

Week9_예습과제_김도희

U-Net: Convolutional Networks for Biomedical Image Segmentation



data augmentation을 이용한 U-Net의 main structure와 training strategy 제안. U-Net은 대칭적인 U자형 구조로 설계되어 있으며, contracting path와 expanding path로 구성

1. Introduction

Biomedical image processing에서는 **localization**이 포함되어야만 한다. 이러한 이유로 sliding window set up이 제시되었다. 이 방식은 localize가 가능하고 patch로의 training data의 수가 training image이 수보다 훨씬 더 많다는 특징이 있다. 하지만 기존의 sliding window 방식은 느리며, 중복이 많고 정확한 localize와 맥락 활용 사이의 균형을 맞추기 어렵다.

따라서, 이 논문에서는 **U-Net**을 제안. 기존의 Fully Convolutional Network를 기반으로 설계되었으며, 적은 수의 학습 이미지로도 효율적인 분할을 가능하게 한다.

Main Idea : 연속되는 layer로 축소되는 네트워크를 보충하는 방식, upsampling하는 부분

결론적으로는 적은 수의 훈련 이미지로도 효과적인 세분화가 가능하며, 다양한 생의학 이미지 처리 문제에 적용할 수 있음을 입증하였다. 새로운 네트워크 아키텍처와 훈련 전략을 통해 생의학 이미지 세분화에서 높은 성능을 달성할 수 있음을 보여주었다.

2. Network Architecture

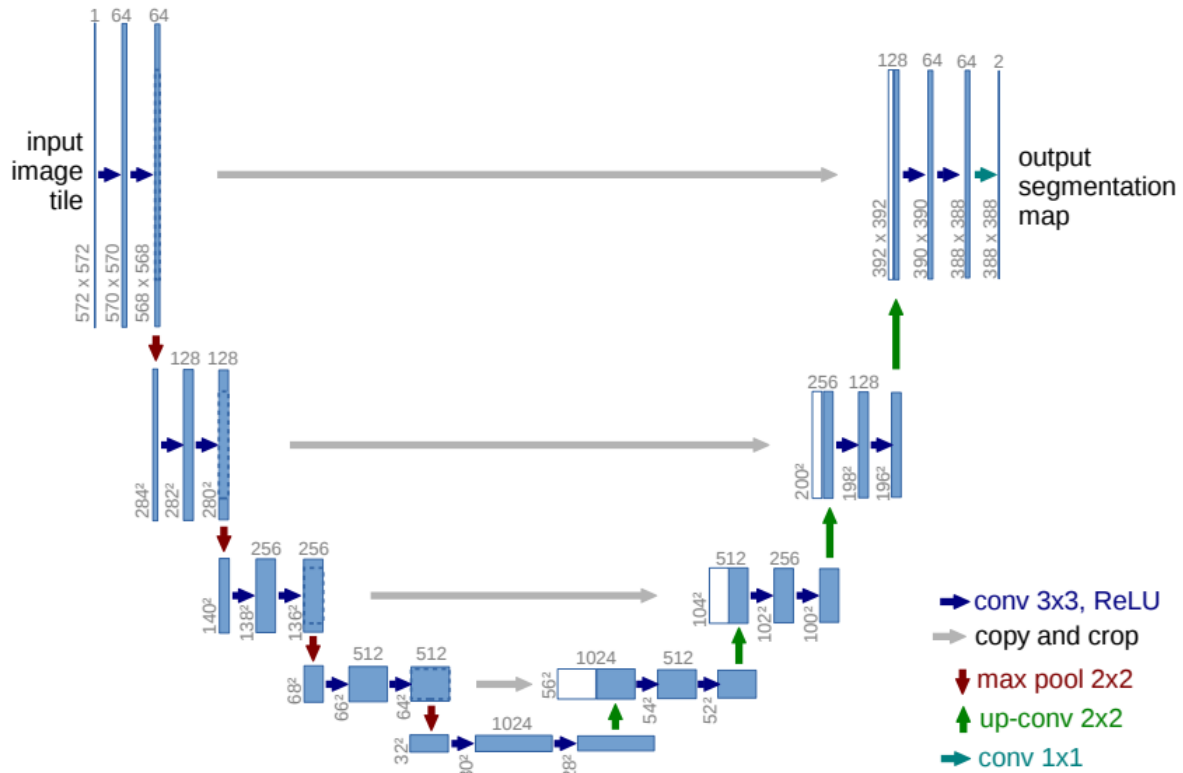


Fig. 1. U-net architecture (example for 32x32 pixels in the lowest resolution). Each blue box corresponds to a multi-channel feature map. The number of channels is denoted on top of the box. The x-y-size is provided at the lower left edge of the box. White boxes represent copied feature maps. The arrows denote the different operations.

- Contracting path(입력의 공간적 컨텍스트를 캡처): 반복되는 2개의 3x3 convolution으로 구성되고 난 후에는 ReLU와 down sampling을 위한 2 striding 2x2 max pooling operation, down sampling시 num of channel 2배
- Expansive path(정확한 현지화를 제공): upsampling과 함께 고해상도 정보를 결합. 대칭적인 U자형 설계는 경계 픽셀의 손실을 최소화

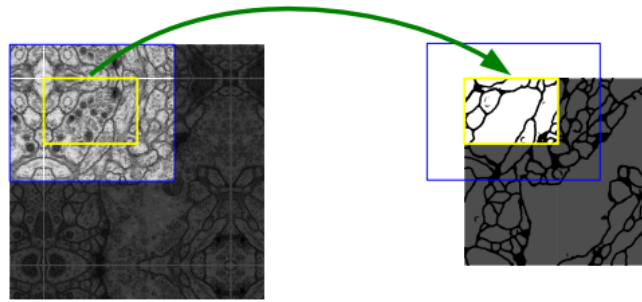


Fig. 2. Overlap-tile strategy for seamless segmentation of arbitrary large images (here segmentation of neuronal structures in EM stacks). Prediction of the segmentation in the yellow area, requires image data within the blue area as input. Missing input data is extrapolated by mirroring

3. Training

- Stochastic Gradient Descent 사용하여 Caffe를 통해 구현됨
- unpadded convolution을 사용하기 때문에 출력 이미지의 크기는 입력 이미지보다 항상 작은 고정된 경계 너비
- 큰 배치 크기보다는 큰 입력 타일을 선호하며, 배치를 단일 이미지로 함.
- 높은 모멘텀 0.99 를 사용
- 에너지함수
 - 최종 특징 맵 위에 픽셀 단위의 소프트맥스와 크로스 엔트로피 손실 함수로 계산

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

- $a_k(x)$: 픽셀 위치 x 에서 k 번째 특징 채널의 활성화 값
- K : 클래스의 수
- $p_k(x)$: 최대 함수의 근사치
- 손실함수

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

가중치 맵은 ground truth 분할에 대해 사전 계산되며, 특정 클래스의 픽셀이 학습 데이터에서 차지하는 빈도를 보상하고, 접촉하는 세포들 사이의 작은 경계들을 네트워크가 학습하도록 한다.

네트워크는 여러 합성곱 층과 경로를 가지므로, 가중치의 초기화가 매우 중요. 그렇지 않으면 네트워크의 일부가 과도한 활성화를 하거나, 다른 부분이 전혀 기여하지 않게 될 수 있다. 초기 가중치는 각 특징 맵이 대략 단위 분산을 가지도록 조정해야 한다. 우리 네트워크에서는 가중치를 가우시안 분포에서 표준편차 $\sqrt{2/N}$ 로 초기화한다. 여기서 N 은 한 뉴런의 입력 노드 수.

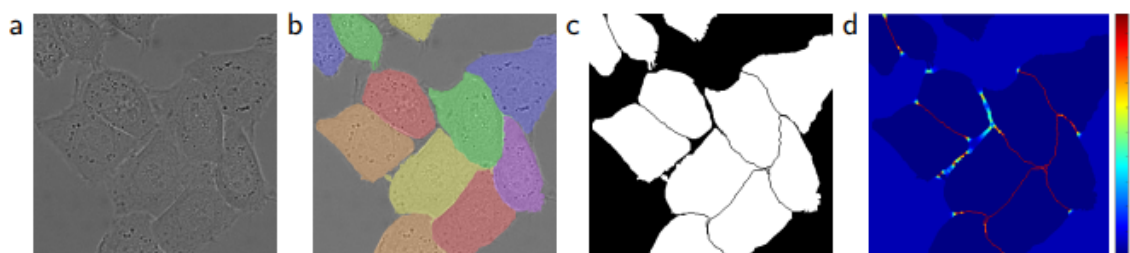


Fig. 3. HeLa cells on glass recorded with DIC (differential interference contrast) microscopy. (a) raw image. (b) overlay with ground truth segmentation. Different colors indicate different instances of the HeLa cells. (c) generated segmentation mask (white: foreground, black: background). (d) map with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

3.1 Data Augmentation

데이터 증강은 학습 샘플이 적을 때 네트워크가 원하는 invariance과 robustness 특성을 학습하도록 하는 데 필수적

- **데이터 증강의 필요성:** 현미경 이미지에서 이동, 회전 불변성 및 변형에 대한 robustness 을 학습하기 위함.
- **랜덤 탄성 변형:** 주요 증강 기법으로 랜덤 탄성 변형을 사용하여 적은 수의 주석 이미지로도 효과적인 학습이 가능. 이 변형은 3x3 격자에서 랜덤 변위를 생성하고, 3-deg intetpolation을 통해 부드럽게 변형합니다.
- **드롭아웃 사용:** 수축 경로의 끝에서 드롭아웃 층을 사용해 추가적인 증강 효과를 부여

4. Experiments

- 해당 연구는 U-Net을 사용하여 세 가지 다른 세분화 작업에 적용한 결과를 설명하고 있음

1. 전자 현미경 이미지에서의 신경 구조 세분화

- Drosophila의 뇌 신경 연구를 위해 제공된 EM 세분화 챌린지 데이터셋을 활용
 - 각 이미지는 완전히 주석이 달린 세포와 막의 분할 맵을 포함
- U-Net은 다양한 회전된 입력 데이터의 평균을 사용하여 선행 또는 후행 처리 없이도 매우 낮은 warping error와 rand error를 달성하였음
 - 이전에 사용된 Ciresan 슬라이딩 윈도우 컨볼루션 네트워크보다 훨씬 우수한 성능을 보였음

2. 광학 현미경 이미지에서의 세포 세분화

: ISBI 세포 추적 챌린지

- **PhC-U373** 데이터셋은 상영 공간 현미경을 사용하여 기록된 Glioblastoma-astrocytoma U373 세포들을 포함하고 있음
- **DIC-HeLa** 데이터셋은 차분 간섭 현미경을 사용하여 기록된 HeLa 세포들을 포함하고 있음
- 각각의 데이터셋에서 U-Net은 높은 IOU(Intersection Over Union)를 달성하여 다른 알고리즘들보다 우수한 성능을 보였음

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

- 3. **성능 평가:** U-Net은 높은 정확도로 세포 분할을 수행하며, 데이터셋 특화된 처리 없이도 탁월한 성과

5. Conclusion

- U-Net 구조는 매우 다양한 생명 과학 세분화 응용 프로그램에서 매우 좋은 성능을 발휘함
 - 탄성 변형을 이용한 데이터 증강 덕분에 매우 적은 수의 주석이 달린 이미지만으로도 높은 성과를 얻을 수 있으며, NVidia Titan GPU(6GB)에서 단 10시간의 합리적인 훈련 시간을 보임
- Caffe 기반의 전체 구현과 훈련된 네트워크를 제공
- U-Net 구조가 다양한 과제에 쉽게 적용될 수 있을 것이라 기대됨