

# U-Net 리뷰

이름: 서 동 관

## 1. 연구의 배경

**CNN**은 시각적 인식 작업에서 인간을 뛰어 넘는 큰 향상을 가져왔다. 그러나 학습을 위해서는 많은 훈련데이터와 라벨링이 필요해 물적 경제적 시간적 비용이 많이 소비되는 단점이 있었다. 본 연구에서는 이러한 단점을 극복하기 위해 데이터 증강이라는 네트워크 전략을 제시하고자 하였다. 이러한 연구는 생물학 분야에 있어 부족한 데이터로 부터 이미지를 분류하는데 이전의 모델을 능가하는 효과와 시간도 많이 감축할 수 있음을 확인하였다.

## 2. 목적

**CNN**의 역사는 오래 되었지만 사용가능한 데이터가 극히 제한적이었으며 데이터의 물리적 양도 많이 필요하기에 성공이 제한적이었다. 이러한 문제를 해결하고자 크리제프스키외의 연구에서 수백만개의 매개변수를 가진 네트워크 모형을 제시하였고, 그 후 더 크고 깊은 네트워크들이 속속 등장하였다. 그러나 **CNN**은 일반적으로 단일 이미지 분류가 대다수 였으며 생물학적인 이미지 처리에 있어 분류는 포함되지 않았다. 따라서 본 연구에서는 생물학적인 이미지 처리의 합리적 방법을 제시하는데 목적이 있다.

## 3. 필요성

본 연구에서는 생물학의 **CNN** 적용에 있어 데이터의 과소를 데이터 증강으로 해결하는 방법을 제시하여 향후 생물학 이미지 분류에 있어 딥러닝의 적용을 한 단계 향상시키는 역할을 할것이며, 보다 빠른 세포 분류로 생물학의 발전에 영향을 미칠것이다.

## 4. 실험 설계:

본 연구에서는 **FCN**을 수정 확장하였으며 주요 아이디어로는 풀링 연산자가 **Up-sampling** 연산으로 대체되는 연속적인 계층으로 네트워크를 구성하였다.

U-Net은 이미지의 전반적인 컨텍스트 정보를 얻기 위한 네트워크와 정확한 지역화(Localization)를 위한 네트워크가 대칭 형태로 구성되어 있다.

Expanding Path의 경우 Contracting Path의 최종 특징 맵으로부터 보다 높은 해상도의 Segmentation 결과를 얻기 위해 몇 차례의 Up-sampling을 진행한다. 다시 말해, Coarse Map에서 Dense Prediction을 얻기 위한 구조이다.

Coarse Map to Dense Map 개념 뿐만 아니라 U-Net은 FCN의 Skip Architecture 개념도 활용하여 얇은 층의 특징맵을 깊은 층의 특징맵과 결합하는 방식을 제안하였다.

이러한 CNN 네트워크의 Feature hierarchy의 결합을 통해 Segmentation이 내제하는 Localization과 Context(Semantic Information) 사이의 트레이드 오프를 해결할 수 있다.

수축형태와 팽창형태의 네트워크 구성 구조를 다음과 같이 편성하였다.

#### The Contracting Path

- 3x3 convolutions을 두 차례씩 반복 (패딩 없음)
- 활성화 함수는 ReLU
- 2x2 max-pooling (stride: 2)
- Down-sampling 마다 채널의 수를 2배로 늘림

Expanding Path는 Contracting Path과 반대의 연산으로 특징맵을 확장한다.

#### The Expanding Path

- 2x2 convolution ("up-convolution")
- 3x3 convolutions을 두 차례씩 반복 (패딩 없음)
- Up-Conv를 통한 Up-sampling 마다 채널의 수를 반으로 줄임
- 활성화 함수는 ReLU
- Up-Conv 된 특징맵은 Contracting path의 테두리가 Cropped된 특징맵과 concatenation 함
- 마지막 레이어에 1x1 convolution 연산

위와 같은 구성으로 U-Net은 총 23-Layers Fully Convolutional Networks 구조이다. 주목해야 하는 점은 최종 출력인 Segmentation map의 크기는 Input Image 크기보다 작다는 것이다.

그 이유는 Convolution 연산에서 패딩을 사용하지 않았기 때문이다.

그리고 U-Net 연구팀은 크기가 큰 이미지의 경우 이미지 전체를 사용하는 대신 overlap-tite 전략을 사용하였다.

이같은 방법은 이미지를 타일로 나눠서 입력으로 사용한다는 점으로, 다음 tile에 대한 Segmentation을 얻기 위해서는 이전 입력의 일부분이 포함되어야 한다.

이러한 이유로 Overlap-Tite 전략이라 칭한다.

이미지의 경계 부분 픽셀에 대한 세그멘테이션을 위해서는 0이나 임의의 패딩값을 사용하는 대신 이미지 경계 부분의 미러링을 이용한 Extrapolation 기법을 사용하였다.

네트워크의 출력 값은 softmax로 loss function은 cross-entropy, 활성화 함수는 relu를 stride는 2로 맞추어 데이터 크기를 줄였다.

U-Net의 가중치는 한 뉴런에 들어오는 노드의 개수를  $N$ 이라고 하면  $\text{root}(2/N)$ 의 표준 편차를 가진 가우시안 분포를 이용해 가중치를 초기화 했다.

학습데이터의 부족으로 데이터 증강을 통해 이를 해결하였는데 여기서 **Data Augmentation**이란 데이터가 한정되어 있을때 데이터를 회전, 반전과 같은 여러 효과를 주어 데이터의 크기를 키우는 방식이다.

**training** 데이터가 한정되어있을때 사용하는 방식인데, **U-net**이 세포 **segmentation**에서 강점을 보이는 **network**이고, 세포 데이터에선 **Data Augmentation**의 적용이 나쁘지 않은 효과를 가져오기 때문에 이를 사용하였다.

## 5. 검증

**Data set**은 **ISBI 2012**에서 시작된 전자파 세분화 과제로 개방 데이터를 활용하였다. 이 훈련 데이터는 항문 신경 코드(**VNC**)에 있는 **Drosophila**의 직력 섹션 전송 전자 현미경의 30개 이미지를 사용했다.

실험결과 **warping error**는 0.000353, **rand error**는 0.0382, **pixel error**는 0.0611를 기록하여 기존에 제시된 결과보다 더 나은 성능을 확인하였다.

그 후 **u-net**을 일반적인 현미경 이미지의 세포 분할 작업에 적용했다.

첫번째 데이터는 **Phc-U373**의 위상 대비 현미경으로 기록된 폴리아크릴리미드 기질에 교아세포종-아스트라세포종 **U373**세포 등으로 라벨링이 된 35개의 훈련 이미지이다.

실험결과 평균 **IOU**가 92%로 나타나서 기존 모델보다 더 나은 결과를 달성했다.

두번째 데이터로 **DIC-HeLa**라는 차동 간섭 대조 현미경으로 기록된 **HeLa**세포이다.

이 이미지는 훈련이미지가 20개로 라벨링이 포함되어 있다.

실험결과 평균 **IOU**는 77.5%를 달성하여 기존 최고 모델보다 더 나은 결과를 확인했다.

## 6. 결론

**U-NET**은 매우 다른 생물의학 분야에서 매우 우수한 성능을 달성 했다.

라벨링은 필요하지 않으며, **NVidia Titan GPU**에서 10시간의 짧은 훈련으로 학습이 가능하고, 더 많은 작업에 쉽게 적용될 수 있다고 주장하였다.