



U-Net: Convolutional Networks for Biomedical Image Segmentation



Data Augmentation을 위한 효율적인 네트워크를 제시하자!

- 문맥을 캡처하기 위한 수축 경로와 정밀한 위치 지정을 가능하게 하는 대칭 확장 경로로 구성

1. Introduction

- 해당 논문은 딥 컨볼루션 네트워크의 성공적인 훈련에 많은 주석이 달린 훈련 샘플이 필요하다는 일반적인 합의에 도전
 - 데이터 증강을 강력하게 활용하여 적은 수의 주석이 달린 샘플로도 효율적인 훈련이 가능한 네트워크와 훈련 전략을 제시

주요 내용

1. 기존 연구 및 한계

- 컨볼루션 네트워크는 대규모 데이터셋에서 우수한 성능을 보였지만, 많은 훈련 데이터가 필요한 단점이 존재
- Ciresan 등의 네트워크는 슬라이딩 윈도우 방식을 사용하여 위치를 지정할 수 있지만, 느리고 중복이 많음

2. 제안된 네트워크

- 수축 경로와 확장 경로로 구성된 완전 컨볼루션 네트워크(Fully Convolutional Network)를 기반으로 수정 및 확장한 U자형 아키텍처를 제안
- 높은 해상도의 특징을 유지하면서 위치 정보를 정확하게 제공하는 업샘플링 기법을 사용

3. 데이터 증강

- 훈련 데이터가 적기 때문에 탄성 변형을 포함한 데이터 증강 기법을 사용하여 네트워크가 다양한 변형에 대해 불변성을 학습할 수 있도록 함

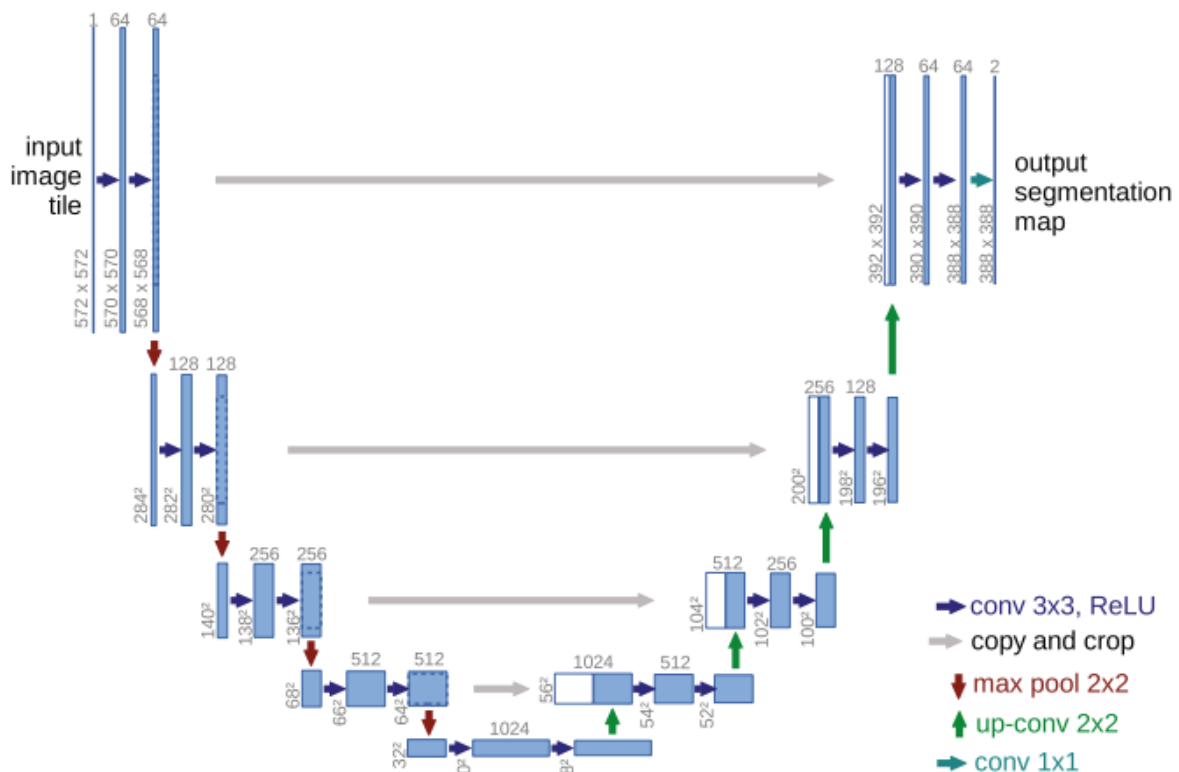
4. 성능 및 응용

- 제안된 네트워크는 ISBI 챌린지의 신경 구조 세분화 및 세포 추적에서 우수한 성능을 보였음
- 특히, ISBI 2015 세포 추적 챌린지에서 큰 차이로 우승하였음

5. 결론

- 적은 수의 훈련 이미지로도 효과적인 세분화가 가능하며, 다양한 생의학 이미지 처리 문제에 적용할 수 있음을 입증
- 새로운 네트워크 아키텍처와 훈련 전략을 통해 생의학 이미지 세분화에서 높은 성능을 달성할 수 있음을 보여줌

2. Network Architecture



• 네트워크 구조

◦ 수축 경로

- 두 개의 3x3 컨볼루션, 렐루(ReLU) 활성화 함수, 2x2 맥스 풀링 연산을 반복하여 다운샘플링을 진행
 - 각 다운샘플링 단계에서 특징 채널 수를 두 배로 증가
- 확장 경로
 - 특징 맵의 업샘플링, 2x2 컨볼루션, 수축 경로의 자른 특징 맵과의 연결, 두 개의 3x3 컨볼루션
 - 각 컨볼루션 뒤에 ReLU 활성화 함수를 적용
 - 마지막 층에서 1x1 컨볼루션을 사용하여 원하는 클래스 수로 매핑
 - 총 23개의 컨볼루션 레이어 활용
- 타일링 전략
 - 출력 세분화 맵의 원활한 타일링을 위해 모든 2x2 맥스 풀링 연산이 짝수 크기의 레이어에 적용되도록 입력 타일 크기를 선택하는 것이 중요함

3. Training

- Caffe를 통해 구현됨
 - 확률적 경사 하강법(SGD)을 활용하여 훈련됨
- GPU 메모리를 최대한 활용하고 오버헤드를 최소화하기 위해, 대량의 배치 크기보다 큰 입력 타일을 활용
 - 배치 크기를 단일 이미지로 축소함
- 매우 높은 모멘텀(= 0.99)을 활용
 - 이전에 본 훈련 샘플의 상당수가 현재 최적화 단계의 업데이트를 결정하도록
- 에너지 함수
 - pixel-wise softmax 함수로 계산됨 (+ cross entropy loss function)

$$p_k(\mathbf{x}) = \exp(\underbrace{a_k(\mathbf{x})}_{\text{activation}}) / \left(\sum_{k'=1}^{\overbrace{K}^{\text{\# of classes}}} \exp(\underbrace{a_{k'}(\mathbf{x})}_{\text{pixel position}}) \right)$$

- $p_k(\mathbf{x}) \approx 1$: maximum activation $a_k(\mathbf{x})$ 를 가질 때
 - $p_k(\mathbf{x}) \approx 0$: 나머지 경우
- 이후 cross-entropy loss는 규제를 적용(penalize)

$$E = \sum_{\mathbf{x} \in \Omega} \underbrace{w(\mathbf{x})}_{\substack{\uparrow \text{weight map}}} \log(p_{\underbrace{\ell(\mathbf{x})}_{\substack{\uparrow \text{true label}}}}(\mathbf{x}))$$

- 각 실제 세분화에 대한 가중치 맵을 사전에 계산
 - 데이터 세트에서 특정 클래스의 픽셀 빈도를 보상하고, 네트워크가 접촉하는 세포 사이에 도입한 작은 분리 경계를 학습하도록 함

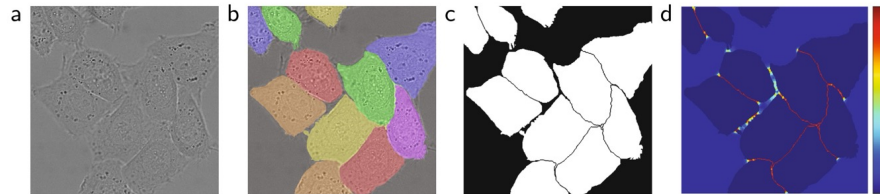


Fig. 3. HeLa cells on glass recorded with DIC (differential interference contrast) microscopy. **(a)** raw image. **(b)** overlay with ground truth segmentation. Different colors indicate different instances of the HeLa cells. **(c)** generated segmentation mask (white: foreground, black: background).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

- 결정 경계는 아래의 식을 통해 계산됨

$$\underbrace{w(\mathbf{x})}_{\substack{\uparrow \text{weight map}}} = \underbrace{w_c(\mathbf{x})}_{\substack{\uparrow \text{class frequency} \\ \text{중요}}} + w_0 \cdot \exp \left(- \frac{\overbrace{(d_1(\mathbf{x}) + d_2(\mathbf{x}))^2}^{\text{distance}}}{2\sigma^2} \right)$$

- 초기 weight 설정이 중요함
 - 모든 부분이 동등하게 반영될 수 있도록 하여야 함
 - Gaussian distribution을 따르도록 가중치 설정

3-1. Data Augmentation

- 매우 적은 수의 annotated 이미지로 세분화 네트워크를 효과적으로 훈련시키기 위해 데이터 증강 기법을 활용함

1. 필요성

- 데이터 증강은 훈련 샘플이 제한적인 경우 네트워크에 원하는 불변성과 견고성 특성을 가르치는 데 중요

2. 적용

- 현미경 이미지에서는 주로 이동 및 회전 불변성, 변형 및 회색값 변동에 대한 견고성이 필요

3. 주요 기법

- 무작위 탄성 변형
 - 거친 3x3 그리드에서 무작위 이동 벡터를 사용하여 부드러운 변형을 생성
 - 이동은 정규 분포에서 샘플링되며, 바이큐빅 보간을 통해 각 픽셀의 이동이 계산됨
- 드롭아웃 레이어
 - 수축 경로의 끝에서 드롭아웃 레이어를 사용하여 추가적인 암묵적 데이터 증강을 수행

⇒ 매우 적은 양의 주석이 달린 이미지로도 뛰어난 세분화 네트워크를 훈련시킬 수 있는 중요한 도구임

4. Experiments

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

1. 전자 현미경 이미지에서의 신경 구조 세분화

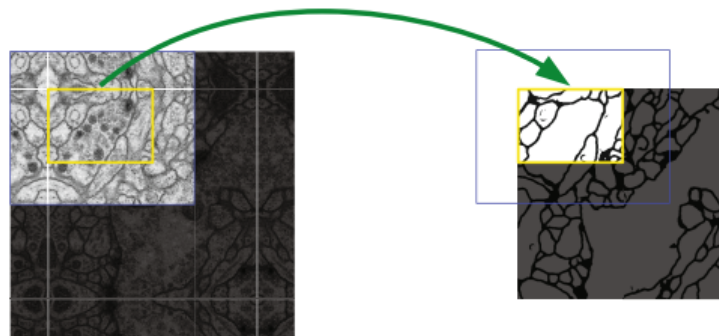


Fig. 2. Overlap-tile strategy for seamless segmentation of arbitrary large images (here segmentation of neuronal structures in EM stacks). Prediction of the segmentation in the yellow area, requires image data within the blue area as input. Missing input data is extrapolated by mirroring

- 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 슬라이딩 윈도우 컨볼루션 네트워크보다 훨씬 우수한 성능을 보였음

2. 광학 현미경 이미지에서의 세포 세분화

- **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

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