

U-Net: Convolutional Networks for Biomedical Image Segmentation

우다연

dyeonwu@gmail.com

Computer Vision

2023/03/14



Contents

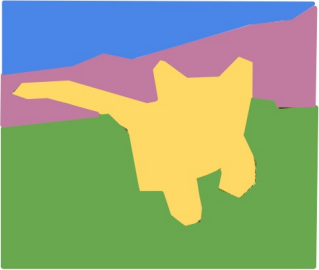
- Introduction
- Architecture of U-Net
- Strategy
- Training
- Experiments

Introduction

Segmentation

- 이미지를 하나의 클래스로 분류하는 일반적인 classification과 달리 이미지의 **각 픽셀마다** 클래스 분류


Semantic Segmentation



GRASS, CAT, TREE, SKY

No objects, just pixels


Instance Segmentation



DOG, DOG, CAT

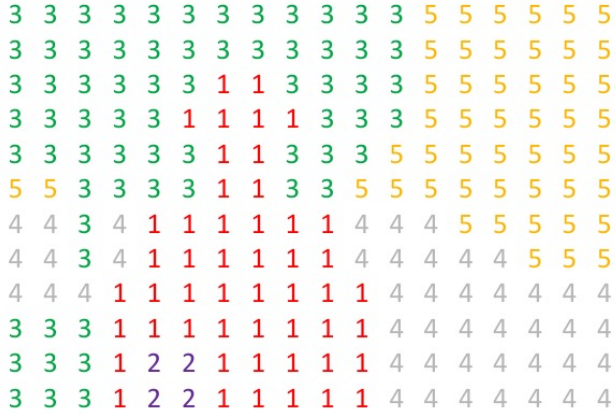
Object

This image is CC0 public domain



Input

segmented →



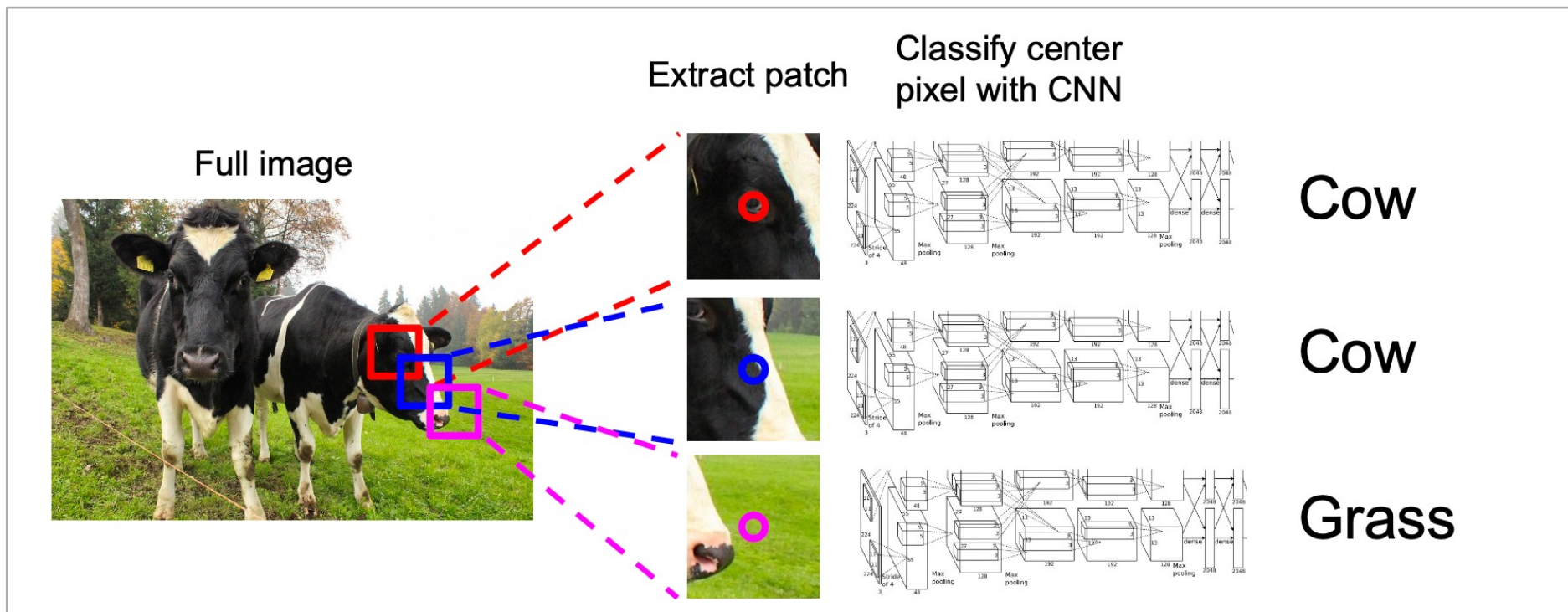
1: Person
2: Purse
3: Plants/Grass
4: Sidewalk
5: Building/Structures

Semantic Labels

Introduction

Sliding window

- Segmentation을 수행하기 위한 다양한 방법 등장, 대표적인 것이 Sliding window

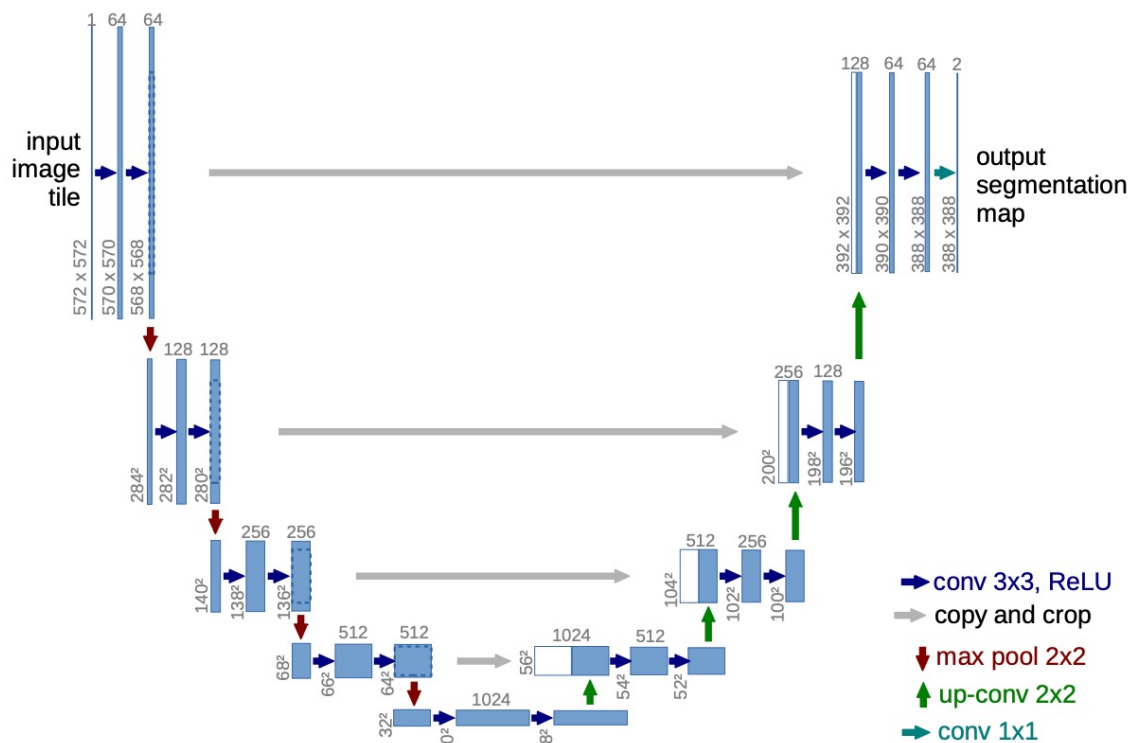


문제: 계산 비용, 시간 ↑, patch 간 중복 영역으로 인한 비효율성 등

Architecture of U-Net

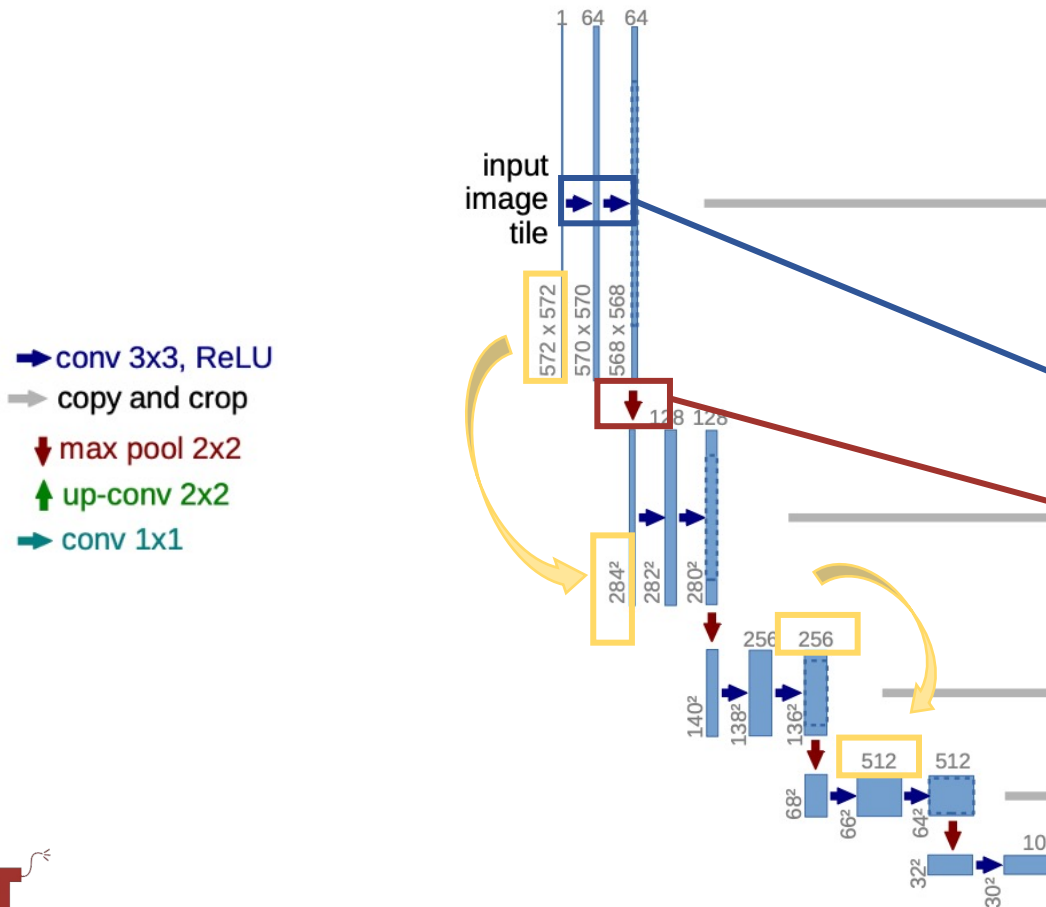
U-Net

- Fully Convolutional network 기반 모델
- 적은 양의 training images 활용,
고도화된 segmentation 수행
- Contracting path & Expanding path 대칭적 형태



Architecture of U-Net

Contracting path

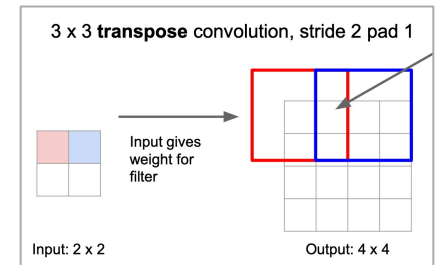
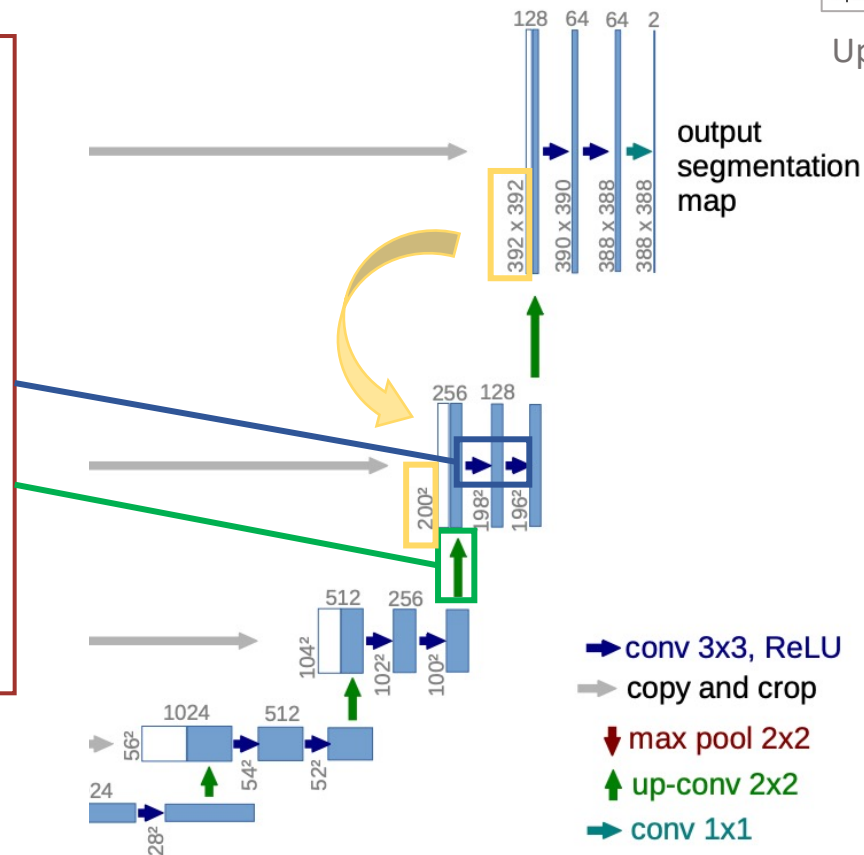


- Downsampling
- 입력 이미지의 context 추출 (주요 특징 추출)
- 각 step마다 3x3 conv 2번 수행 후 ReLU 연산
- 2x2 max-pooling (stride 2)를 통해 이미지 크기 줄임
- 채널의 크기 2배씩 증가

Architecture of U-Net

Expanding path

- Upsampling
- **정확한 localization 수행 (어디에 위치하는지)**
- 각 step마다 3x3 conv 2번 수행 후 ReLU 연산
- 각 step마다 2x2 up-conv 연산을 하여 이미지 확대
- 채널의 크기 2배씩 감소

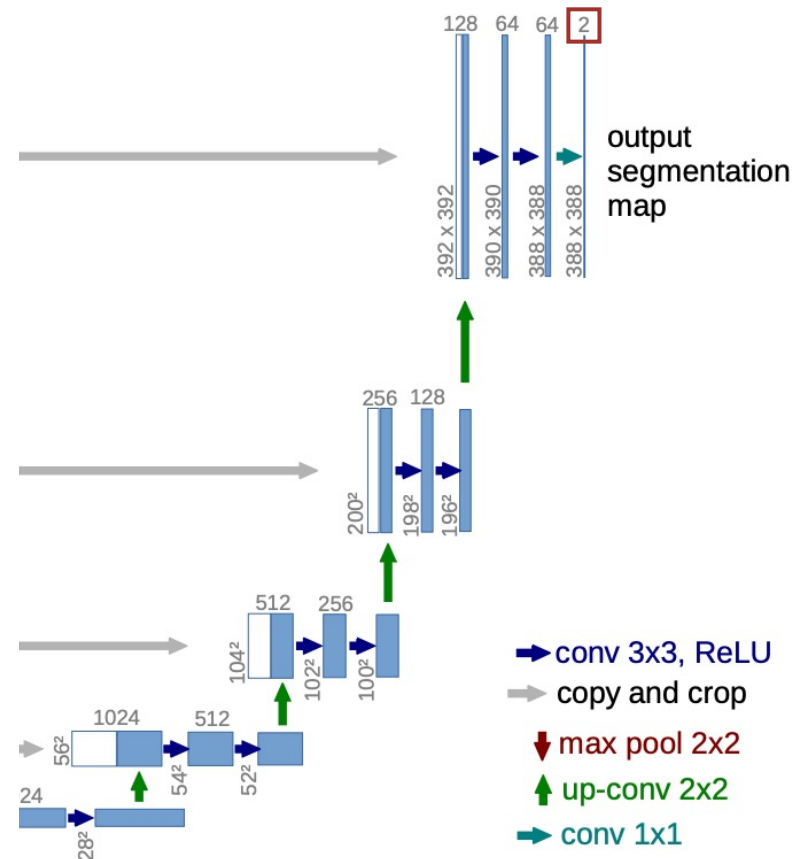


Up-conv (transposed conv)

Architecture of U-Net

Expanding path

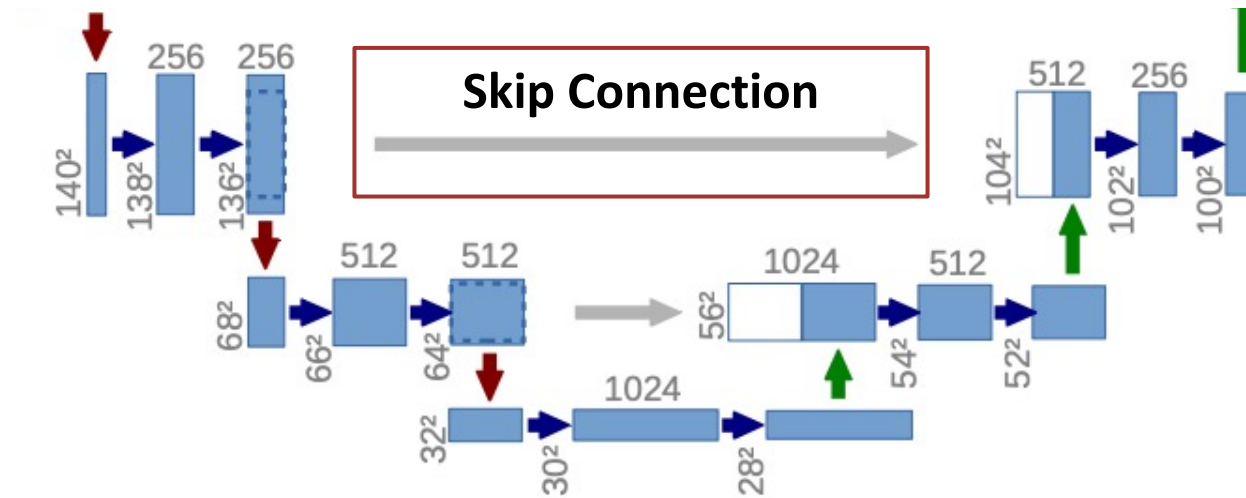
- fc-layer 없음
- 1x1 conv 연산을 통해 배경과 배경 아닌 것 구분



Architecture of U-Net

Skip connection

- 수축 경로와 확장 경로의 feature map을 합치는 일 (**Concatenation**)

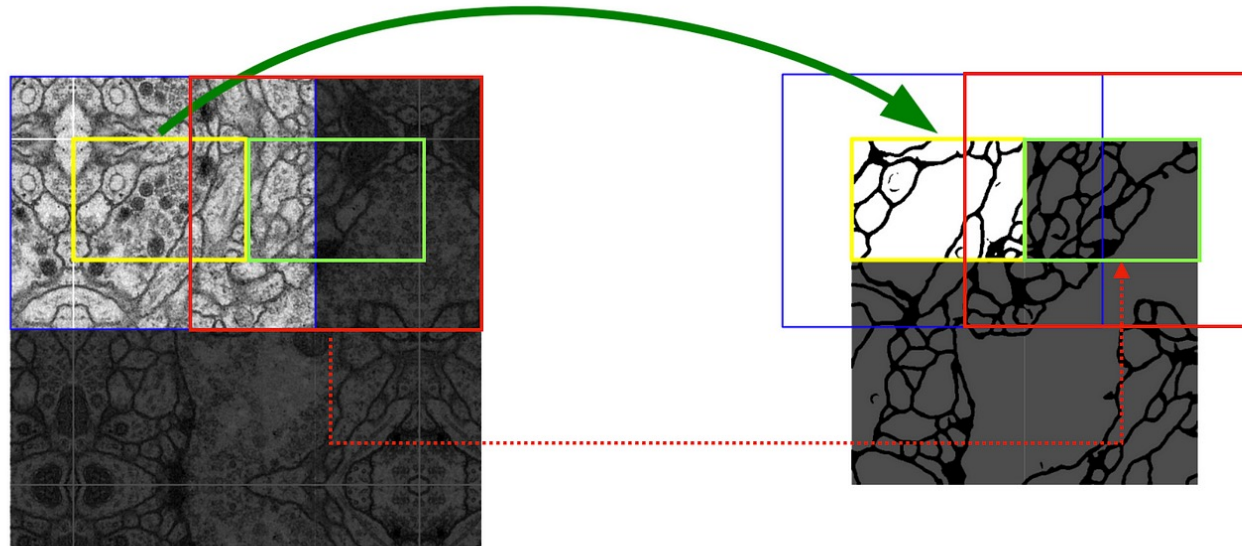


- 저차원 & 고차원 특징을 모두 활용함으로써 이미지 정보 추출 & 정확한 위치 정보 파악 가능
- 수축 경로와 확장 경로의 feature map 사이즈를 맞춰주기 위해 **crop** 수행

Strategy

Overlap-tile

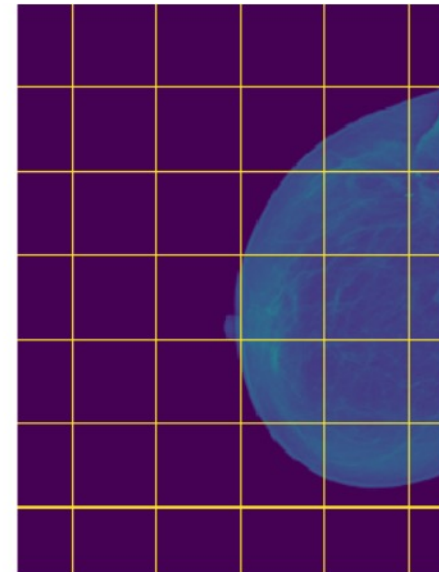
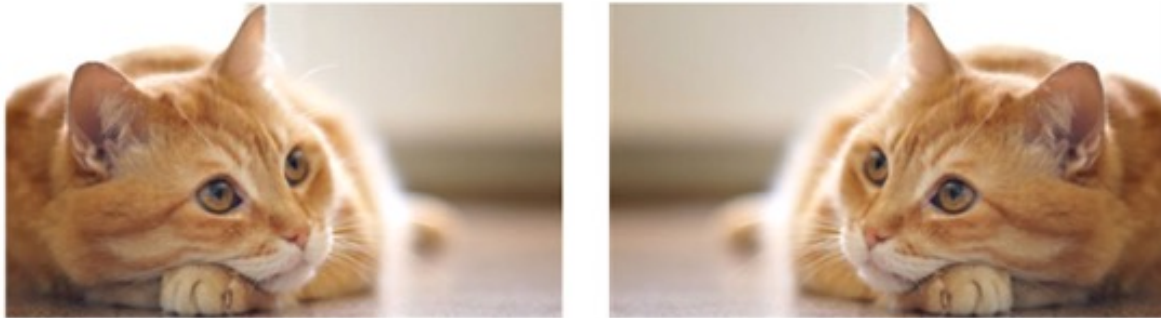
- U-Net architecture에서 input 대비 output의 이미지 해상도가 작아진다는 점 고려 (padding X)
- 이미지 일부분을 크롭하고 이보다 넓은 범위를 경계로 설정
- 이미지가 존재하지 않는 부분은 미러링 기법 수행 (extrapolation)



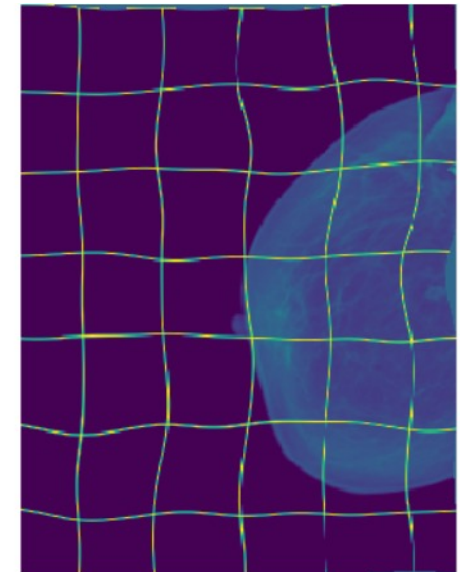
Strategy

Data augmentation

- Training image가 적을 때 다양성을 높이고 모델을 견고하게 만들 수 있도록 함
- Grid에 비선형성을 더하는 **Elastic deformation** 방법 사용



(a) Original

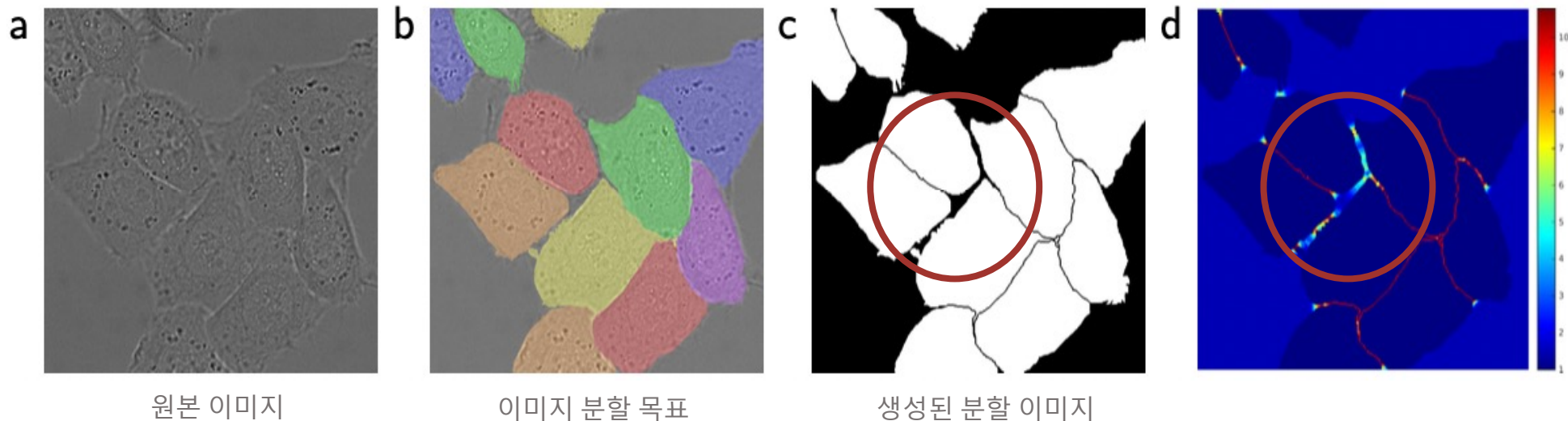


(b) Deformed

Training

Touching cells sepeation

- 동일 클래스의 접촉 개체 분할 작업 중요



Training

Touching cells sepeation

- 가중치 맵 $w(x)$ 사전 계산
- 네트워크가 작은 분리 경계 학습하도록 함
- 각 픽셀마다 클래스 분포가 다른 것 보완
- 인접한 세포 사이 배경 레이블에 높은 가중치 부여
- 세포까지의 거리값 작을수록 가중치 높아짐

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

- w_c : 클래스 출현 빈도의 균형을 맞추기 위한 가중치 맵
- d_1 : 픽셀 x 에서 가장 가까운 세포의 경계까지의 거리
- d_2 : 픽셀 x 에서 두 번째로 가까운 세포의 경계까지의 거리

Training

학습 방법

- 픽셀 단위의 **softmax** 함수 사용하여 예측

$$p_k(x) = \frac{\exp(a_k(x))}{\sum_{k'=1}^K \exp(a_{k'}(x))}$$

픽셀 x 가 클래스 k 에 속할 확률

- x : 픽셀의 위치
- k : k 번째 특징 채널 (클래스)
- K : 클래스의 전체 개수
- $a_k(x)$: k 번째 채널의 x 위치의 activation 값

- Loss function: **Cross-Entropy**

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

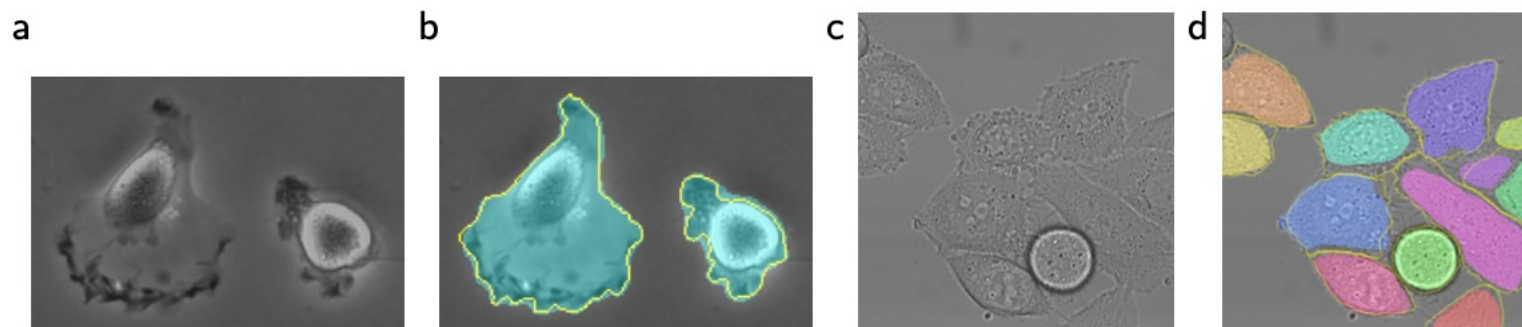
- $\ell(x)$: 픽셀 x 의 true label (정답 값)

Experiments

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

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

* IOU 수치



* 노란색 테두리: ground-truth



TRAIN AND TEST