

Breast Cancer Tumor Detection

With Attention U-net and Contour Labeling

박기승 , 정운영 , 장주현

1. Introduction

2. Data

3. EDA

4. Attention U-net

5. Bounding box with Contours

6. Dacon

Introduction

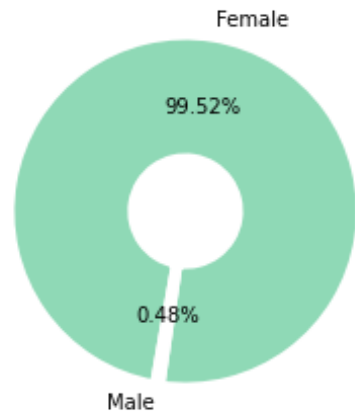
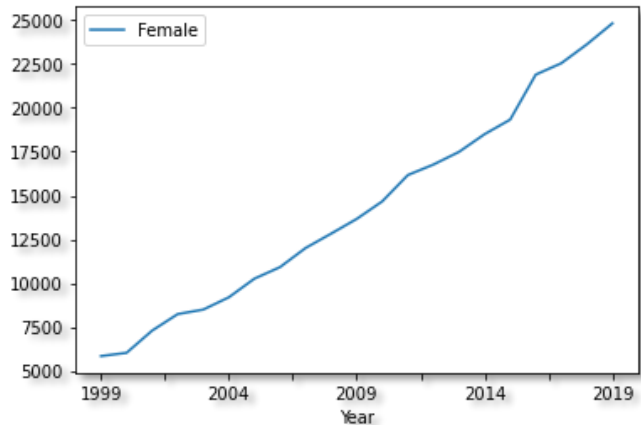
- 여성인구 대비 유방암환자는 해를 거듭할수록 증가하는 추세
- 여러 암들의 공통적인 원인을 제외하곤 뚜렷한 원인이 밝혀지지 않은 유방암
- 물리적인 절제술이 유일한 예방법이며 종양에 대한 의학적 초기 진단이 중요
- 국내에서도 전체 오진 10건 중 6건이 암 오진이며 여성의 경우 유방암 오진 피해가 가장 많음
- 따라서 의사의 진단과 컴퓨터 보조 진단 시스템 과정을 추가로 거쳐 더욱 정확한 검출이 필요함 Misdiagnosis(2017)

현재 유방암 분야의 한계

유방암 대표적 표준검사인 유방촬영술은
유방 전체가 하얗게 보이고 경계가 모호해
종양을 찾아내기 힘들
Diagnosis and Reatment of Breast Cancer

연구 목적

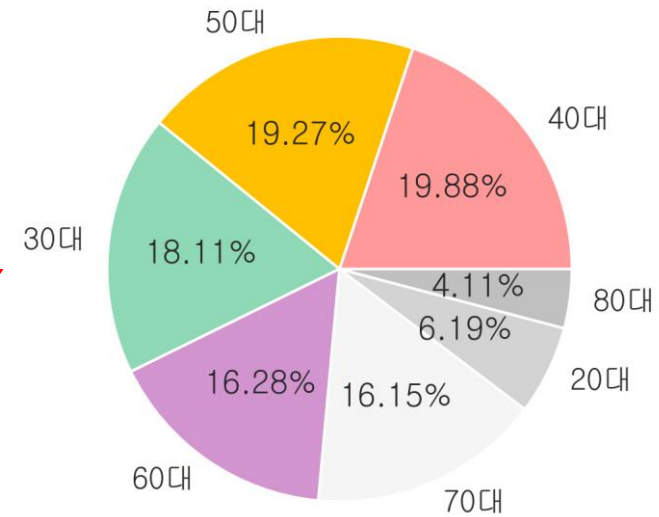
Attention U-net을 통한 종양 파악 및
Contour 기술을 적용한
종양 Level 부여를 진행하고자 함



여성환자 비중 99.52%

여성환자 수는 점점 증가하는 추세

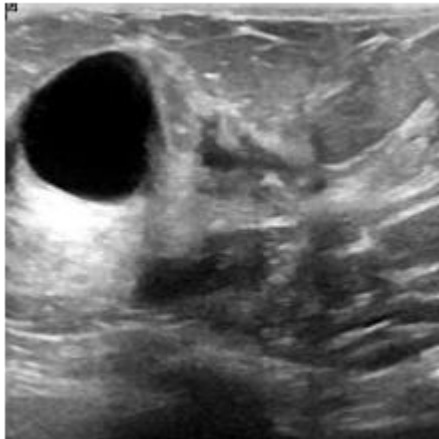
여성환자의 연령 별 발병 빈도



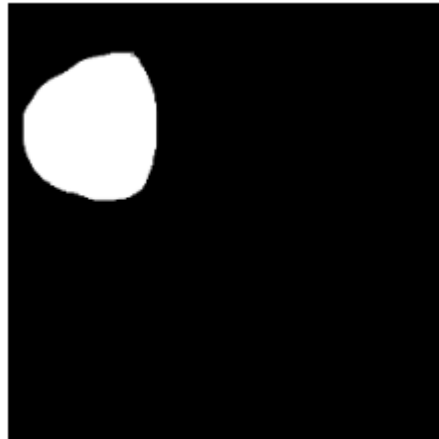
중년 여성의 경우에서 가장 많은 빈도

Breast Cancer Image

Image

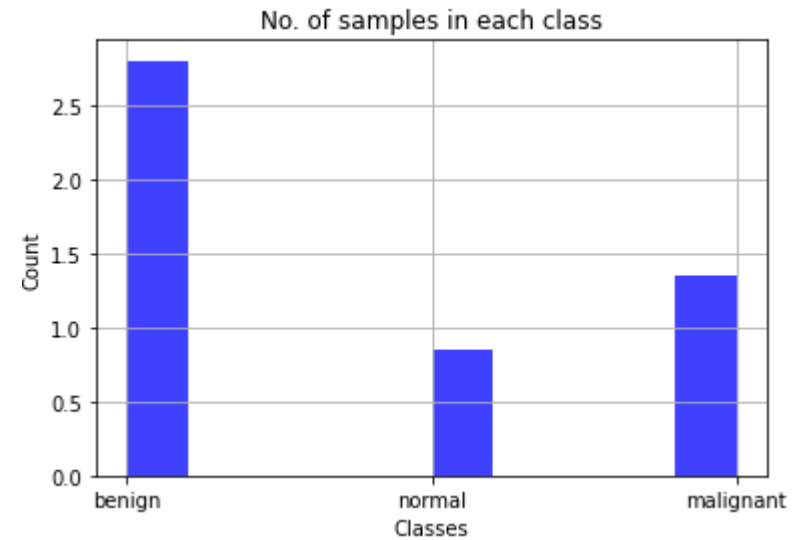


Mask



25-75세 사이 여성 유방암 환자 600명 초음파 사진

500*500의 이미지 780장

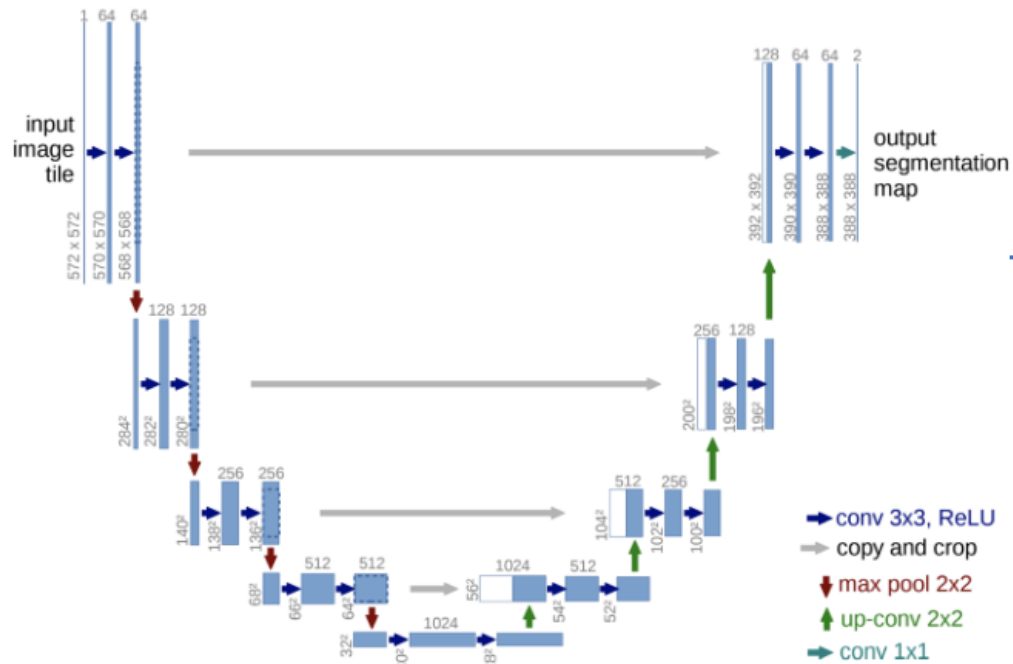


Benign : 437
Malignant : 210
Normal : 133

U-net

‘Convolutional Networks for Biomedical Image Segmentation’ 논문에서 제안된 구조

적은 수의 학습 데이터로도 높은 정확도의 Image Segmentation 성능을 보여주는 모델



신경망 구조를 스킵 연결을 통해
좌우 대칭이 되도록 레이어를 구성

해당 모습이 U자 형태라
U-net이라 부름

Image Segmentation의 인코더-디코더 관계

인코더 -> 채널 수 **증가** / 차원 **축소** : 인코더 단계 이미지만의 특징 포착

디코더 -> 채널 수 **감소** / 차원 **증가** : 인코딩 과정에서 축소된 차원을 고차원으로 복원

위 인코더-디코더 관계의 취약점

1. 인코딩 단계의 차원 축소를 거치며 자세한 좌표 정보를 손실하게 됨
2. 디코딩 단계에서 저차원의 정보를 가져오기 때문에 손실된 정보에 대한 회복이 어려움

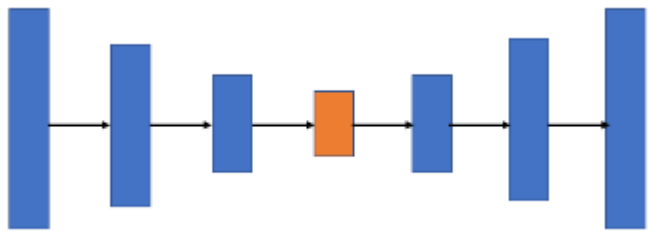
U-net의 인코더-디코더 관계

인코더와 디코더 사이의 연결 방식을 스킵 연결(Skip Connection) 방식으로 적용

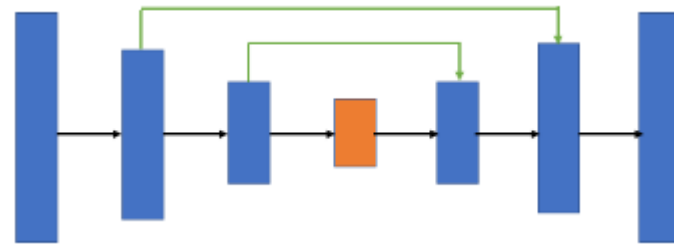
저차원과 고차원의 정보를 함께 이용해 최종 이미지의 특징을 추출하여 정확한 위치 정보 파악을 가능하게 함

일반연결 스킵연결

스킵연결 : 인코딩 단계의 각 레이어에서 얻은 특징을 디코딩 단계의 각 레이어에 합치는 방법

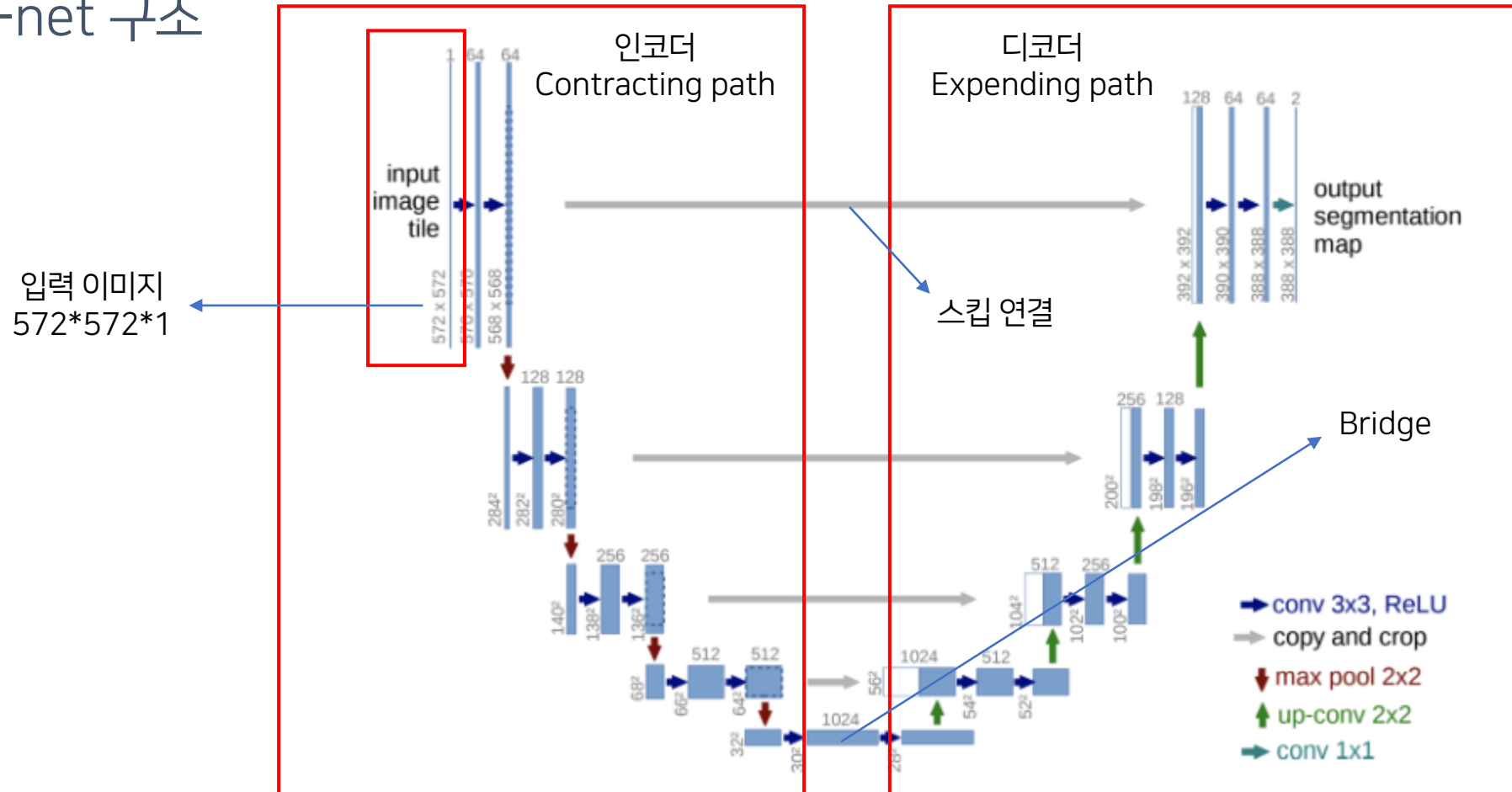


일반 연결방식의 인코더-디코더



스킵 연결방식의 인코더-디코더

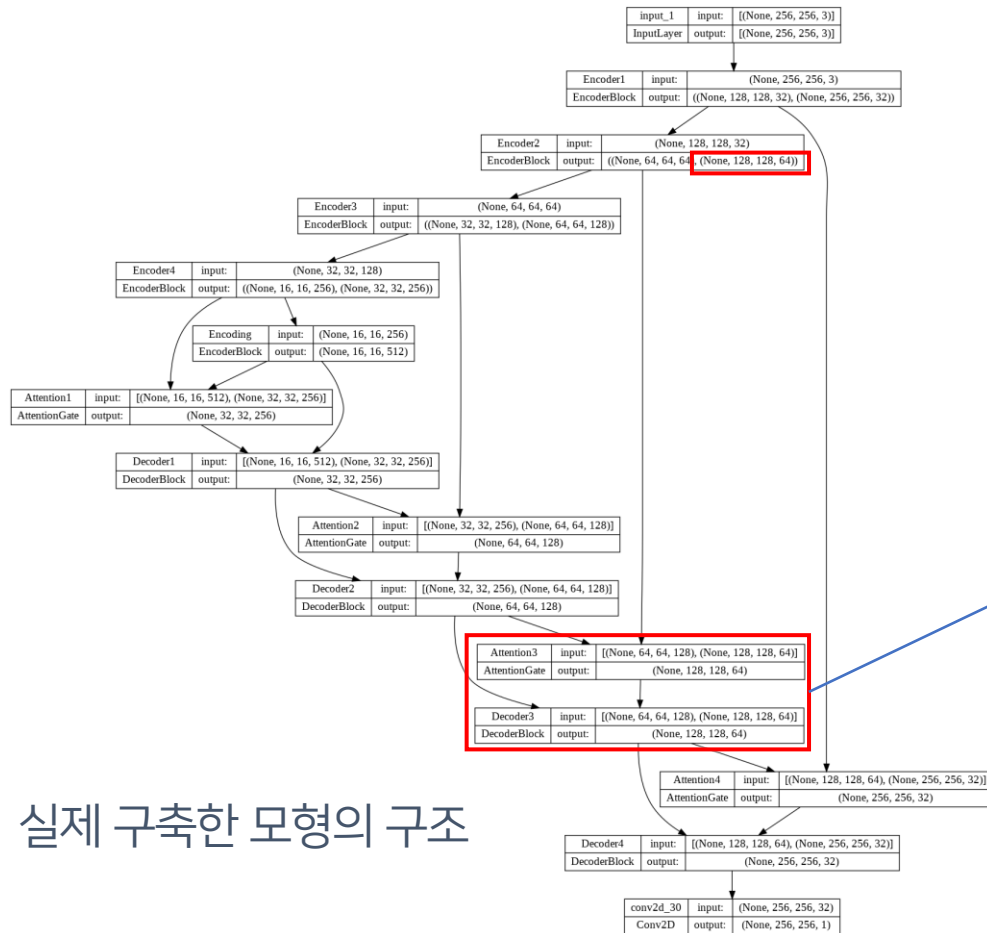
U-net 구조



Attention U-net

Attention

디코더의 해당하는 레이어의 인코더 이미지를 집중(Attention)하여 최종 출력을 결정하므로 Attention이라고 부름



다음 Decoder로 차원이 증가하기 전

Attention Gate를 통해
해당 레이어의
Encoder 이미지 정보를 이용

실제 구축한 모형의 구조

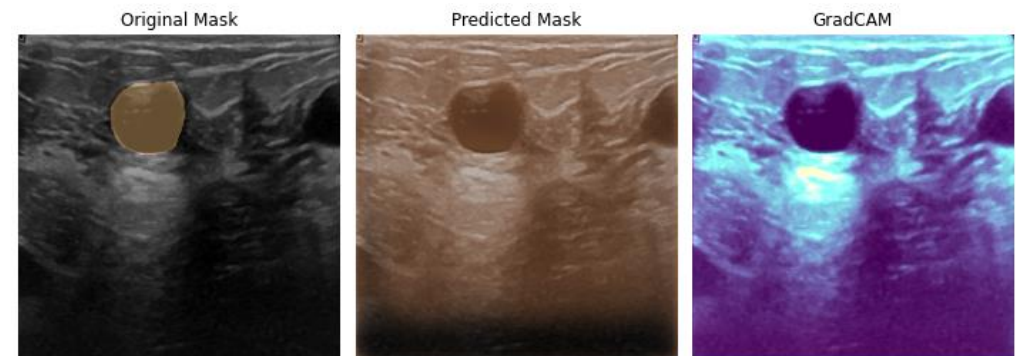
Attention U-net

Model Train

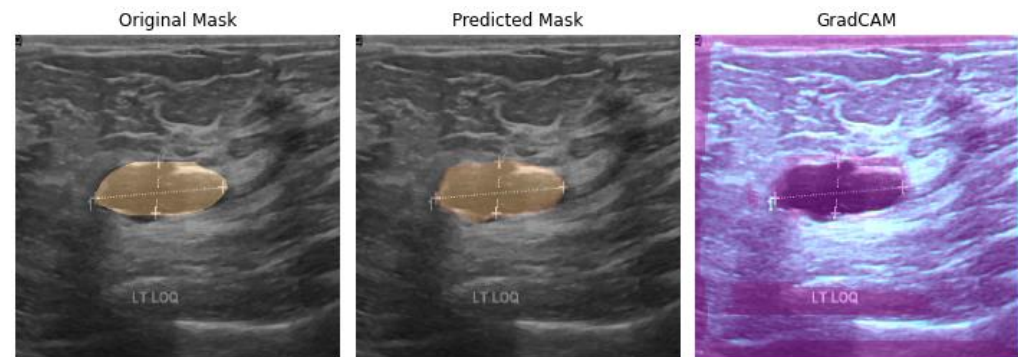
Batch Size : 8 / Epoch : 100 / Optimizer : Adam

여기서 Steps_per_epoch를 활용해
Epoch마다 매개변수를 업데이트해주며 학습함 (9,700번)

Epoch 1

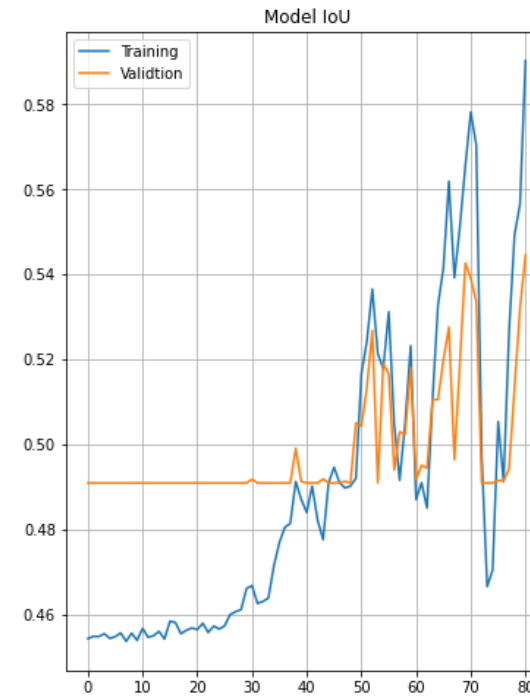
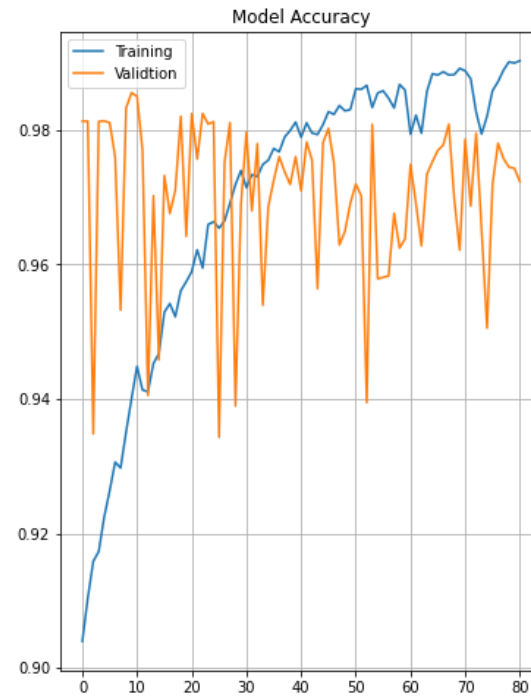
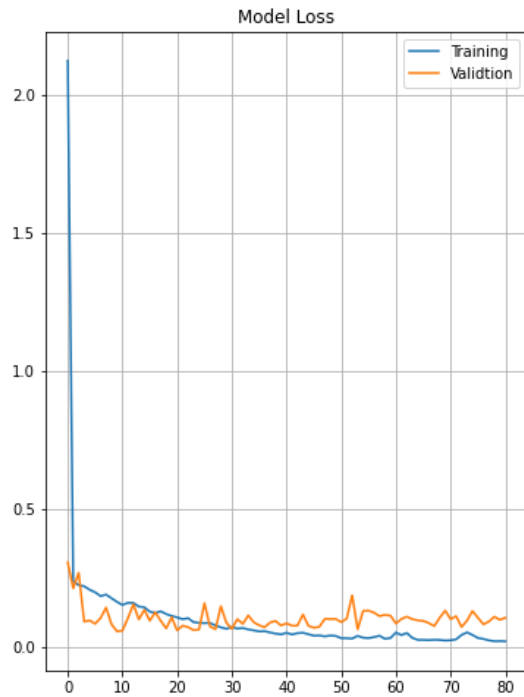


Epoch end



Attention U-net

Model Train



Accuracy

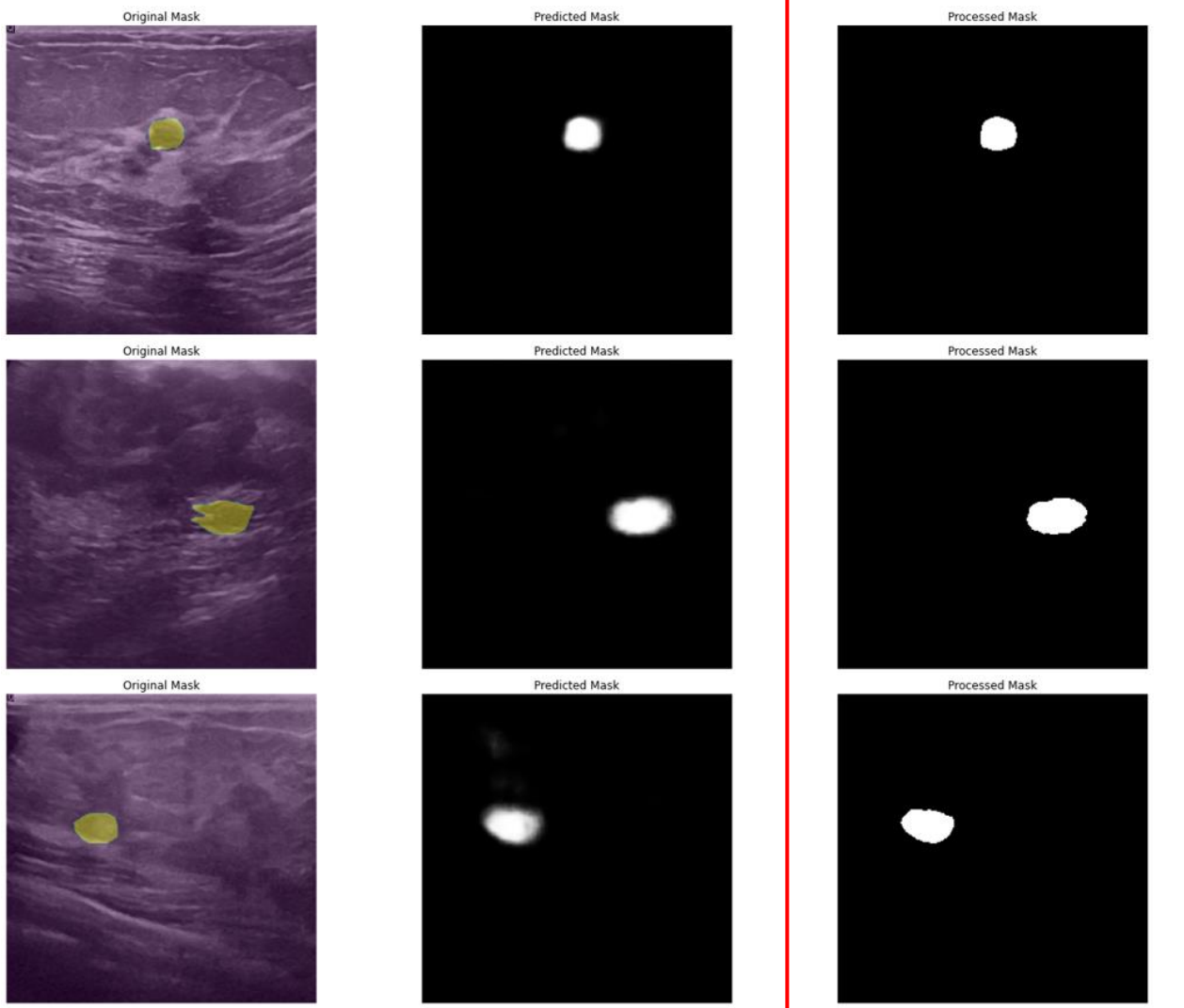
Train: 99.03
Val : 97.23

IoU

Train: 59.03
Val : 54.45

Contours

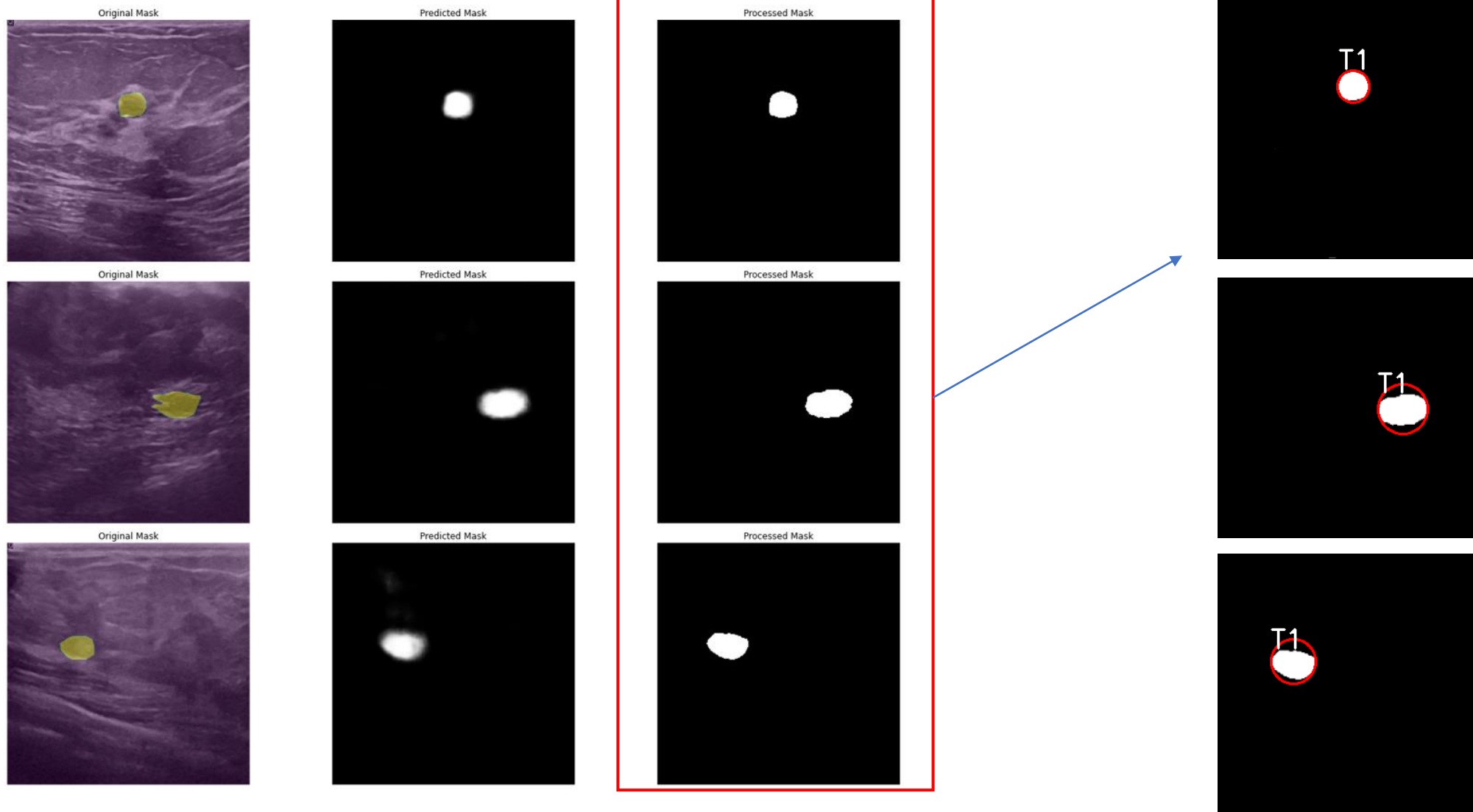
Model Evaluation



최종 예측 후
희미한 가장자리 부분을 제거하는
후처리를 진행하여
최종 예측 이미지 생성

Contours

Model Evaluation



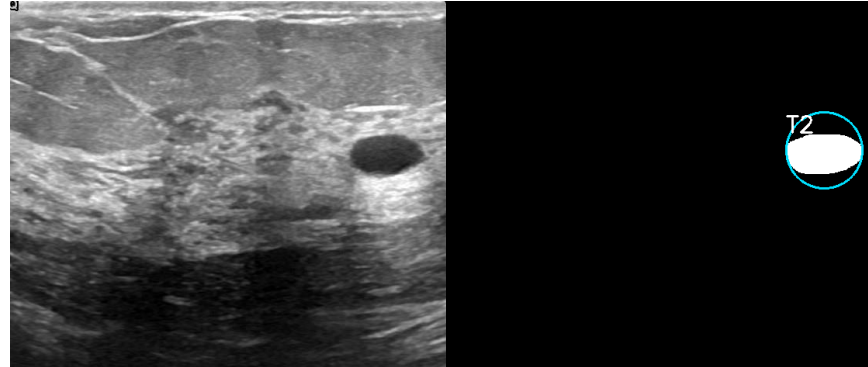
Contours

종양 크기에 따른 컬러라벨을 부여한 최종 결과

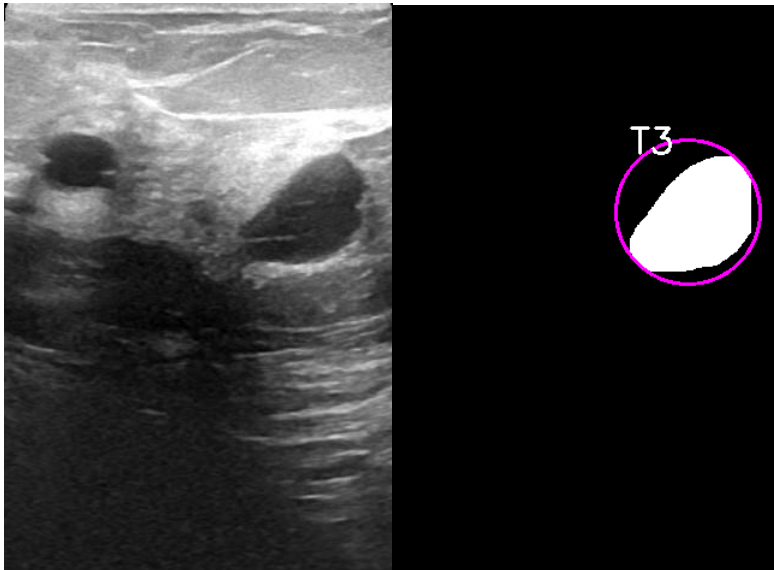
T1



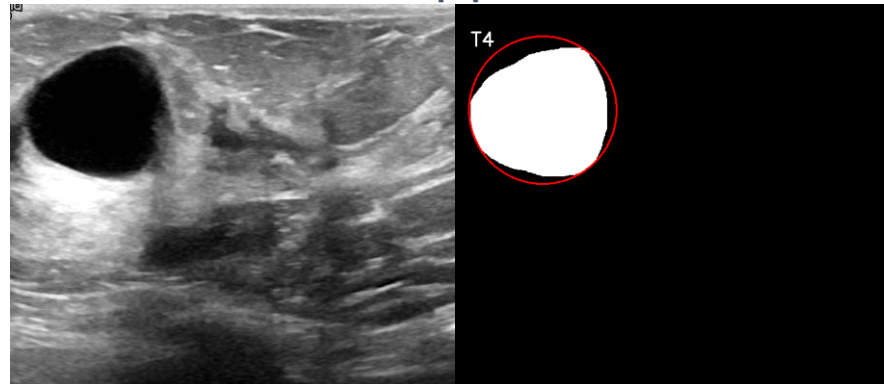
T2



T3



T4



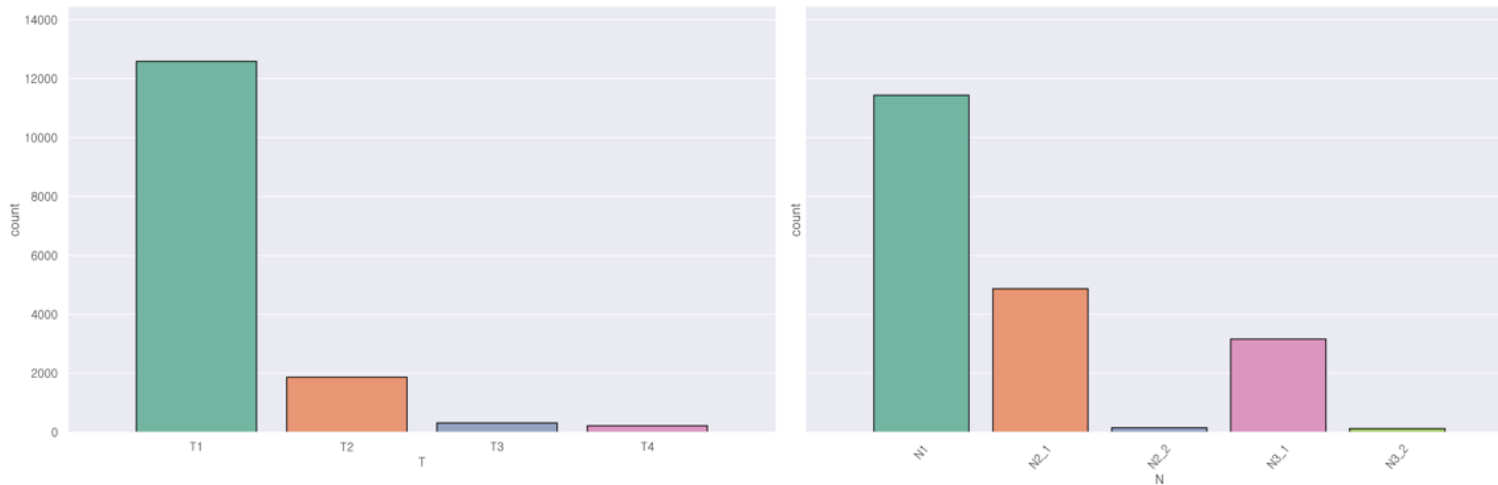
T1-4

종양 크기에 대한 Stage

Dacon 유방암의 임파선 전이 예측



유방암 및 임파선 전이 Plot



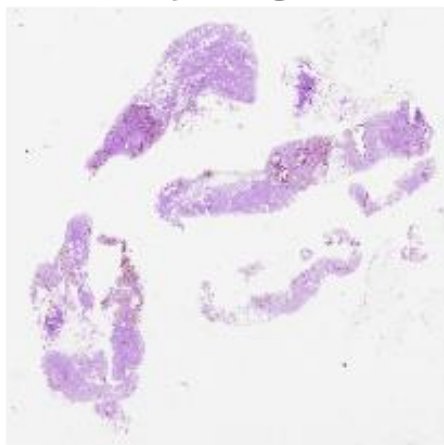
T(Tumor) : 유방암 종양 크기
N (Node) : 전이된 임파선의 정도

유방암 조직은 임파선을 통해
다른 부위로 전이 될 가능성이 높음

유방암 진단에 있어서
임파선 전이에 대한 파악은 중요함

Modeling

Input image



Histogram of Oriented Gradients



임파선 Masking



Image

Train : 1,000

Test : 250

Mask : 58개

Model : Mobile Net V2

F1-Score : 0.6421