하는가?(background 제외)

$$y_{ij} = f_{ks}(\{\mathbf{x}_{si+\delta_i, sj+\delta_j}\}_{0 \leq \delta_i, \delta_j \leq k-1})$$

- 합성곱층에서 following layer의 data vector는 위와 같이 구해진다.(k는 kernel의 사이즈, s는 stride의 크기, f는 층의 종류에 따른 함수)
- 2개의 합성곱층을 지나가는 것을 하나의 합성곱층을 지나는 것처럼 위 함수를 합성하고 싶을 때 빈칸에 들어갈 식은 어떻게 되는가?

$$f_{ks} \circ g_{k's'} = (f \circ g)_{\quad, \quad}$$

U-Net: Convolutional Networks for Biomedical Image Segmentation

문제

- U-net은 zero-padding 대신 mirroring extrapolation을 사용한다. 이 방식을 사용해도 되는 이유는?

- 아래 문장들의 T/F를 구분하시오

1. U-Net의 구조(21 slide 참고)는 총 23개의 convolutional layers가 존재한다. (T/F)
2. U-Net은 whole image training을 사용한다. (T/F)
3. U-Net에서는 cropping이 필수이다. (T/F)