

유방암 진단을 위한 조직병리학적 이미지 인식 정확도 개선 연구

산업공학 종합설계 2

김예빈

양효용

최보미

최윤정

2020. 11. 15.

Contents

1. Introduction.....	3
1.1 유방암 진단 이미지 인식 연구 필요성	3
1.2 유방암 진단 이미지 인식 연구 주제.....	6
2 Literature Review	6
3 Research Methodology.....	12
3.1 데이터	12
3.2. 연구 방법.....	14
3.2.1 CNN architecture 에 따른 유방암 진단 정확도 비교.....	14
3.2.2 조직병리학적 진단 요인을 고려한 이미지 특징강화 방법 적용에 따른 정확도 비교.....	16
3.2.3 CAM(class activation map)시각화를 통한 설명가능성 향상.....	30
3.2.4 Grad-CAM 을 통한 설명가능성 향상.....	31
4. 연구 결과	30
4.1 CNN architecture 에 따른 유방암 진단 정확도 비교	30
4.2 조직병리학적 진단 요인을 고려한 이미지 특징강화 방법 적용에 따른 정확도 비교.....	33
4.3 CAM(class activation map)으로 설명가능한 진단결과 확인.....	34
4.4 Grad-CAM 결과.....	39
5 Conclusion & Recommendations	41
6 References	42

유방암 진단을 위한 조직병리학적 이미지 인식 정확도 개선 연구

1. Introduction

1.1 유방암 진단 이미지 인식 연구 필요성

의료에서 가장 중요한 질환 질환 중 한가지는 암질환이며, 전체 사망에서 26.5% (2018 년 통계청)이 암이 원인이며, 여성의 유방암은 전 세계적으로 매년 200 만명 이상이 발생하는 중요한 문제이다. 유방암의 조기진단을 할 때 조직검사는 병리과 전문의 판독을 통하여 암 진단의 여부가 결정하는데 유방조직 이미지를 이용하여 인공지능으로 조기진단을 할 수 있게 된다면 의료비 절감 및 오진률을 감소시킬 수 있어 의료서비스에 도움을 줄 수 있을 것이다. 실제로 분당서울대병원에 따르면 유방촬영술에서 악성이 의심되어 조직검사를 시행했을 때 실제로 악성으로 확진 되는 경우는 30%로 조직검사의 이미지를 통한 검사가 필수적이라고 한다. (분당서울대병원 유방암센터, n.d.). 또한 조직검사 전에 시각적 분석을 위해 염색 후 진단 과정에서 결정을 내릴 때 집중적으로 보는 부위를 전체 슬라이드 조직 스캔으로 결정한다고 한다.

I 조직병리검사 과정



Figure 1. 조직검사에 소요시간은 작은 조직은 약 1-2 일, 큰 조직은 약 3-4 일이 소요된다고 한다.

즉 조직검사 시간 단축을 위한 연구가 필요한 상황이다. (홍은경, 2013)

또한 위의 그림처럼(Fig.1) 조직검사에 소요되는 시간이 약 1-4 일이 소요되며(홍은경,2013) 다양한 배율로 현미경 이미지를 분석하는데 병리학자들은 패턴, 질감 및 다양한 형태학적 특성, 배율이 다른 이미지 들에 대해서 검사를 해야 하기 때문에 만약 이 일부 작업을 자동화하는데 도움을 줄 수 있는 알고리즘이 필요하다고 주장한다. (Alom et.al, 2020) 따라서 이 연구에서는 유방촬영술을 통한 이미지가 아닌 조직검사 이미지를 사용해 인식할 수 있는 cnn architectue 를 개발하고 정확도를 높일 수 있는 방법과 함께 의사들의 업무 프로세스를 돕기 위해 이미지 인식에 영향을 준 부분에 대해 제안할 수 있는 알고리즘을 제안하고자 한다.

또한 본 저자들이, 현미경에서 얻은 조직 이미지를 전처리 없이 8 가지 진단 결과 (악성(malignant): 유관상피내암(Ductal carcinoma in-Situ, DCIS), 소엽암(lobular carcinoma), 유방점액성암(Mucinous Carcinoma), 유두상암종(papillary_carcinoma), 양성(benign): 선증(adenosis), 유방섬유선종(fibroadenoma), 엽상 종양(phyllodes tumor), 관상선종(tubular_adenoma)를 class 로 cnn 으로 다음과 같이 epoch 20 번으로 학습시켰을 경우, validation accuracy 는 64%를 보였지만 test accuracy 는 12%로 매우 심각한 과적합을 볼 수 있었다.

Layer (type)	Output Shape	Param #
input_1 (InputLayer)	[(None, 150, 150, 3)]	0
conv2d (Conv2D)	(None, 148, 148, 32)	896
max_pooling2d (MaxPooling2D)	(None, 74, 74, 32)	0
conv2d_1 (Conv2D)	(None, 72, 72, 64)	18496
max_pooling2d_1 (MaxPooling2D)	(None, 36, 36, 64)	0
conv2d_2 (Conv2D)	(None, 34, 34, 128)	73856
conv2d_3 (Conv2D)	(None, 32, 32, 64)	73792
max_pooling2d_2 (MaxPooling2D)	(None, 16, 16, 64)	0
flatten (Flatten)	(None, 16384)	0
dense (Dense)	(None, 512)	8389120
dense_1 (Dense)	(None, 8)	4104

Table 1. 유방암 조직이미지를 8 가지 진단결과(악성(malignant):유관상피내암(Ductal carcinoma in-Situ, DCIS), 소엽암(lobular carcinoma), 유방점액성암(Mucinous Carcinoma),

유두상암종(papillary carcinoma), 양성(benign): 선증(adenosis), 유방섬유선종(fibroadenoma), 엽상종양(phylloides tumor), 관상선종(tubular adenoma)) class 로 학습시킨 cnn architecture

이는 8 가지 class 에 대해 12%정확도는 랜덤하게 결과를 내보내는 결과와 마찬가지로 매우 낮은 것을 확인할 수 있었는데 wenzhang(2020)의 논문에서는 배율이 높아질수록 검증 정확도와 테스트 정확도가 차이가 많이 나는 문제가 있다고 지적했으며, AlexNet, ResNet, Inception module 을 합치는 방식을 제안하는 등 과적합을 막기위한 연구들이 진행되고 있다. 또한 조직 검사에서 진단을 내릴 때 세포핵이 주요 의사결정 지점이 되는데 세포핵은 일반적으로 흐릿하게

보이기 때문에 핵과 주변 사이의 경계를 높여 이를 CNN 이 인식하게 만들어야 한다는 점을 Loukas(2013) 등의 논문을 통해 떠올려 이를 해결할 수 있는 방법을 제안한다.

1.2 유방암 진단 이미지 인식 연구 주제

본 저자들은 현미경 조직검사 이미지를 통해 악성(malignant):유관상피내암(Ductal carcinoma in-Situ, DCIS), 소엽암(lobular carcinoma), 유방점액성암(Mucinous Carcinoma), 유두상암종(papillary_carcinoma), 양성(benign): 선증(adenosis), 유방섬유선종(fibroadenoma), 엽상종양(phyllodes tumor), 관상선종(tubular_adenoma), 8 가지 진단 결과의 정확도를 높이기 위한 cnn architecture 와 불투명한 세포핵을 분리하여 정확도를 개선할 수 있는 방법 등을 연구하였다. 이를 위해서 유방암 이미지 인식 연구들의 방법을 논문을 통해 조사하고, 기존의 조직병리학에서는 암을 진단하는 기준을 조사하였다. 그리고 데이터를 학습시킬 때 층의 깊이에 따른 정확도와, 인셉션모듈을 사용했을 때의 정확도, 배율에 따른 정확도를 비교해보았고 기준값을 넘으면 최대값 아니면 0 임계처리, 기준값을 넘으면 원래값 아니면 0, 흑백변환, 평균 적응임계처리, 가우시안블러 적응임계처리로 특징을 부각한 후에 정확도를 비교해보았다. 그 후 설명가능성을 높이기 위해 class activation map 을 시각화하여 각 레이어의 필터의 가중치를 시각화 해보았다.

2 Literature Review

2.1. 유방암 이미지 인식 연구

먼저, 유방암에 대한 논문을 살펴보았을 때 "유방 초음파 영상에서의 영상학적 특징들을 이용한 유방암 진단.(권보준,2013)" 논문에서는 유방 영상의 병변의 영역으로부터 특징들을 추출하여 기하, 통계, 밝기, 텍스처, 경계, 모양에 대한 특징들을 연구하였다. 긴 병변보다 위 아래로 분포한 병변이 암일 확률이 높았고 병변의 경계가 불분명할 경우 암이 될 확률이 높아서 이를 나타낼 수

있는 특징을 디자인하여 경계의 명확성을 경계 안과 밖이 얼마나 차이가 큰가를 주변 픽셀들에 대해 밝기 변화량을 비교하여 경계 픽셀들에 대해 히스토그램으로 만들었다. 그리고 랜덤 포레스트 방법으로 특징들을 가지고 학습을 진행하였다고 한다. (권보준. "유방 초음파 영상에서의 영상학적 특징들을 이용한 유방암 진단." 국내석사학위논문 서울대학교 대학원, 2013.)

통계적으로 병변 내 영역의 밝기 값의 최소, 최대 평균, 표준편차, 파워, 최빈값, 백분위 등으로 병변 내부의 밝기 값의 종합적인 분포를 나타낼 수 있는 값들을 사용하였다. 밝기 분포는 영역 내 밝기 값의 히스토그램으로 표현한다.

Wenzhong 의 논문에서는 ResNet 과 Inception 을 사용해서 모델을 만들면 시스템의 정확도가 배율이 증가함에 따라서 점차 감소했다고 밝혔다. 그러면서 DeepBC 모델을 제안하면서 레즈넷과 인셉션모듈, 알렉스넷을 모두 합친 아키텍처를 사용해 기존의 정확도보다 높은 성능을 보였다고 말했다. (Wenzhong, 2020)

Loukas 는 kNN, SVM 및 PNN 을 사용해서 유방암 진단 이미지 인식 방법을 제안하였다. 현미경을 통한 조직학적 단면 이미지에서 겹치는 세포를 제거하거나 총체적인 구조를 제거하여 반영해야 한다고 한다. (Loukas,2013)

Alom 의 논문에 따르면 (Alom et al., 2019) 논문에서는 Inception Recurrent Residual Convolutional 신경망 (IRRCNN) 모델을 사용하여 악성, 양성 이진 분류 및 진단결과에 따른 다중클래스로 인식을 하였다. 유방암 진단을 위해서 SIFT (Scale Invariant Feature Extraction), LBP (Local Binary Patterning), LPQ (Local Phase Quantization), GLCM (Gray-Level Co-occurrence Matrix), 임계 값 인접 통계(TAS) 및 Parameter Free TAS (PFTAS) 등이 연구에 사용되어 정확도가 85% 정도로 진행되어 왔지만 이 연구에서는 Recurrent Convolutional Layers 를 추가하여 더 나은 성능을 보였고, 크기조정, 이미지 자르기, 데이터 증강기술 등을 통해 최고 99% 성능을 보였다고 한다.

기존의 조직병리학 이미지 분석 방법은 대부분 세포핵 분할 및 분류를 기반으로 한다.

이미지에서 세포핵을 인식할 때에는 Laplacian blob detector 알고리즘을 사용했다고 한다. 하지만 조사한 George의 논문에서는 자동 유방암 검출을 위해 'NucDeep'을 이용한 핵 특성 추출 방법을 이용했는데 이미지에서 감지된 오버랩되지 않는 핵 패치를 사용하면 피쳐 추출에 필요한 낮은 복잡도 CNN을 설계할 수 있었다고. 또한 SVM(FF + SVM)를 사용하는 피쳐 퓨전 접근 방식을 이용해 유방 종양 이미지를 추출된 CNN 기능을 기반으로 분류하고 피쳐 퓨전 방법이 변환시킨다. 또한 로컬 핵 피쳐를 컴팩트한 영상 수준 피쳐로 변환하여 분류기 성능을 향상시킨다. 이미지 레벨 분류를 위해 패치 클래스 확률 기반 의사결정 체계(NucDeep + SVM + PD)도 사용한다. 이 NucDeep + FF + SVM 모델은 기존의 여러 가지 방법을 능가하는 것으로 확인되었으며, 교육 복잡성이 낮음에도 불구하고 유사한 수준의 예술 성능을 보여주었다.

(Computer assisted recognition of breast cancer in biopsy images via fusion of nucleus-guided deep convolutional features, Kalpana George, 2020)

Kumar의 논문에서는 현미경 이미지의 형상, 배경 세포의 세분화, 특징추출 등을 사용해서 구조의 특징을 강조하기 위한 방법을 사용했다고 한다. 텍스처 기능, 색상 기반 기능, 그레이 레벨 동시 발생 매트릭스 (GLCM), Law's Texture energy (LTE) 방법, Tamura의 특징추출 방법 등으로 정확도를 높였다. (Kumar, 2015)

Alzubaidi에서는 환자에 따른 과적합 영향을 줄이기 위해 따라 다양한 이미지 처리 기술로 이미지를 확대하는 기술로 정확도를 높였다. 오른쪽 그림과 같이 회전 및 밝기를

변화시키는 방법으로 데이터 증강기술을 사용하여 과적합을 줄이는 방법을 사용하였다. 데이터 증강기술을 사용하여

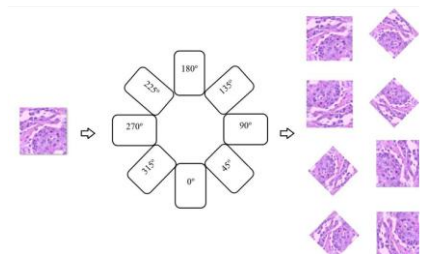


Figure 2. 논문에서 제시한 세포의 회전에 따른 과적합을 줄일 수 있는 방법 관련 figure.

세포의 회전에 따른 과적합을 줄일 수 있어 본 종합설계 연구에서도 필요성을 느꼈다. (Alzubaidi, 2020)

Deep Neural Network with Support Value (DNNS)을 이용해 더 나은 품질의 이미지를 생성하고 다른 성능 파라미터를 수정하기 위해 도입 데이터 부족 문제를 해결하고자 했다. 또한 더 나은 예측과 진단을 위해 히스토그램 값, Sigmoid 함수 및 Histo-sigmoid 함수의 계산을 위한 수학 방정식이 수정 및 업데이트를 하고 나머지 방정식은 표준 BC 검출 알고리즘과 유사하게 진행한다. 크게 세 가지로 DNNS를 수행한다. 첫 번째 단계에서는 Pre-processing(사전 처리) 단계에서는 효과적인 필터링 기술을 이용해 입력 세포학 영상이 사전 처리하여 노이즈를 제거한다. 두 번째 단계에서는 사전 처리된 영상에서 엔트로피, 기하학적 및 텍스트 기능을 추출한다. 세 번째 단계는 Histo-sigmoid 기반 퍼지 클러스터링을 사용해 추출된 영상에서 유방 종양을 분할한다.

(Breast cancer detection by leveraging Machine Learning, Anji Reddy Vaka, 2020)

미리 훈련된 DCNN(Deep Convolution Neural Networks)을 사용하여 다중 네트워크 피쳐 추출 모델을 구현하고, 효과적인 피쳐 치수 감소 방법을 개발하며, 앙상블 서포트 벡터 머신(E-SVM)을 교육한다. 첫째, 스케일 변환 및 색상 향상 방법을 통해 기록영상을 사전 처리한다. 둘째, 멀티 네트워크 기능은 사전 교육된 DCNN 4개(예: DensityNet121, ResNet-50, 멀티 레벨 In 셉션 V3, 멀티 레벨 VGG-16)를 사용하여 추출된다. 셋째, 성능 향상 및 과충격 완화를 위해 DOL(Dual Network Orthogonal Low-Rank Learning)이 추가로 개발한다. 마지막으로, E-SVM은 퓨전된 특징과 투표 전략을 통해 교육을 받고 분류 작업을 수행한다. 분류 작업은 이미지를 네 가지 등급(예: 양성, 상황암, 침습성 암 및 정상)으로 분류한다. 제안된 방법을 공개 ICIAR 2018 Challenge 데이터 집합에서 유방암의 역사학 영상에 대해 평가하여 97.70%의 높은 분류 정확도를 달성한다. (Breast

Cancer Image Classification via MultiNetwork Features and Dual-Network Orthogonal Low-Rank Learning, YONGJUN WANG, 2019)

A Review on Gynecomastia and Male Breast Cancer for Radiologists, Eun Jin Hwang, MD and Joo Hee Cha, 2020 Sep;81(5):1096-1108. Korean. Published online Sep 01, 2020. 저에코의 불규칙한 모양과 침상형, 각진 형이나 미세소엽형 경계를 보이는 고형종괴로 유두 한쪽에 치우쳐 보인다.

Yuan 등(16)이 시행한 125개(양성 100개, 악성 25개)의 남성 유방 종괴의 수술 전 초음파 소견에 대한 후향적 분석에서 유두 한쪽에 치우친 종괴, 불규칙형 모양, 에코성 테두리, 종괴 내 혈류가 보이는 소견이 남성 유방암의 유의한 소견이었다. 그러나 초음파에서 경계가 잘 그려지는 병변이라도 남성에서는 소엽에서 기원하는 낭종이나섬유선종과 같은 BI-RADS 범주 2 나 3 병변이 드물기 때문에 유방암의 가능성이 있으므로 철저하게 검사해야 한다고 한다.

2.2 유방암 조직검사 병리학적 해석 진단 방법

국립암센터의 자료에 따르면 진단을 위한 검사로는 조직병리검사, 세포병검사, 조직화학검사 면역병리검사 등이 있다고 한다. 그 중에서도 본 종합설계 논문에서는 조직병리검사로 얻은 이미지를 사용했다. 우선 종양과 암의 구분 요인으로는 암세포의 형태학적이상, 세포간의 관계이상, 세포들이 질서정연하지 않다는 것으로 확인할 수 있다고 한다. 또한 종양의 크기, 주위 조직,장기 침윤, 혈관 림프관 침윤 등을 통해 확인할 수 있다고 한다. (홍은경, 2013)



Figure 3. 암의 분화도는 암세포의 구조와 기능이 정상세포와 얼마나 비슷한지를 나타내는 지표로 위의 사진에서 G1은 고분화(정상세포와 비슷한 상태), G2(중등도 분화), G3(저분화, 정상세포와 비슷한 점이 거의 없는 상태), G4(미분화, 어떤 세포에서 기원했는지 구별이 불가능한 상태)이다. 분화도가 낮아 단계가 높을수록 암이 더 공격적인 양상을 나타낸다고 한다. (National Cancer Institute, n.d.)

또한 Johns Hopkins pathology 에 따르면 유방암 등급을 결정하는데 사용될 수 있는 Nottingham Histologic Score 시스템 ("the Elston-Ellis modification of Scarff-Bloom-Richardson grading system")은 병리학자가 고려하는 3 가지 요인을 정리했는데 gland formation 의 양(세포 분화도와 종양세포가 정상 샘을 재생하는 정도), 핵특징(다형성 및 종양 세포의 고르지 못한 모양), 종양 세포가 분열하거나 증식하는 정도로 나눈다고 정리했다. (Johns Hopkins pathology, n.d.)

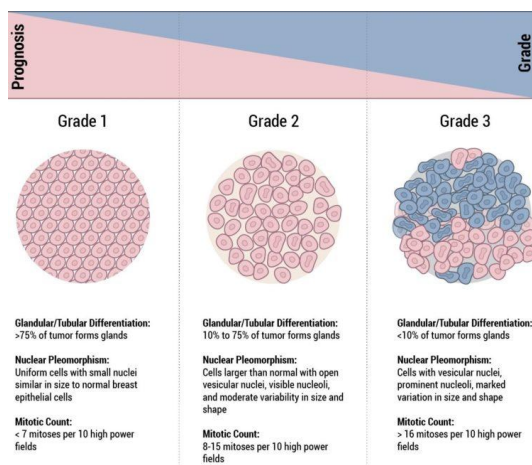


Figure 4 유방암 등급에 따른 병리학적 세포모양 특징(Johns Hopkins pathology, n.d.)

또한 등급에 따른 조직검사 이미지를 다음과 같이 분석했다.

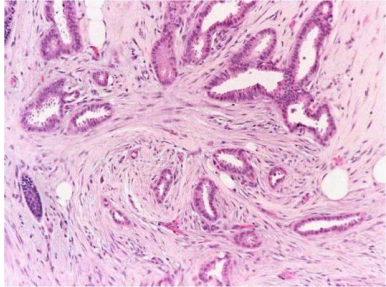
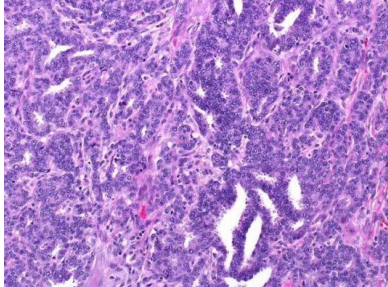
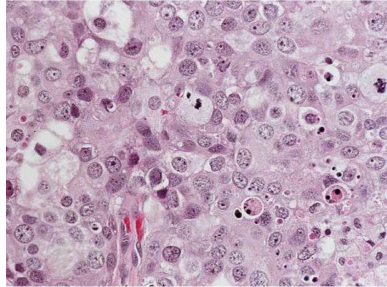
1 급	2 급	3 급
		
Histologic Grade I Invasive Ductal Carcinoma. 침윤성 유관 암종은 균일한 핵을 가진 각진 nuclei 로 구성되어 있다고 한다.	tubular gland structures 를 형성하고 있지만 일부 영역이 비정형 핵을 포함하는 nests of cells 이 잘 형성되지 않은 모습을 보인다고 한다.	가장 심각한 암 등급의 경우에는 뚜렷한 핵 이형성과 유사 분열을 가진 nests of cells(nests cells with marked nuclear atypia and mitotic activity) 특징을 보인다고 한다.

Table 2 유방암 등급에 따른 조직검사 이미지에서 병리학적 특징.

따라서 본 종합설계 논문에서도 이러한 병리학적 형태 특징을 강화할 수 있는 방법을 통해 진단 정확도를 높이하고자 한다.

3 Research Methodology

3.1 데이터

데이터는 BreakHis(Breast Cancer Classification Challenge 2015 데이터 세트) 로 데이터는 40X, 100X, 200X, 400X 현미경 배율별로 8가지 진단결과를 가진 악성과 양성 데이터로 분류되어 있으며 진단결과에 대해 조사한 결과이다.

1) malignant 악성

- 유관상피내암(Ductal carcinoma in-Situ, DCIS)

도관(duct)에서 시작된 암으로, 유관상피내암은 유관을 이루는 세포에서 생긴 암으로 유관의 기저막을 침범하지 않고 유관 안에 있는 제 0 기 암이라고도 한다.

폴더명: ductal_carcinoma

- 소엽암(lobular carcinoma)이라고 부릅니다. 엽(lobe)에서 시작된 암. 암이 되기

전(pre-cancerous) 세포가 유방 세포조직 안에서 발견될 수 있다.

폴더명: lobular_carcinoma

- 유방 점액성 암(Mucinous Carcinoma)

유두 유방암 은 모든 유방암의 1 % 미만을 차지하는 매우 드문 유형의 침습성 유관 유방암입니다. 이 이름은 현미경으로 세포를 볼 때 보이는 손가락 모양의 돌기 또는 구진에서 유래합니다.

폴더명: mucinous_carcinoma

- papillary_carcinoma

유방의 유두상 암종(papillary carcinoma)은 여성 유방 암에서 약 0.3-2%를 차지하는 비교적 드문 종양이며 유두암종(papillary carcinoma)과는 방사선 검사, 세포 검사, (핵심)총생검 조직검사로는 구별이 어렵다.

폴더명: papillary_carcinoma

2) 양성 benign

- 선증(adenosis)은 유방 질환에서 흔한 조직학적 소견이지만 드물게 주변의

종말관소엽단위(terminal duct/lobular unit: TDLU)와 합쳐져 임상적으로 만져지거나 육안으로 인지되는 선증종양(adenosis tumor)을 형성한다.

폴더명: adenosis

- fibroadenoma 유방 섬유선종은 유방에서 발생하는 종양 중 가장 흔한 양성종양으로 정확한 원인은 밝혀지지 않았지만 젖을 분비하는 유선 말단부위의 과다한 증식과 이에 따른 유방조직의 변형 때문에 생긴다고 알려져 있다.

폴더명: fibroadenoma

- 엽상 종양(phyllodes tumor)이라는 말을 사용하고 암인 경우를 악성 엽상종양, 암이 아닌 경우를 양성 엽상종양이라고 부른다.

폴더명: phyllodes_tumor

- 관상선종 tubular_adenoma

크기가 1cm 이상의 선종, 현미경 소견에서 용모 형태(villous adenoma)를 보이는 선종, 세포가 덜 분화된 선종은 암으로 발전할 가능성이 높은 선종이다. 흔하지만, 악성화 경향은 낮은 양성 종양이다.

폴더명: tubular_adenoma

3.2. 연구 방법

3.2.1 CNN architecture 에 따른 유방암 진단 정확도 비교

첫번째로, 아래 Figure 와 같이 층의 깊이에 따른 모델을 구축하여 층의 깊이에 따른 정확도를 비교해보았다. 데이터셋은 40 배율로 진행하였고 1855 이미지와 8 개의 클래스로 학습시켰다. epoch 는 30 으로 진행하였고 optimizer 는 SGD 로, 학습률은 0.0001, momentum 은 0.9, loss 함수는 'categorical_crossentropy'로 진행하였으며 metrics 는 정확도로 비교하였다. ImageDataGenerator 를 이용해 이미지 정규화를 진행하였으며 배치사이즈는 16 으로 학습시켰다.

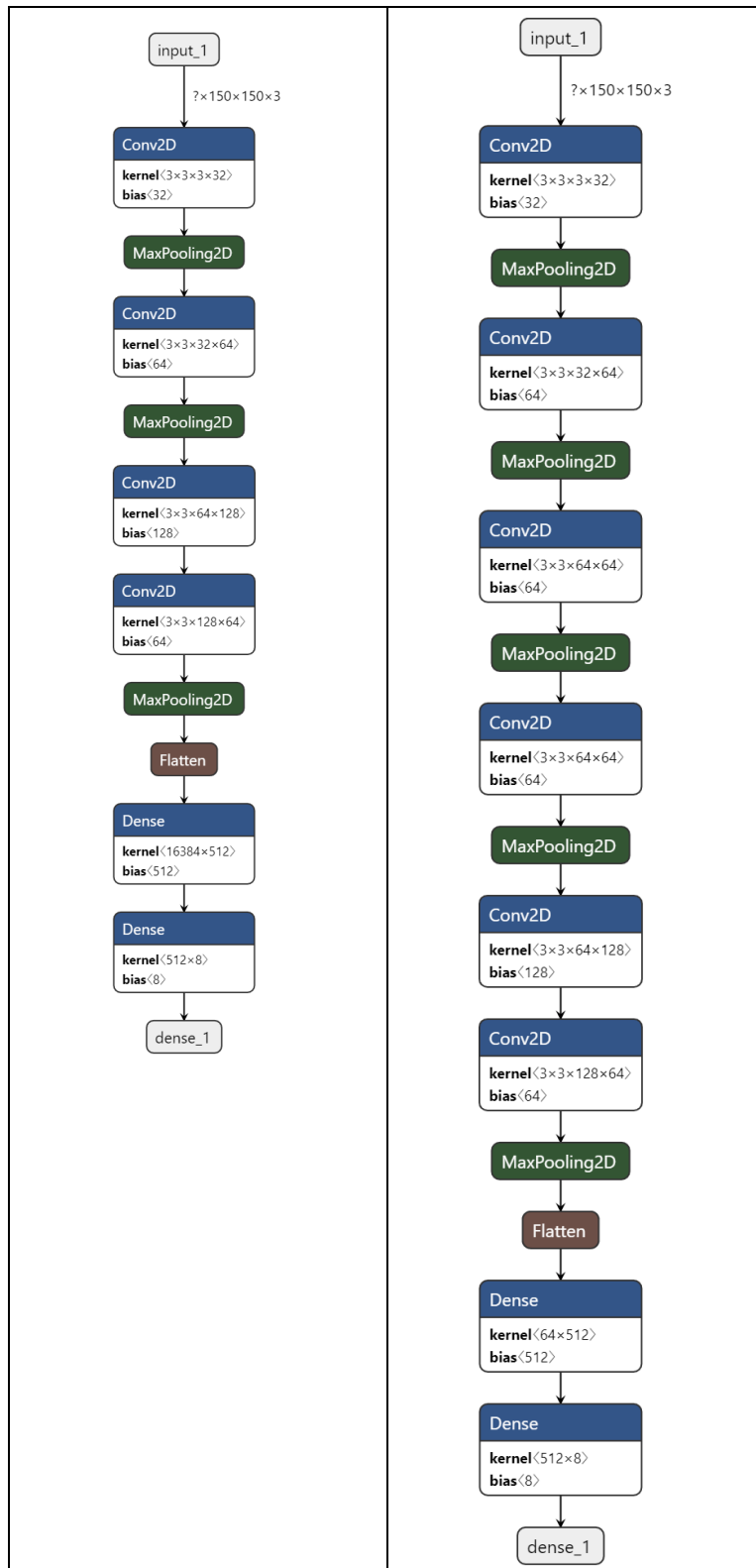


Figure 5. 층에 따른 정확도 비교를 위한 모델 구축.

두번째로, 적합한 CNN architecture 를 비교하기 위해 VGG 와 유사한 모델과 Inception 모듈을 추가한 GoogLeNet 으로 모델을 비교해보았다. 세번째로는 두번째에서 적합한 모델이 인셉션

모듈을 추가한 모델임을 확인한 후 배율에 따라서 정확도를 비교해보았다.

3.2.2 조직병리학적 진단 요인을 고려한 이미지 특징강화 방법 적용에 따른 정확도 비교

1) 임계처리를 통한 세포핵 경계 임계 특징 강화

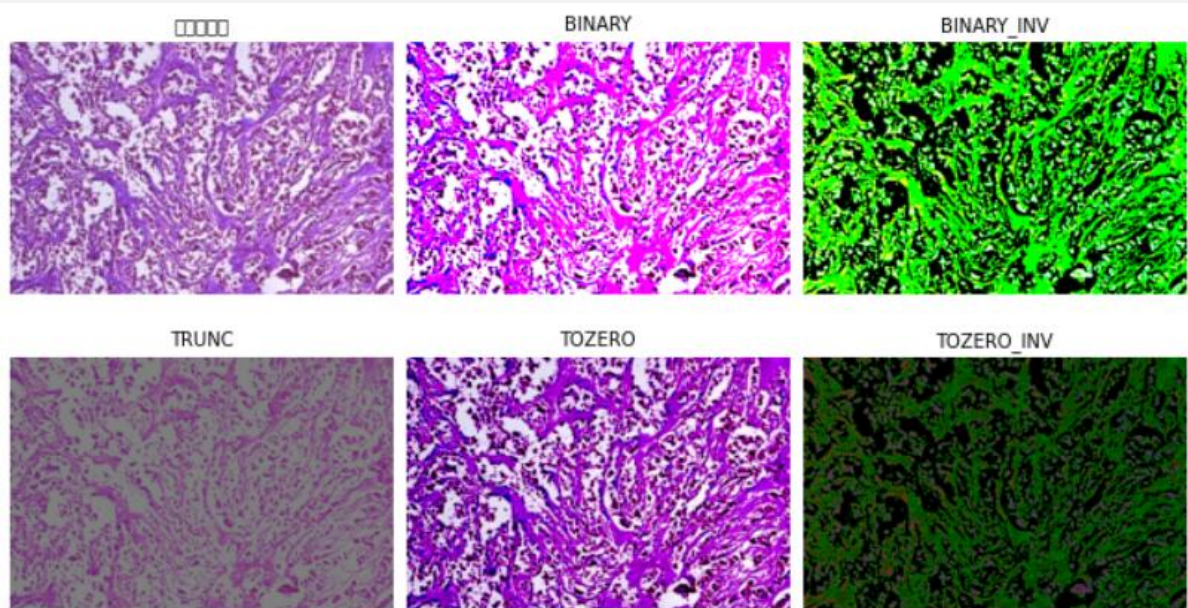
유방암 이미지 데이터들의 픽셀값을 중간값 threshold 을 사용하여 이진화(binanzation)하는 필터를 사용해서 유방암 세포핵의 불분명한 경계로 인한 과적합을 줄이고 암세포가 가지고 있는 불분명한 세포 rest 의 특징을 강화하여 정확도를 높이고 과적합을 줄일 수 있는 방법을 고안하여 적용하였다.

- 기준값을 넘으면 최대값 아니면 0
- 기준값을 넘으면 0 아니면 최대값
- 기준값을 넘으면 기준값 아니면 최대값
- 기준값을 넘으면 원래값 아니면 0
- 기준값을 넘으면 0 아니면 원래값

위의 총 5 가지의 규칙으로 임계처리를 해보았을 때 다음 사진과 같이 나왔다.

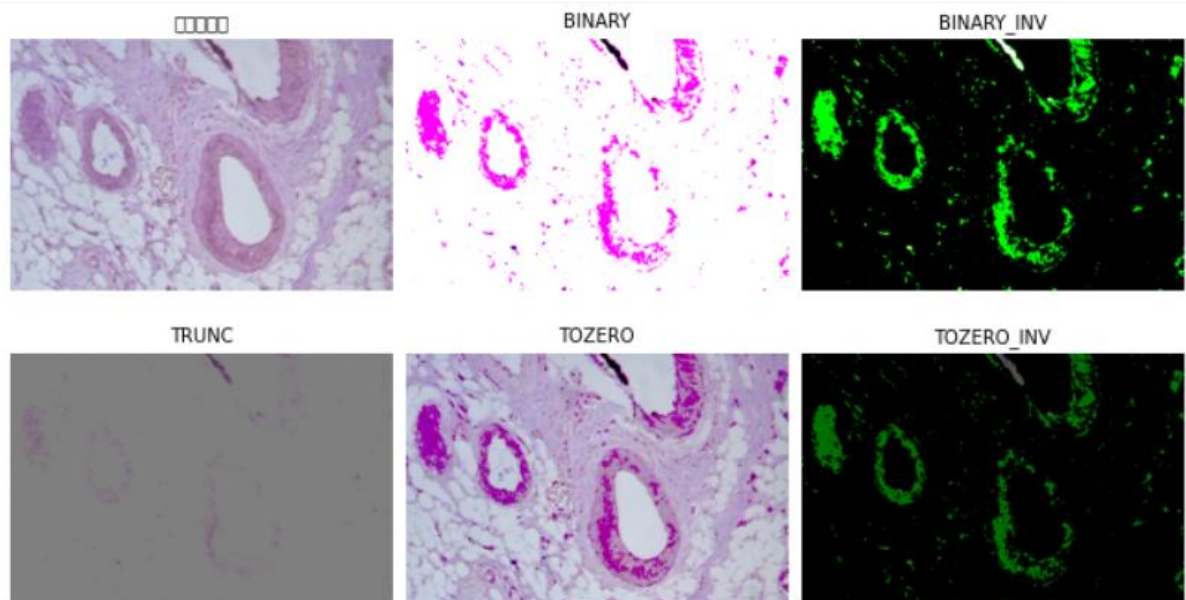
1) 악성 40 배율

ductal_carcinoma



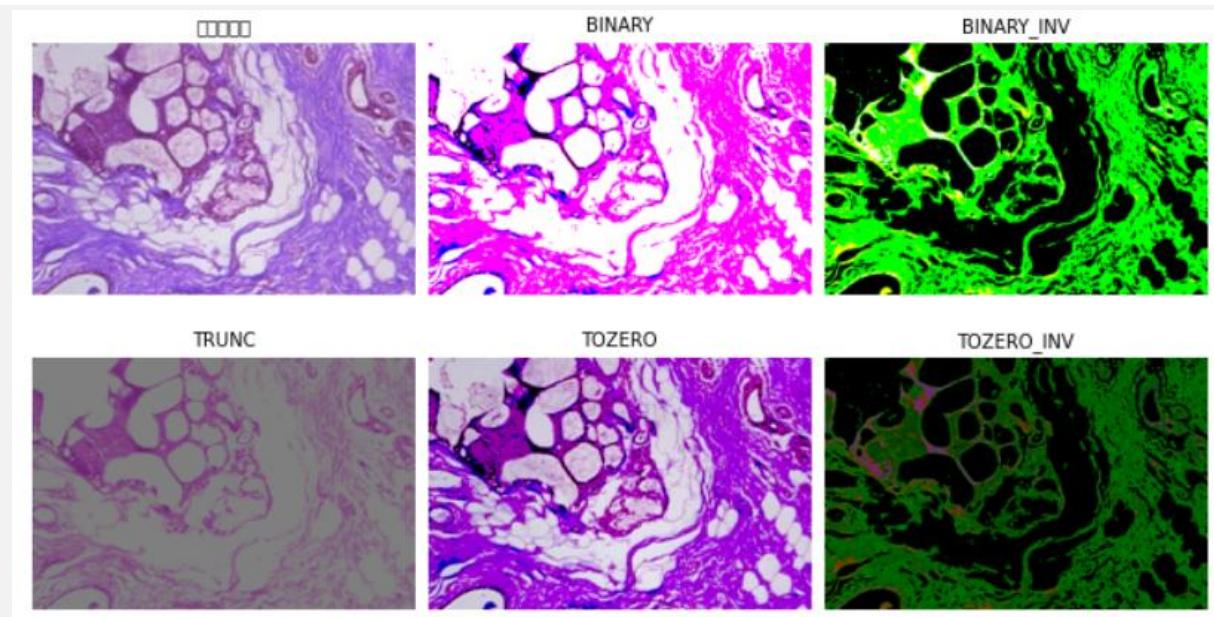
위의 사진과 같이 40 배율에서는 악성의 세포핵 경계의 질서정연하지 못한 모습을 더 부각시킬 수 있었고 그 경계의 임계값을 바이너리와 TOZERO 방식이 명확하게 잘 구분할 수 있는 것을 확인 할 수 있었다.

lobular_carcinoma

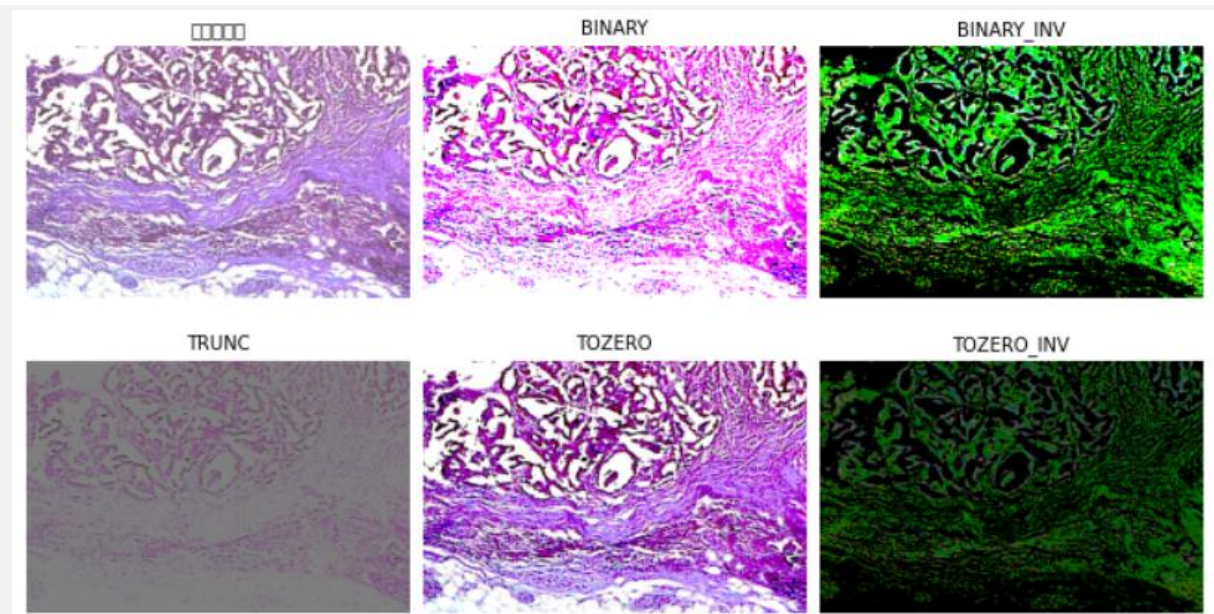


악성의 한가지 특징 중에 세포경계 막이 두꺼워지는 현상을 잘 나타낼 수 있는 Binary, binary_inv, tozero 를 확인할 수 있었다.

mucinous_carcinoma



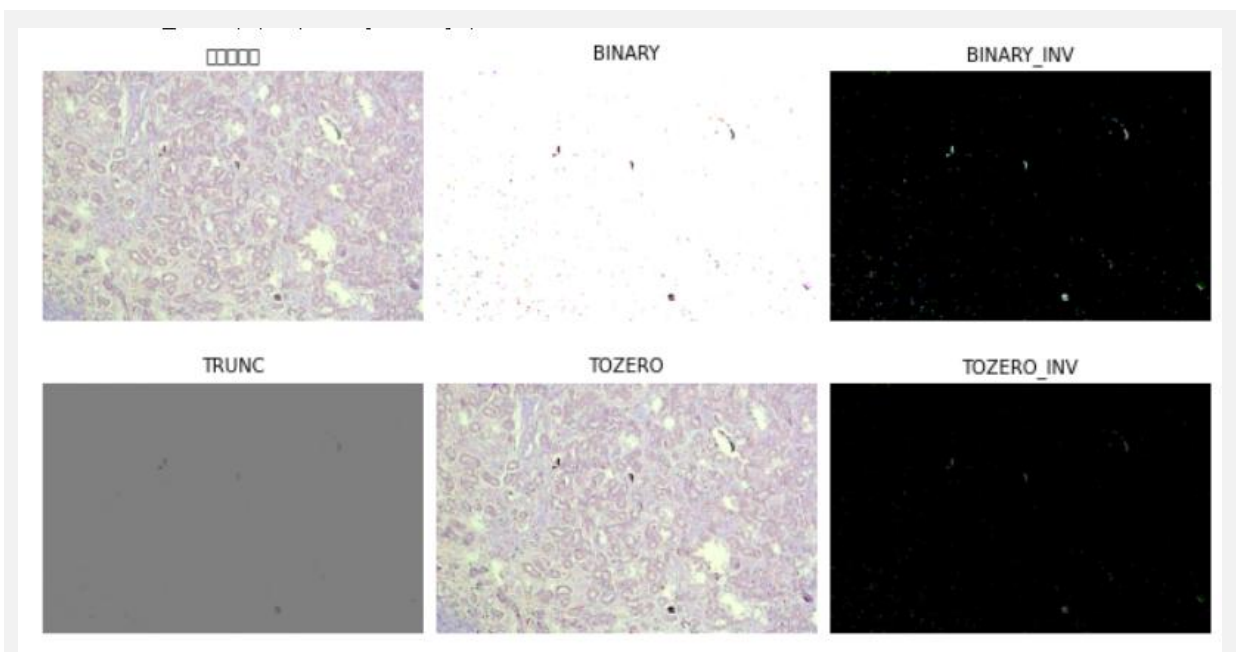
papillary_carcinoma



경계의 차이를 뚜렷하게 만들어 주는 것으로 노이즈를 줄이는데 활용될 수 있음을 보여주었다.

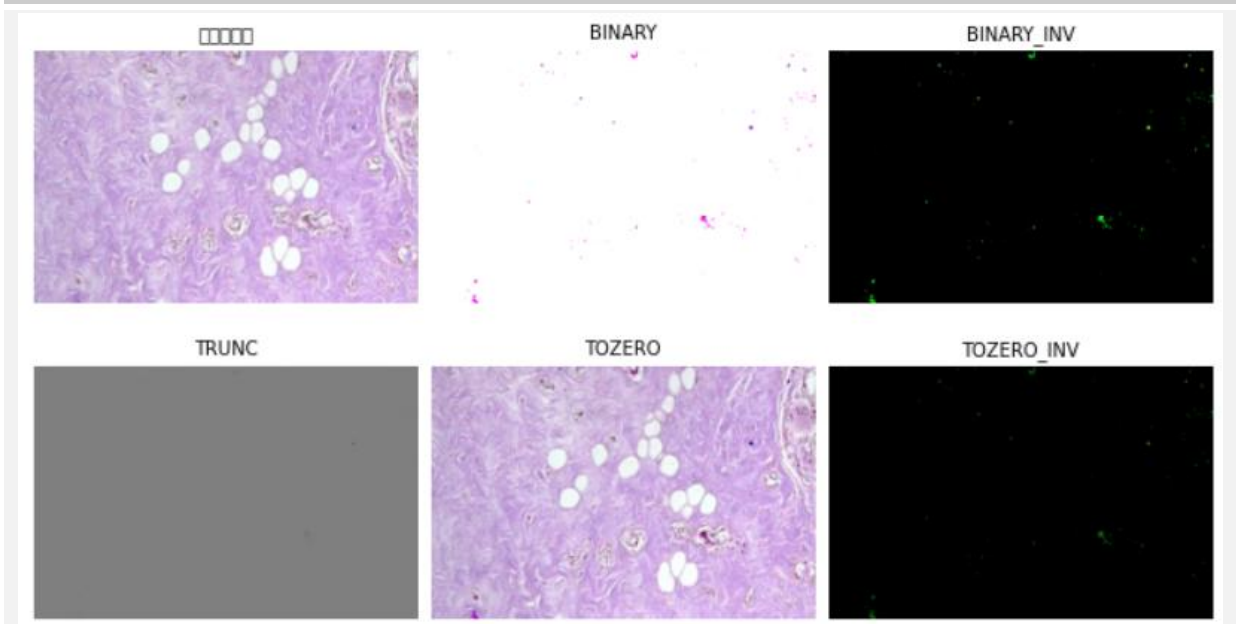
2) 양성 40 배율

adenosis

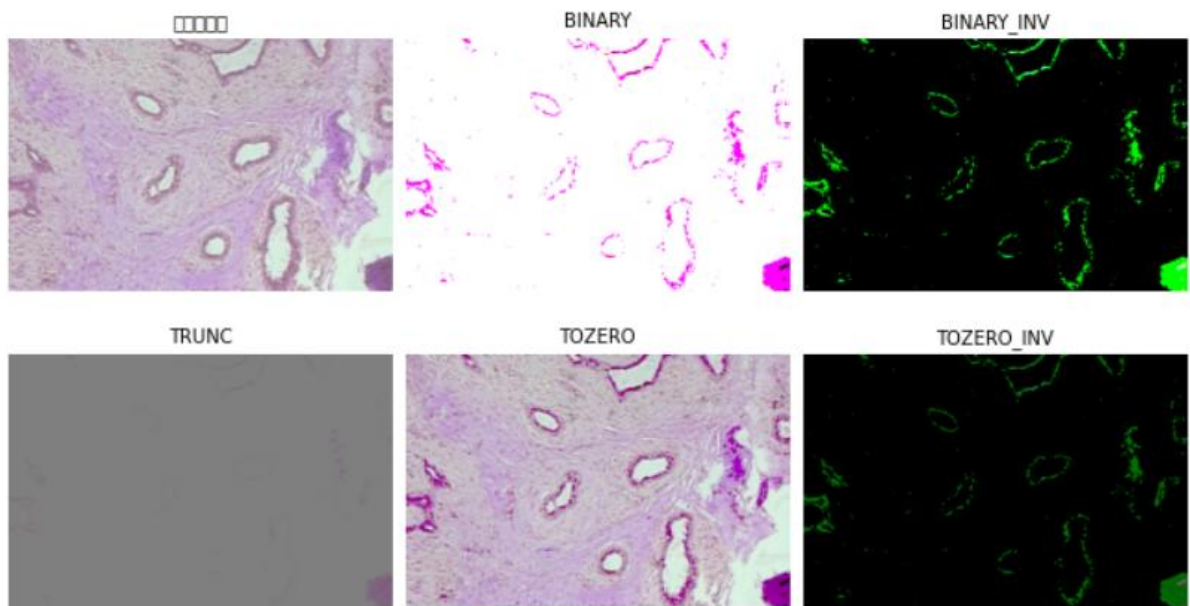


염색이 잘 되지 않아 명도가 낮을 경우에 문제점이 생기는 것을 확인할 수 있었다.

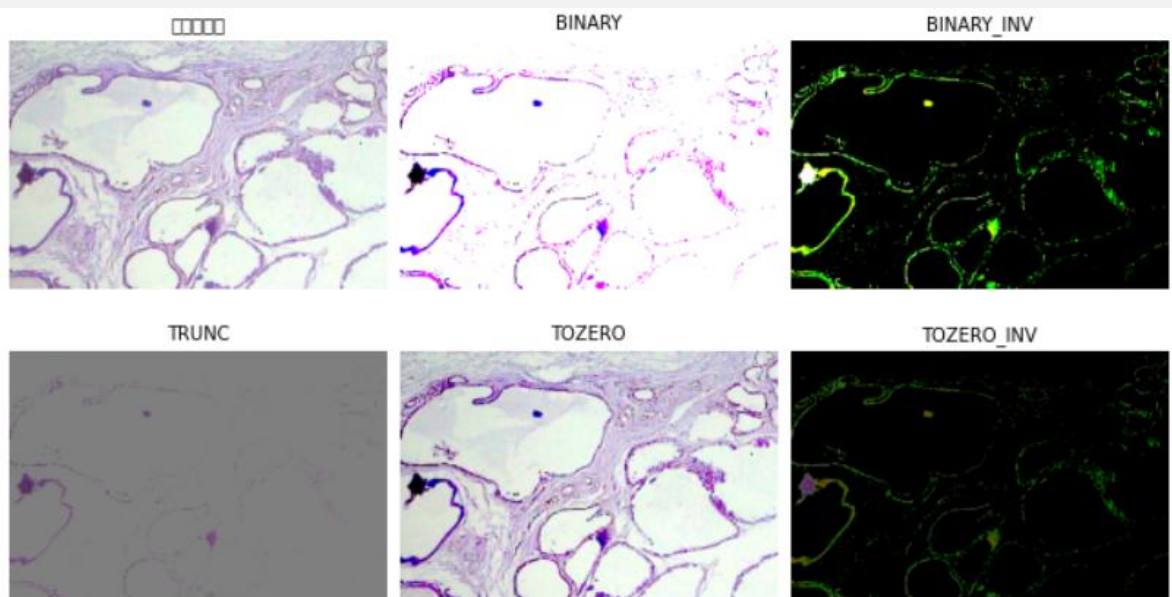
fibroadenoma



phyllodes_tumor

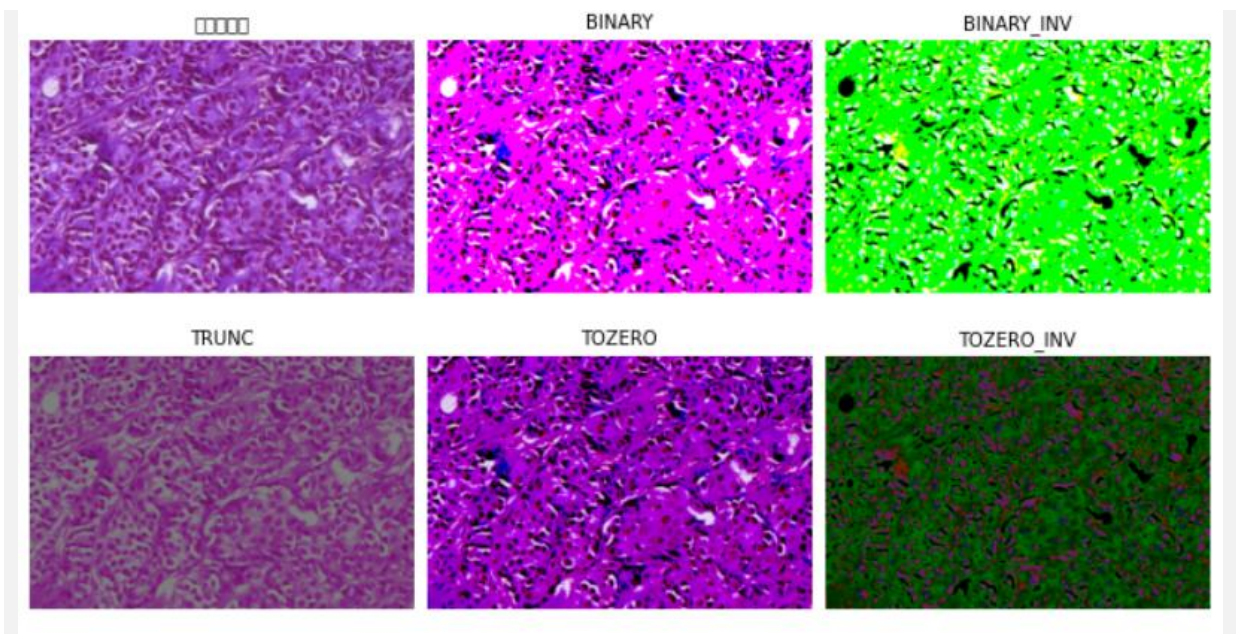


tubular_adenoma

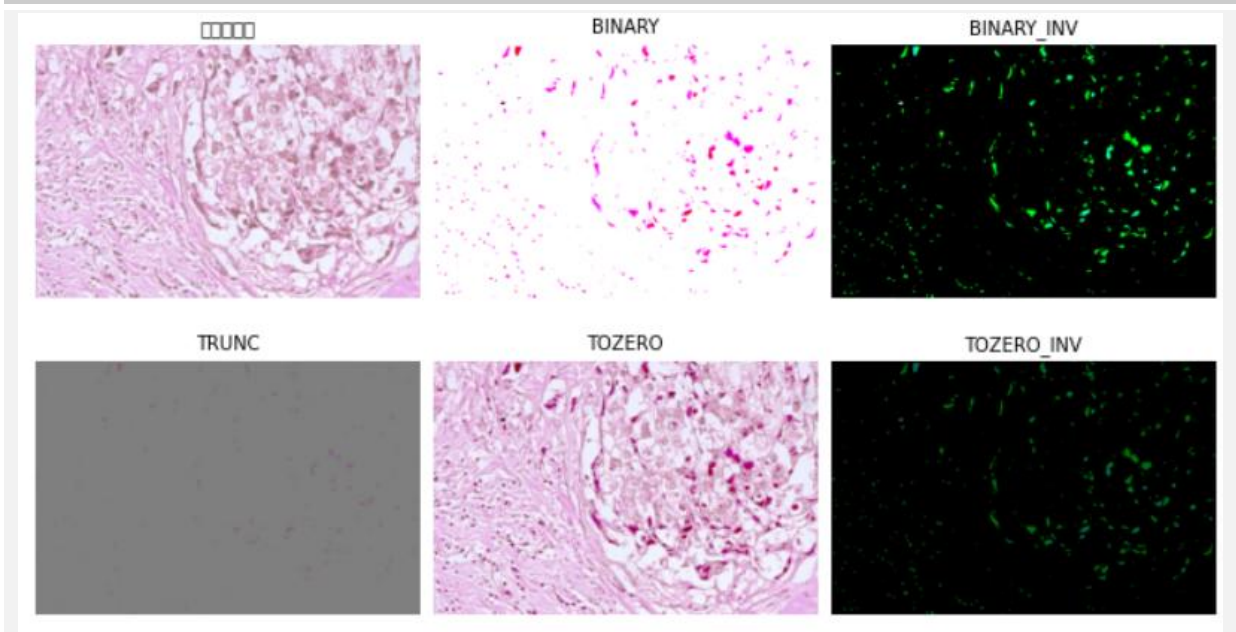


3) 악성 100 배율

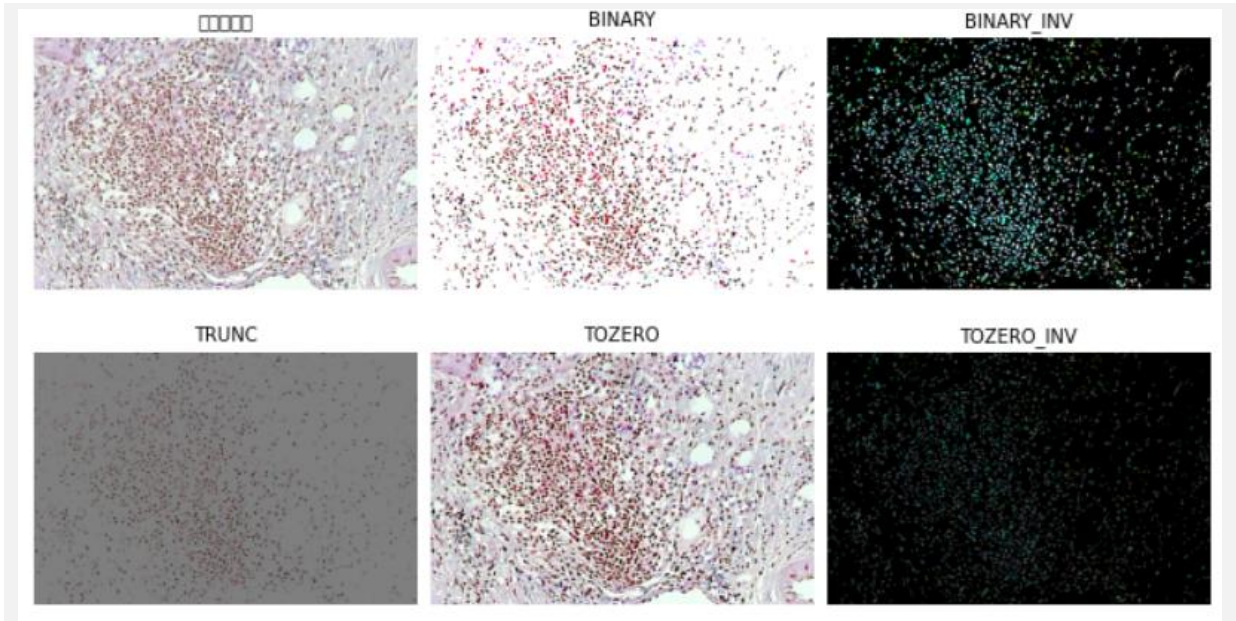
ductal_carcinoma



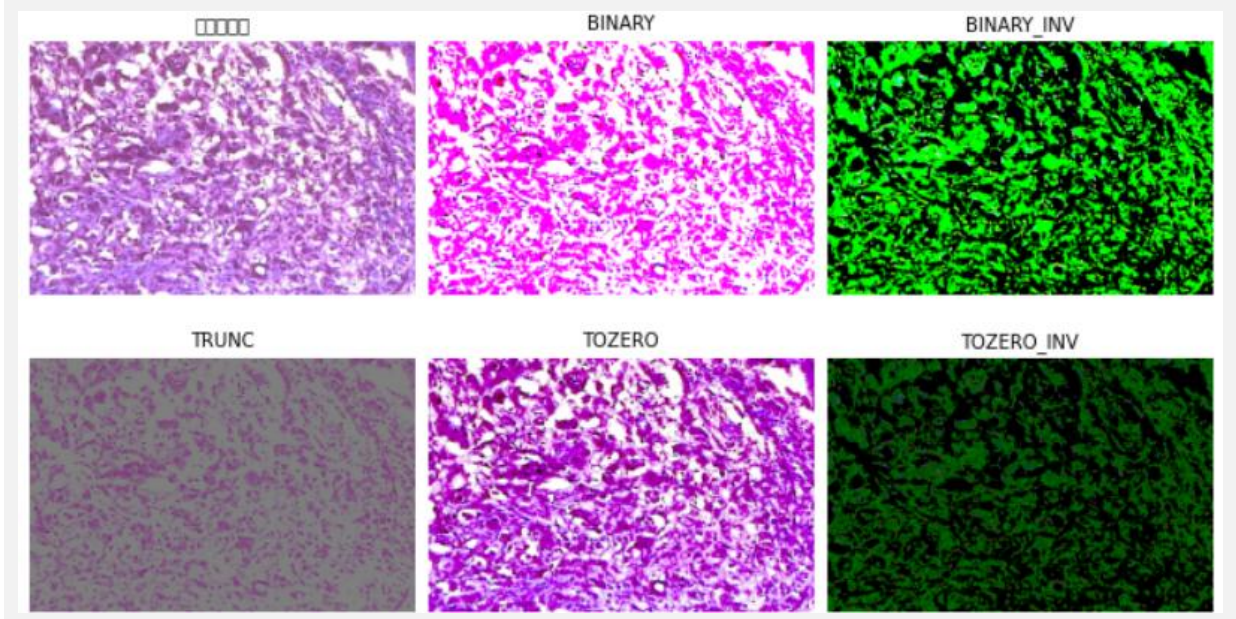
lobular_carcinoma



mucinous_carcinoma

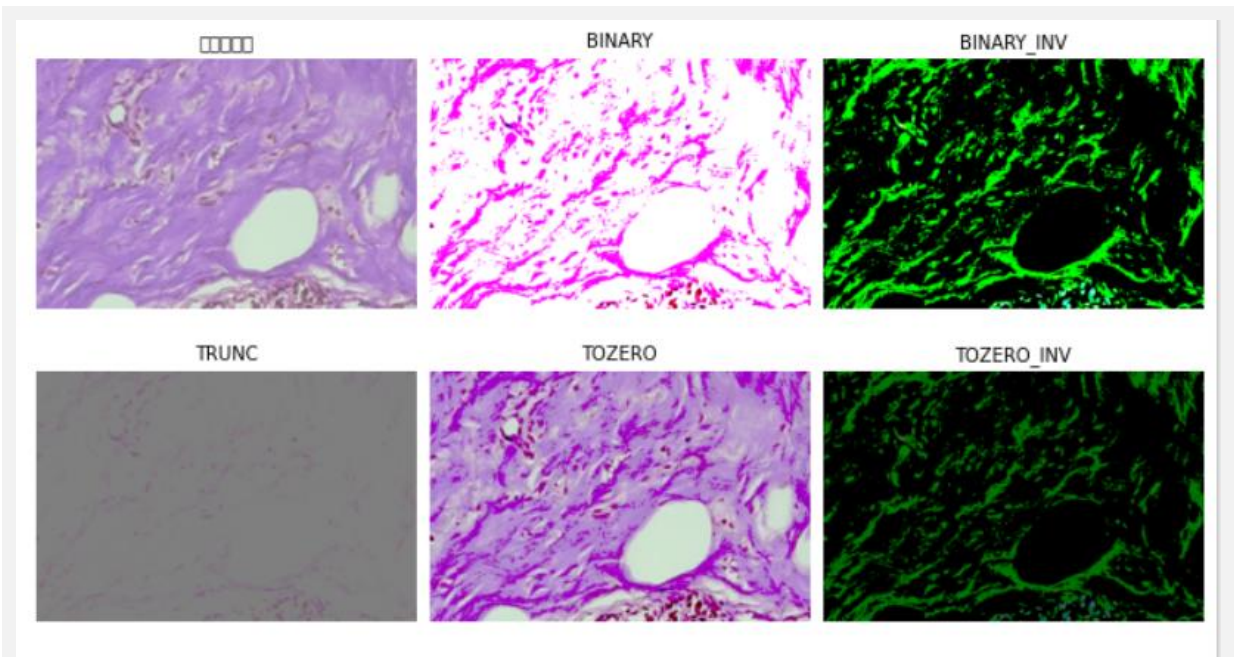


papillary_carcinoma

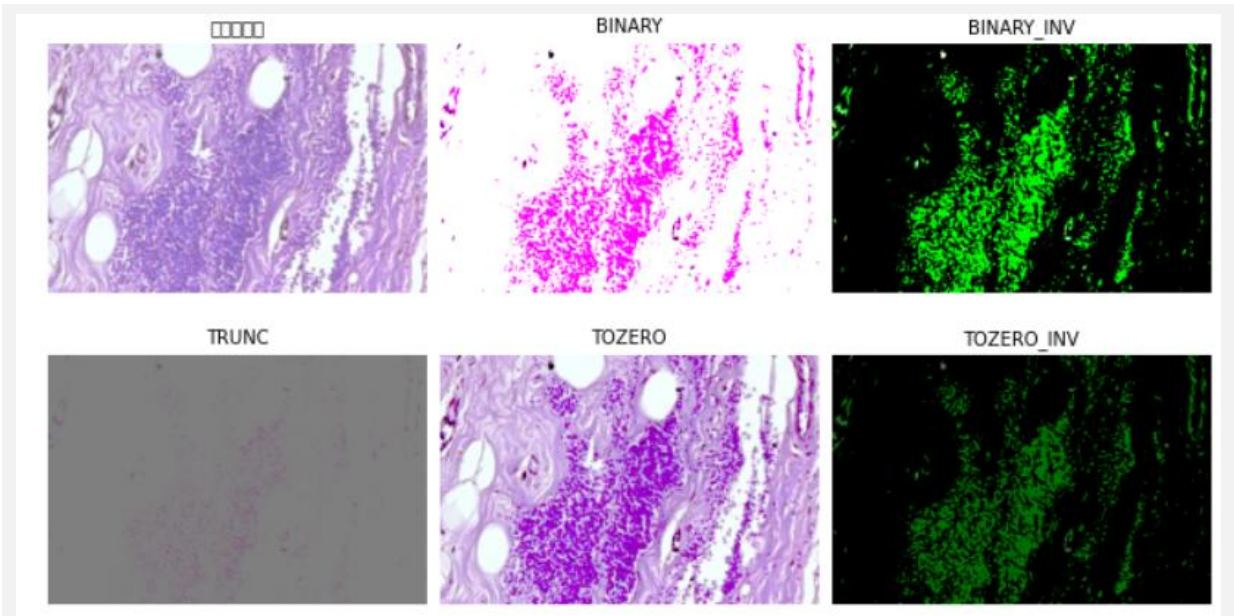


4) 양성 100 배율

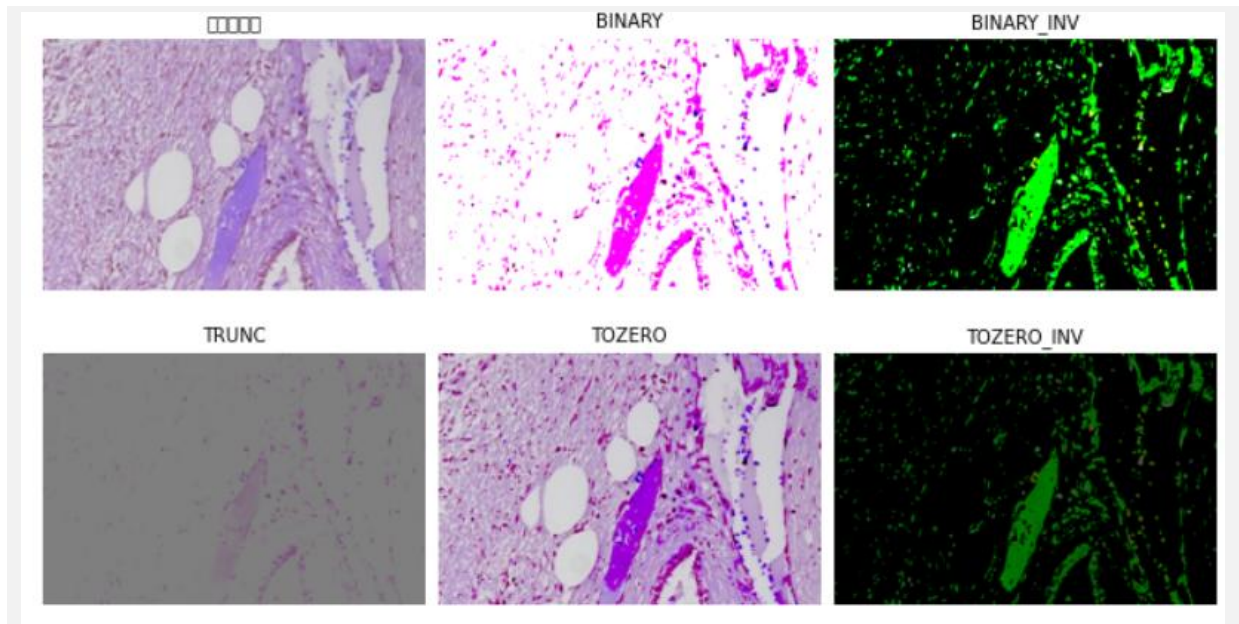
adenosis



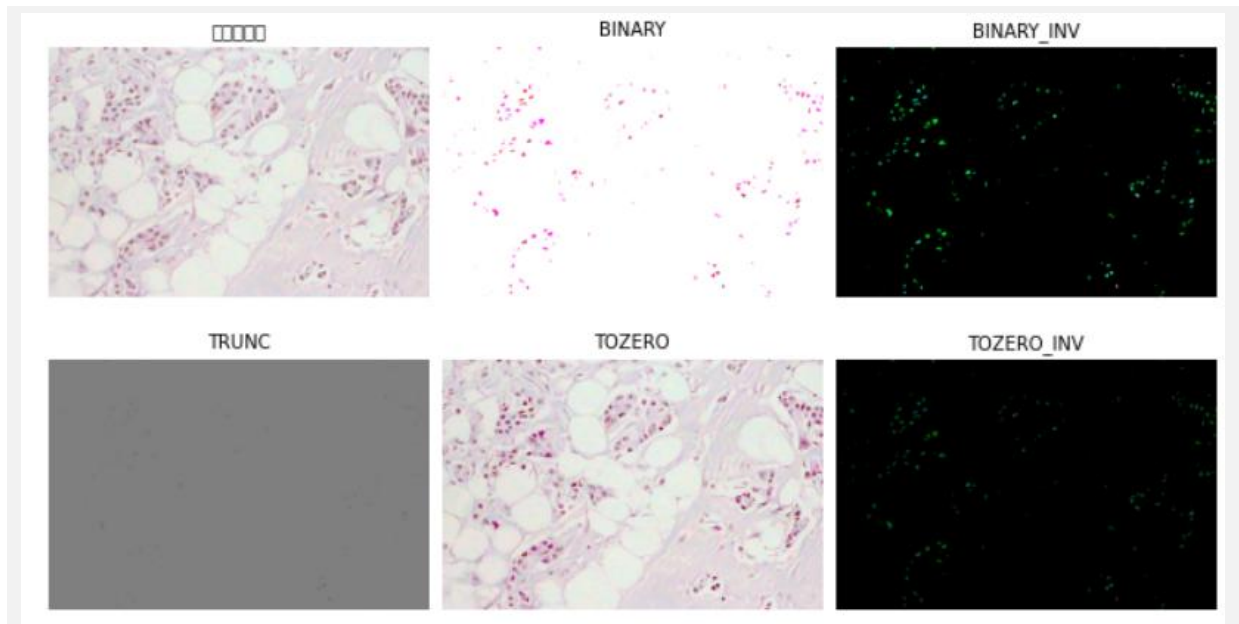
fibroadenoma



phyllodes_tumor



tubular_adenoma

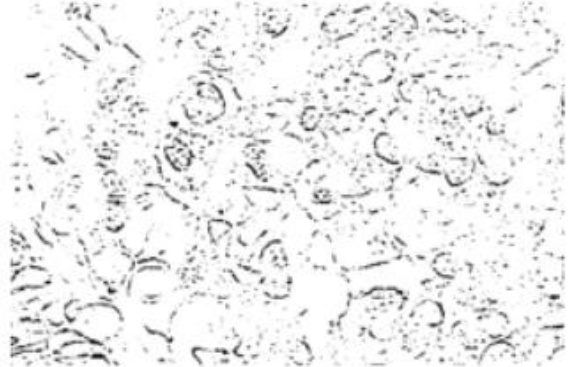
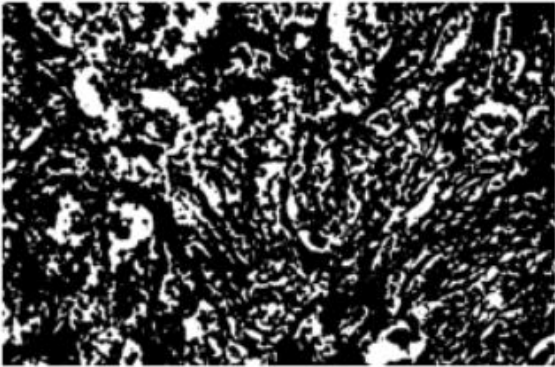


2. 적응 임계처리를 통한 특징강화

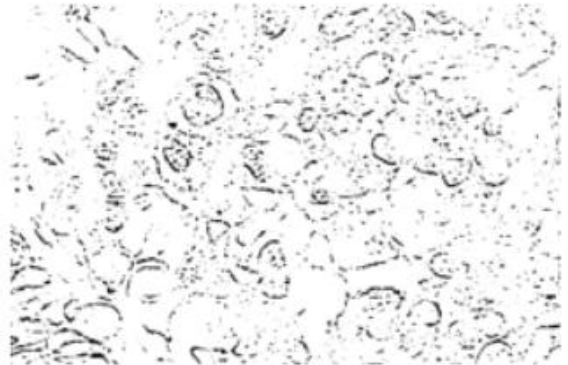
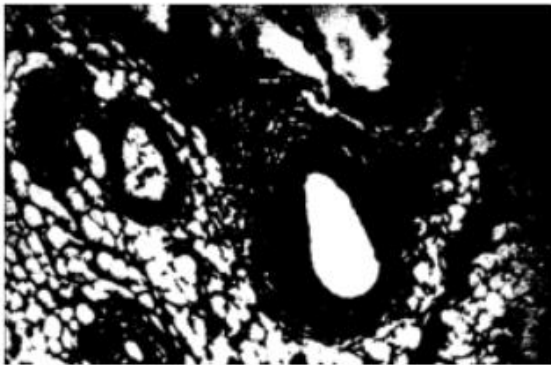
임계처리의 경우 이미지 전체에 하나의 기준값을 적용해서 일정한 영역내의 근접해있는 픽셀값들을 고려하지 못한다는 점을 확인할 수 있었다. 염색에 따라서 위에서 특징이 오히려 부각이 안되는 경우가 생겨서 본 논문에서는 근접해있는 픽셀값들을 이용해서 해당 영역에 적용할 기준값을 자체적으로 계산하는 적응임계처리 방식도 사용하기로 하였다. 아래 사진은 왼쪽은 평균 적응 임계처리한 것이고 오른쪽은 가우시안블러 적응임계처리로 처리한 것이다.

1) 악성 40 배율

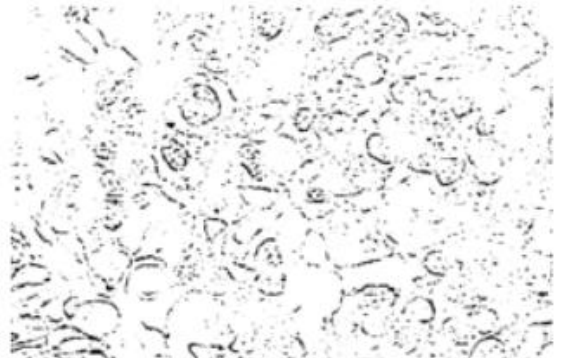
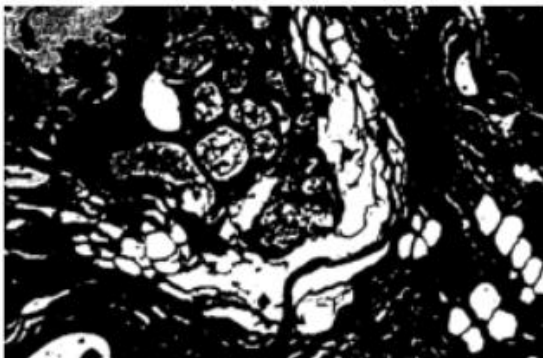
ductal_carcinoma



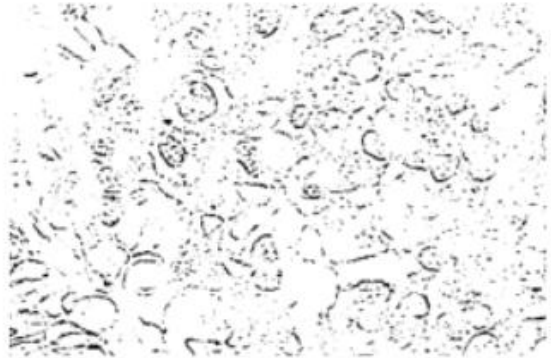
lobular_carcinoma



mucinous_carcinoma

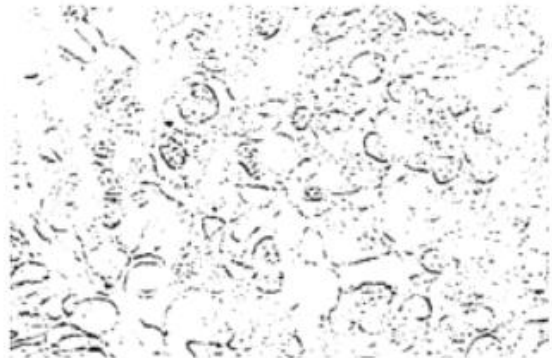
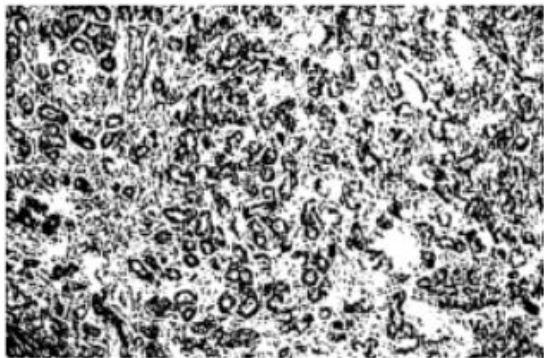


papillary_carcinoma

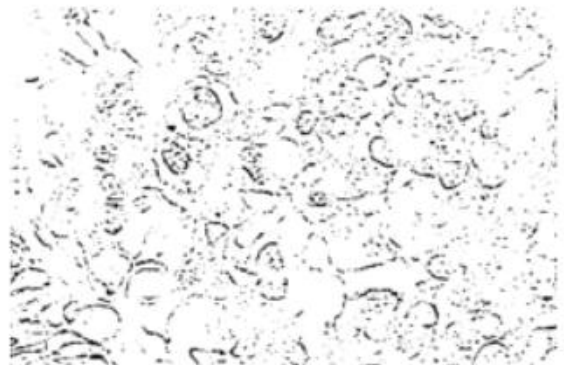
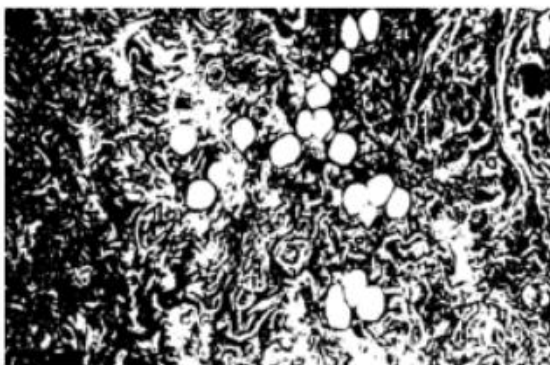


2) 양성 40 배율

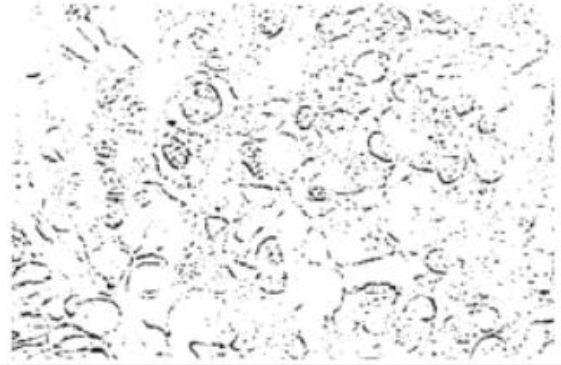
adenosis



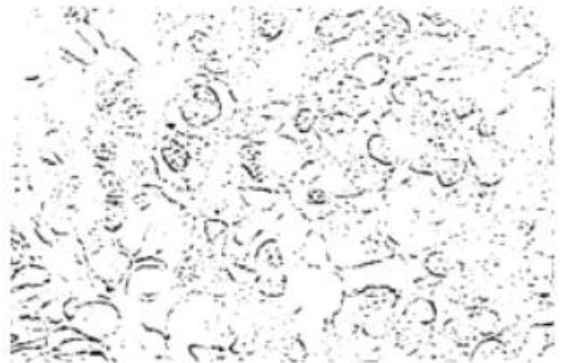
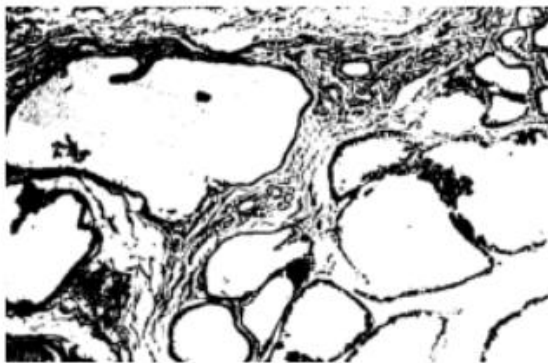
fibroadenoma



phyllodes_tumor

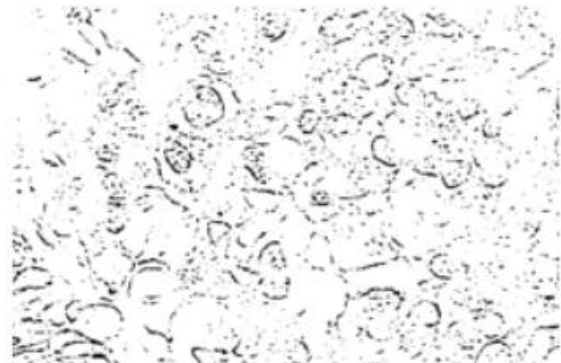


tubular_adenoma

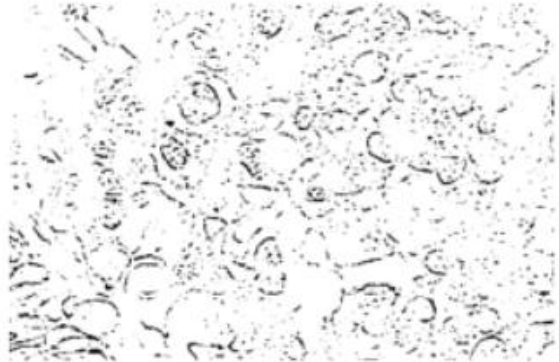


1) 악성 400 배율

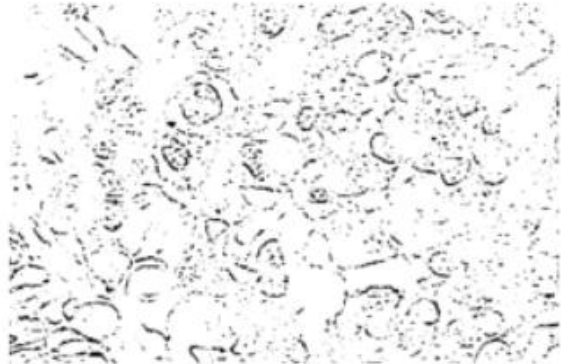
ductal_carcinoma



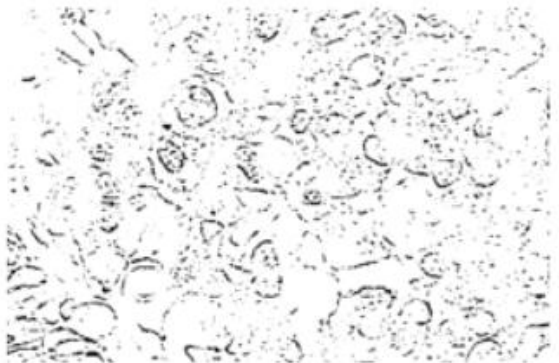
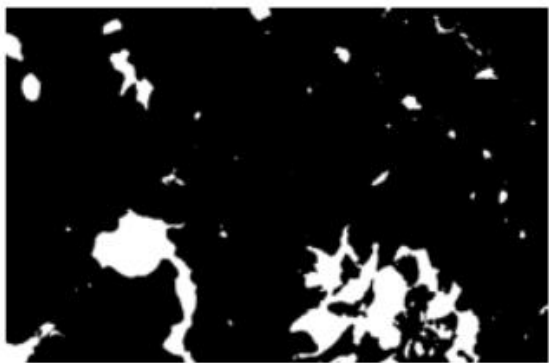
lobular_carcinoma



mucinous_carcinoma

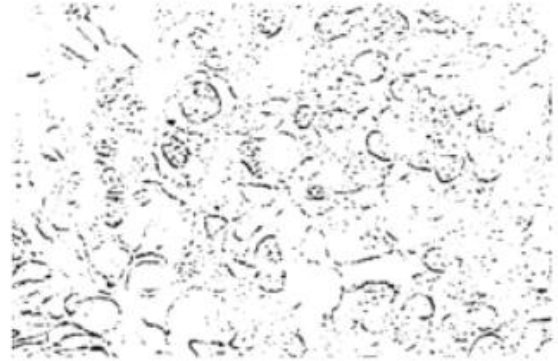


papillary_carcinoma

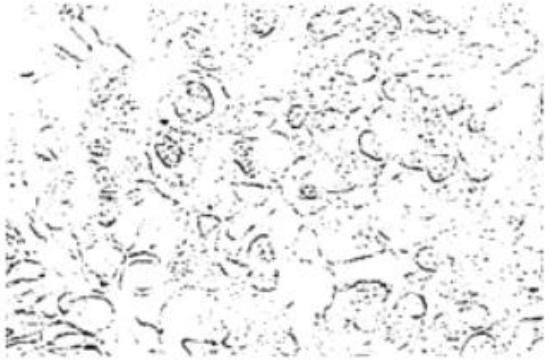
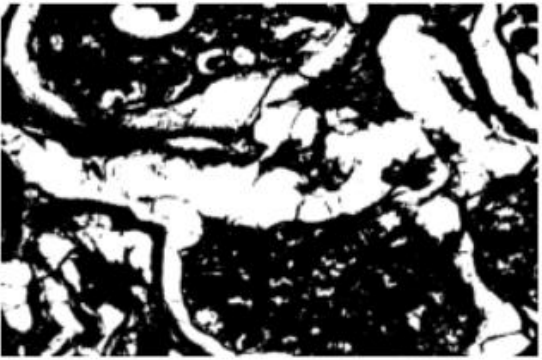


2) 양성 400 배율

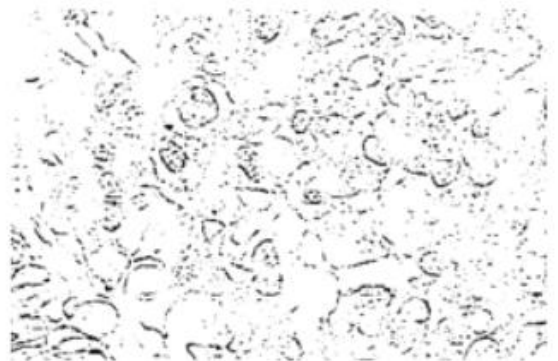
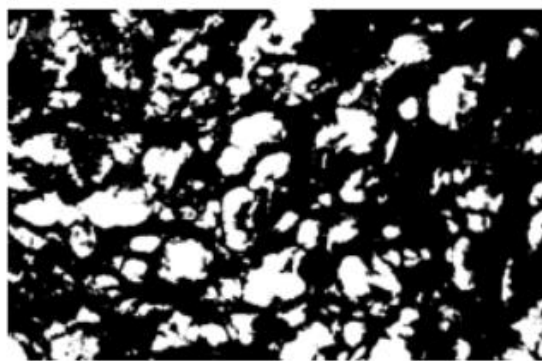
adenosis



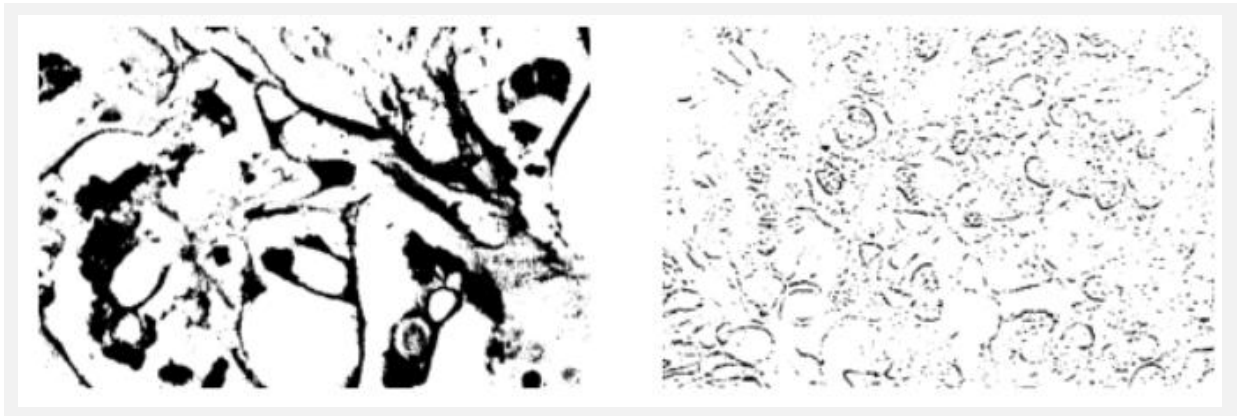
fibroadenoma



phyllodes_tumor



tubular_adenoma



3.2.3 CAM(class activation map)시각화를 통한 설명가능성 향상

암의 진단에 있어서 예측력만큼 중요한 것이 의사의 진단을 돕기 위해서는 설명력이 중요하다고 판단하여 class activation map 을 시각화해서 모델이 어느 부분에서 가중치가 높아서 올바른 예측을 했는지 판단하기 위해서 CAM 을 시각화 하는 방법을 사용했다.

3.2.4 Grad-cam 을 사용한 설명가능성 향상

layer 를 통과한 output 이미지에 convolutional layer 의 gradient 를 시각화하여 확인하여 보았다.

4. 연구 결과

4.1 CNN architecture 에 따른 유방암 진단 정확도 비교

1) 층의 깊이에 따른 비교실험 (epochs=30)

Figure5 의 첫번째 모델처럼 쌓았을 때는 0.4832 정확도, 두번째 모델은 0.4658 의 정확도를

확인할 수 있었다. 데이터셋이 40 배율 같은 경우에는 적은 픽셀이 세포 핵 이미지 부분이기

때문에 풀링(pooling)을 거친 뒤에는 feature map 의 개수가 2 배로 커지면서 필요한 파라미터가

과도하게 많아져, 파라미터가 많았을 때 생기는 고질적인 문제인 gradient vanishing,

overfitting 등의 문제가 발생하여 이러한 현상이 나타난 것이라고 본 저자들이 생각하였다. 그래서 이 문제를 해결하기 위해서 인셉션 모듈을 사용하여 비교해보았다.

2) 적합한 CNN architecture 비교실험

유방암 데이터셋 같은 경우는 층이 깊어질수록 작은 영역에 있는 세포핵들의 특징을 추출하는 것이 힘들어 과적합 문제가 생기는 것으로 본 저자들이 생각했는데 지역적인 receptive 한 field 에서 더 많은 피처를 추출하기 위해 여러 개의 convolution 을 1*1 conv, 3*3conv, 5*5conv, 풀링 레이어를 나란히 놓아서 다양한 scale feature 를 추출하도록 고안하여 적용하였다. feature 추출 등의 과정은 최대한 sparse 하게 유지하고 행렬 연산 자체를 이들을 합쳐서 최대한 dense 하게 만들었다는 점에서 앞의 간단한 모델보다 훨씬 좋은 성능을 나타낼 수 있었다. Keras 및 tensorflow 패키지를 사용했고 ImageNet 데이터 세트에서 훈련된 아키텍처 (VGG, Inception, ResNet 등)를 갖춘 cnn 을 사용해 전이학습을 시켰다. 유방암 데이터셋이 기존 이미지넷과 많이 다른 이미지들이기 때문에 수렴속도가 빨라지는 것이 가능할 지 확인해보았고 사전 훈련된 모델을 사용하여 이미 학습된 가중치를 사용하고 상단에 몇 개의 레이어를 추가하여 모델을 유방암 데이터로 미세 조정한 결과 수렴 속도가 빨라지고 처음부터 학습한 모델과 비교할 때 시간과 계산이 절약된 것을 확인할 수 있었다. inception 모듈을 추가한 GoogLeNet 을 활용한 모델을 만들어 확인해보았다. 아래 Figure 는 구축한 모델이다.

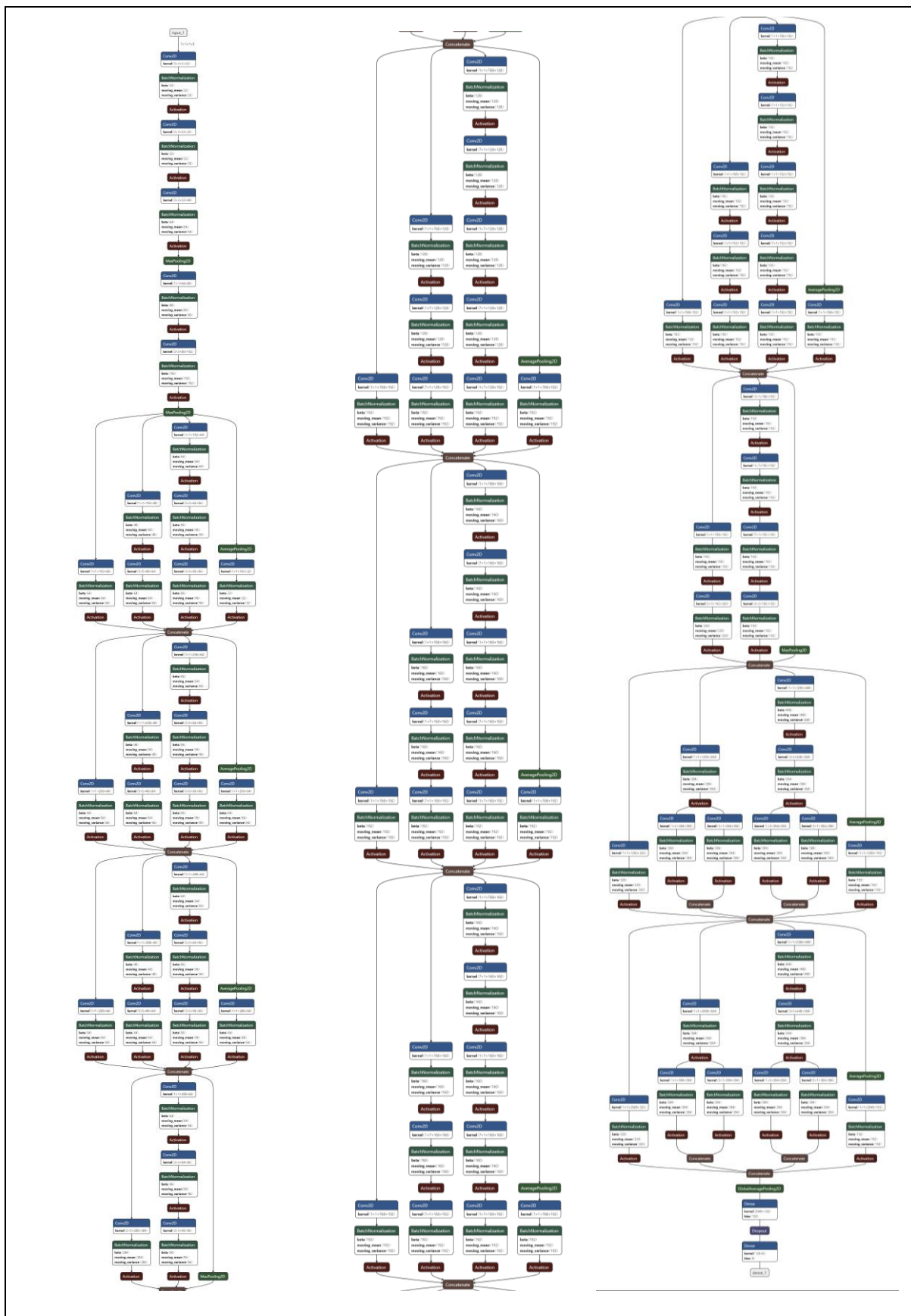


Figure 6 . Inception 모듈을 추가한 모델 사진

3) 배율에 따른 정확도 비교

배율에 따른 정확도를 비교하기 위해 위에서 인셉션 모듈을 추가한 모델을 사용하여 배율에 따라서 비교해보았다.

40 배율	100 배율	200 배율	400 배율
0.9326	0.9683	0.9578	0.9657

학습완료한 hdf5 모델 가중치 파일:

drive.google.com/drive/folders/1L9H_BgA6xgMnomq1HvMIId9oaL0VNuD3?usp=sharing

모두 epoch 를 50 으로 하여 학습시켰을 때 다음과 같은 정확도가 나왔는데 epoch 를 늘려도 accuracy 가 잘 증가하지 않았다. 40 배율일때가 가장 정확도가 낮은 것을 확인할 수 있었다.

하지만 배율에 따라서 상관관계를 정확하게 확인하기는 어려웠다.

4.2 조직병리학적 진단 요인을 고려한 이미지 특징강화 방법 적용에 따른 정확도 비교

	40 배율	100 배율	200 배율	400 배율
기준값을 넘으면 최대값 아니면 0 임계처리	0.9354	0.8687	0.9622	0.9685
기준값을 넘으면 원래값 아니면 0	0.9685	0.9736	0.9600	0.8877
흑백변환	0.9602	0.9500	0.9565	0.9331
평균 적응임계처리	0.9808	0.9769	0.9734	0.9640
가우시안블러 적응임계처리	0.9716	0.9683	0.9745	0.9602

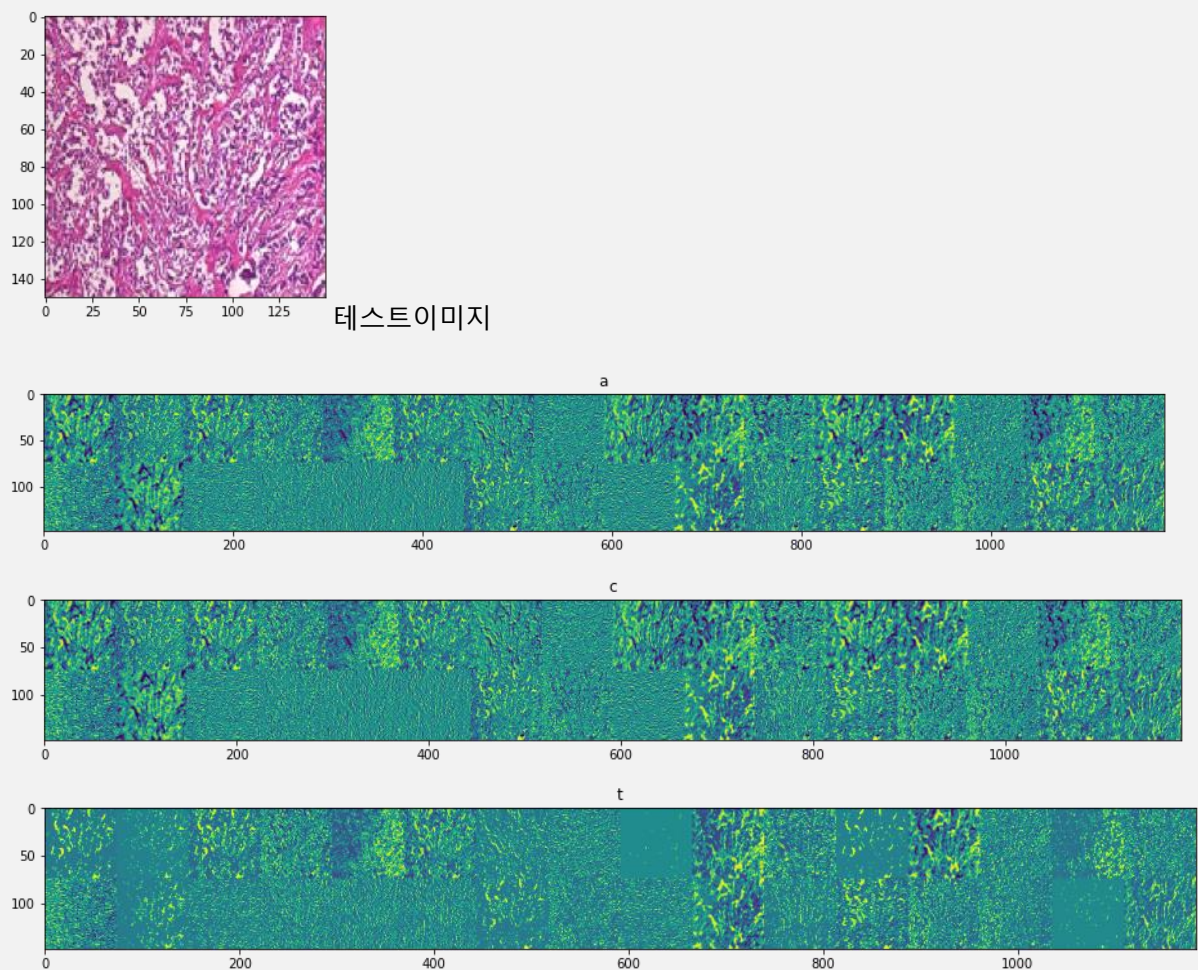
가우시안블러 적응임계처리와 평균적응 임계처리가 모든 배율 데이터에서 가장 높은 정확도를 보였고 원본이미지를 학습시킨 것보다 높은 정확도를 보임을 확인할 수 있었다.

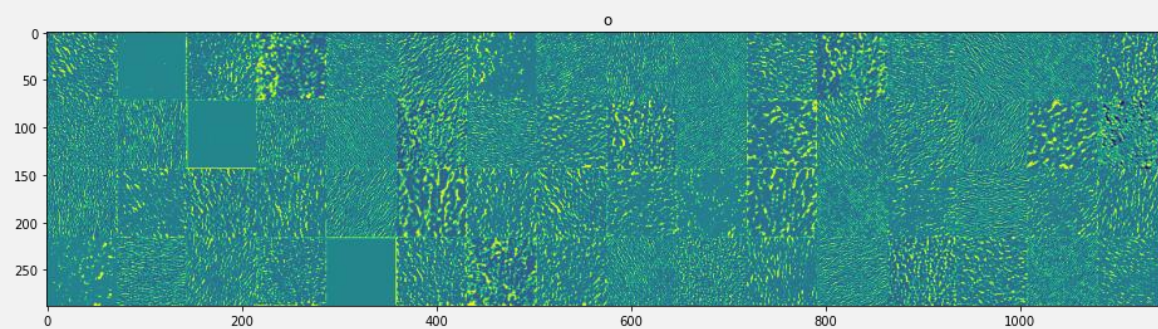
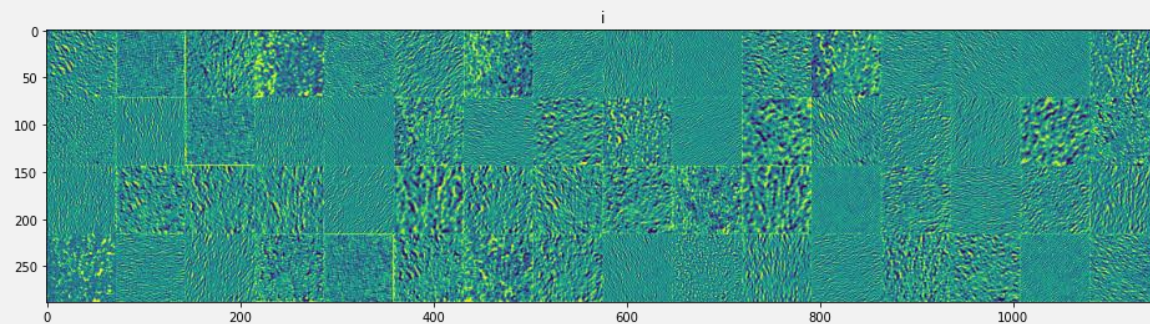
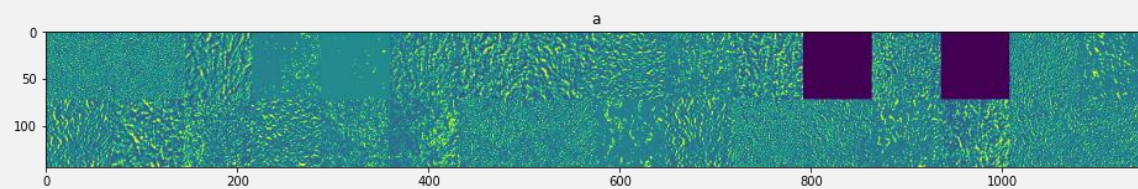
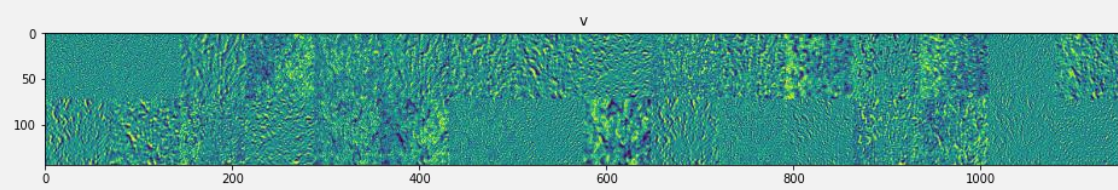
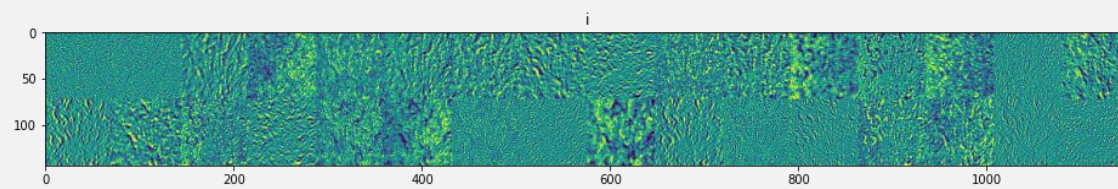
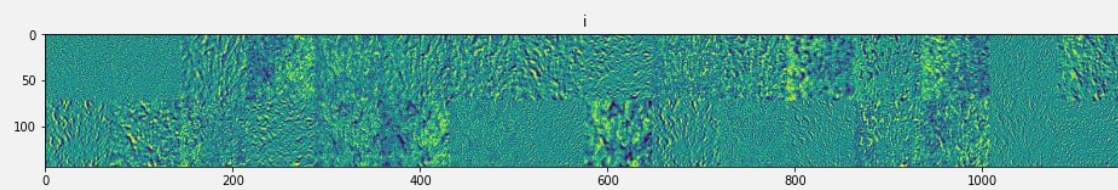
4.3 CAM(class activation map)으로 설명가능한 진단결과 확인

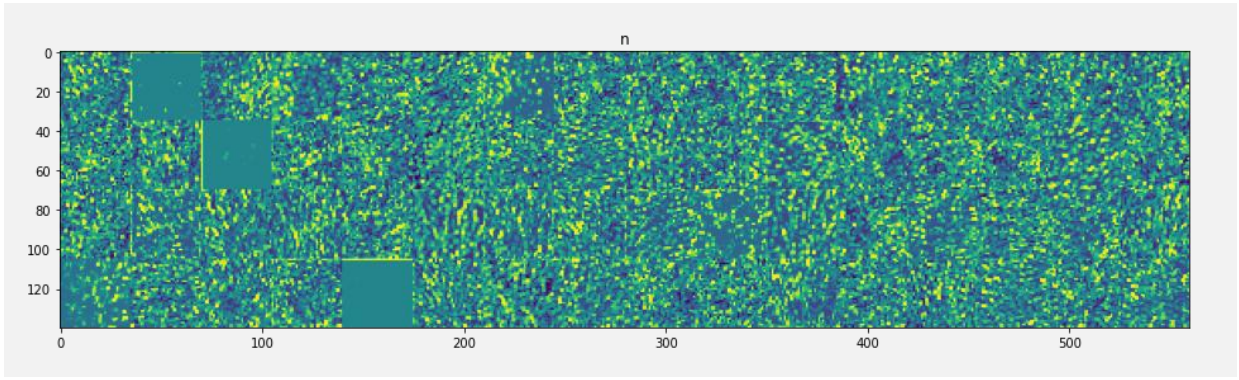
'conv2d', 'batch_normalization', 'activation', 'conv2d_1', 'batch_normalization_1', 'activation_1', 'conv2d_2', 'batch_normalization_2', 'activation_2', 'max_pooling2d'으로 층 구성이 되어있고 이 피쳐맵을 시각화하여 진단에 도움이 되는 설명력을 높이하고자 하였다.

1) 악성 40 배율

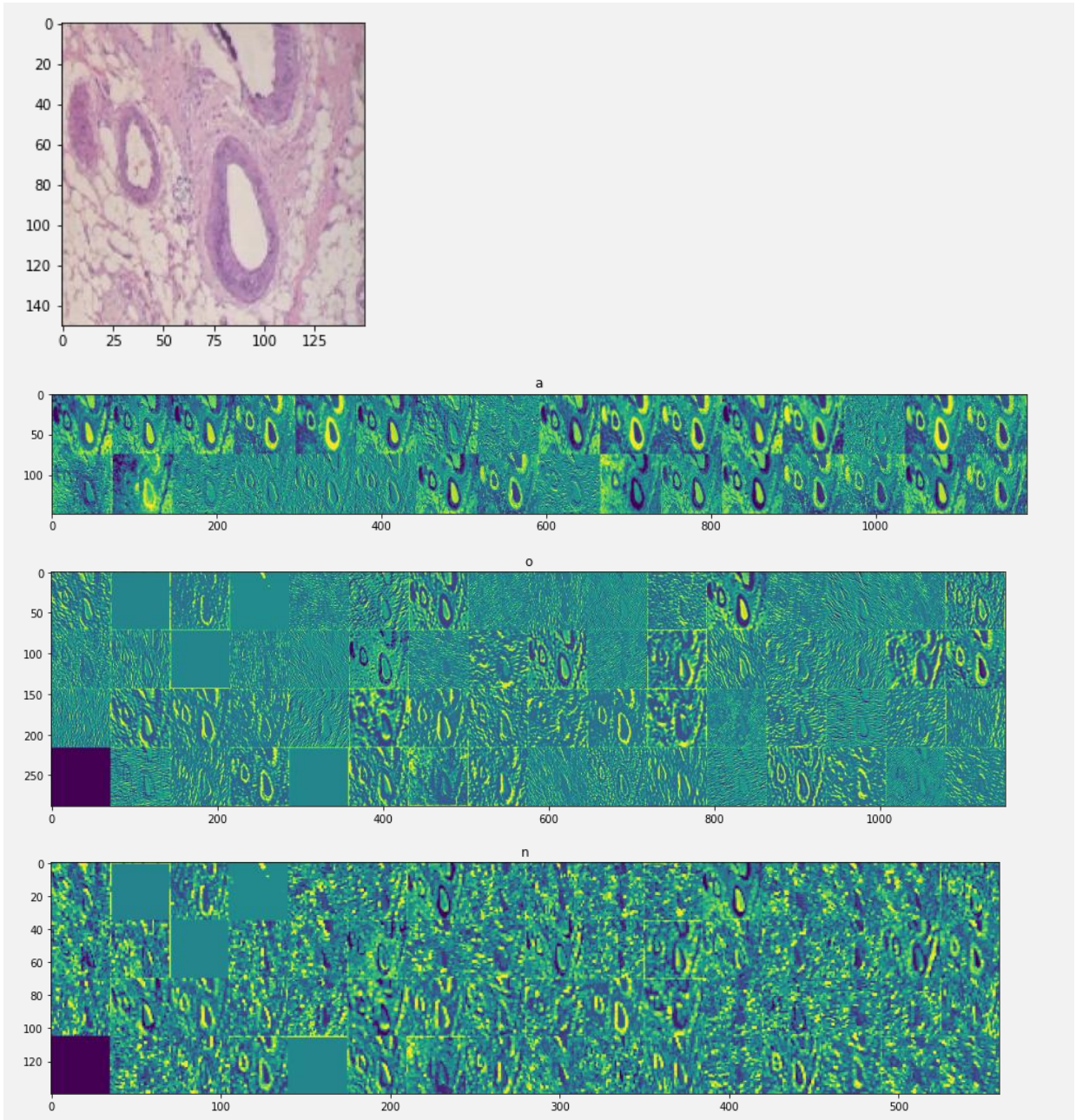
ductal_carcinoma



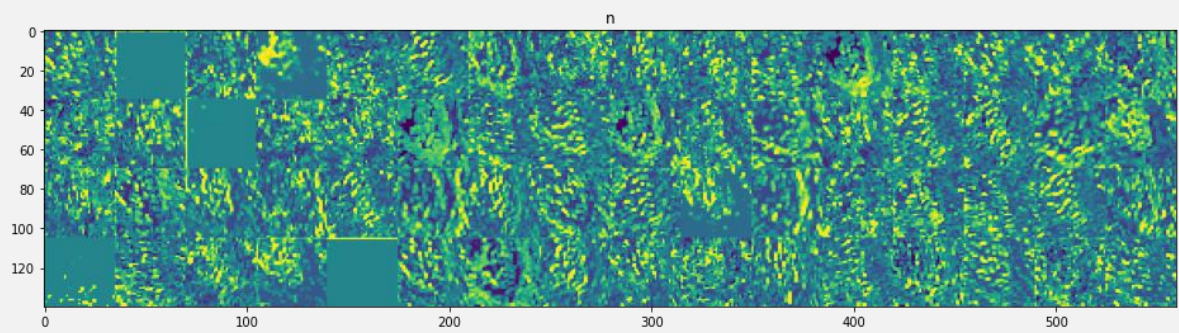
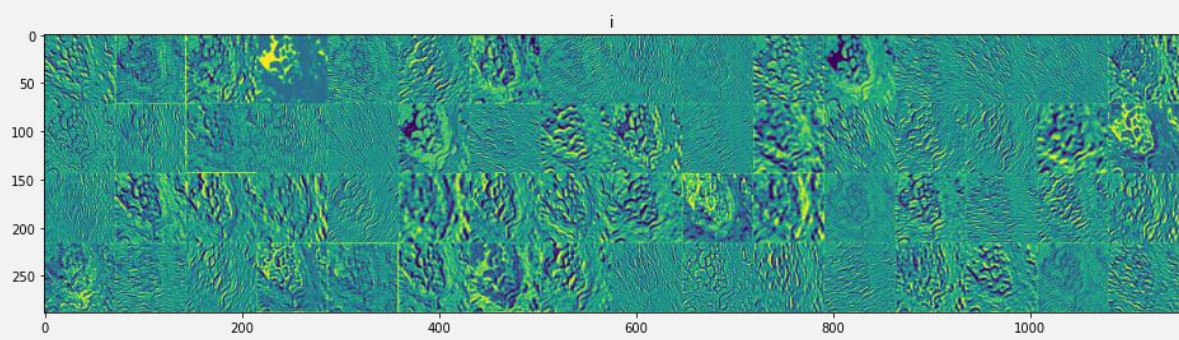
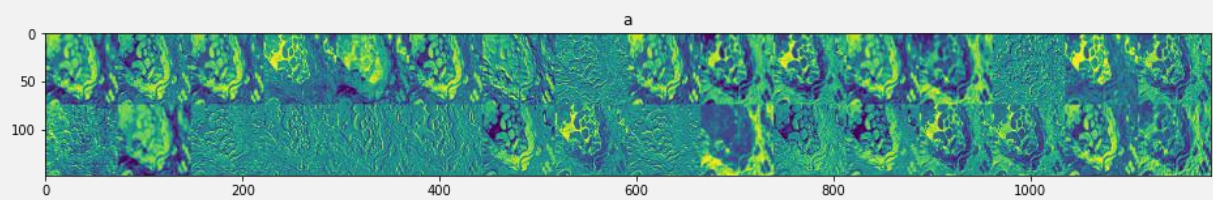
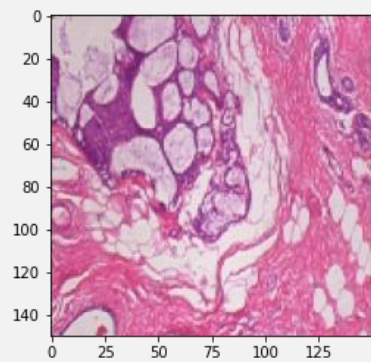




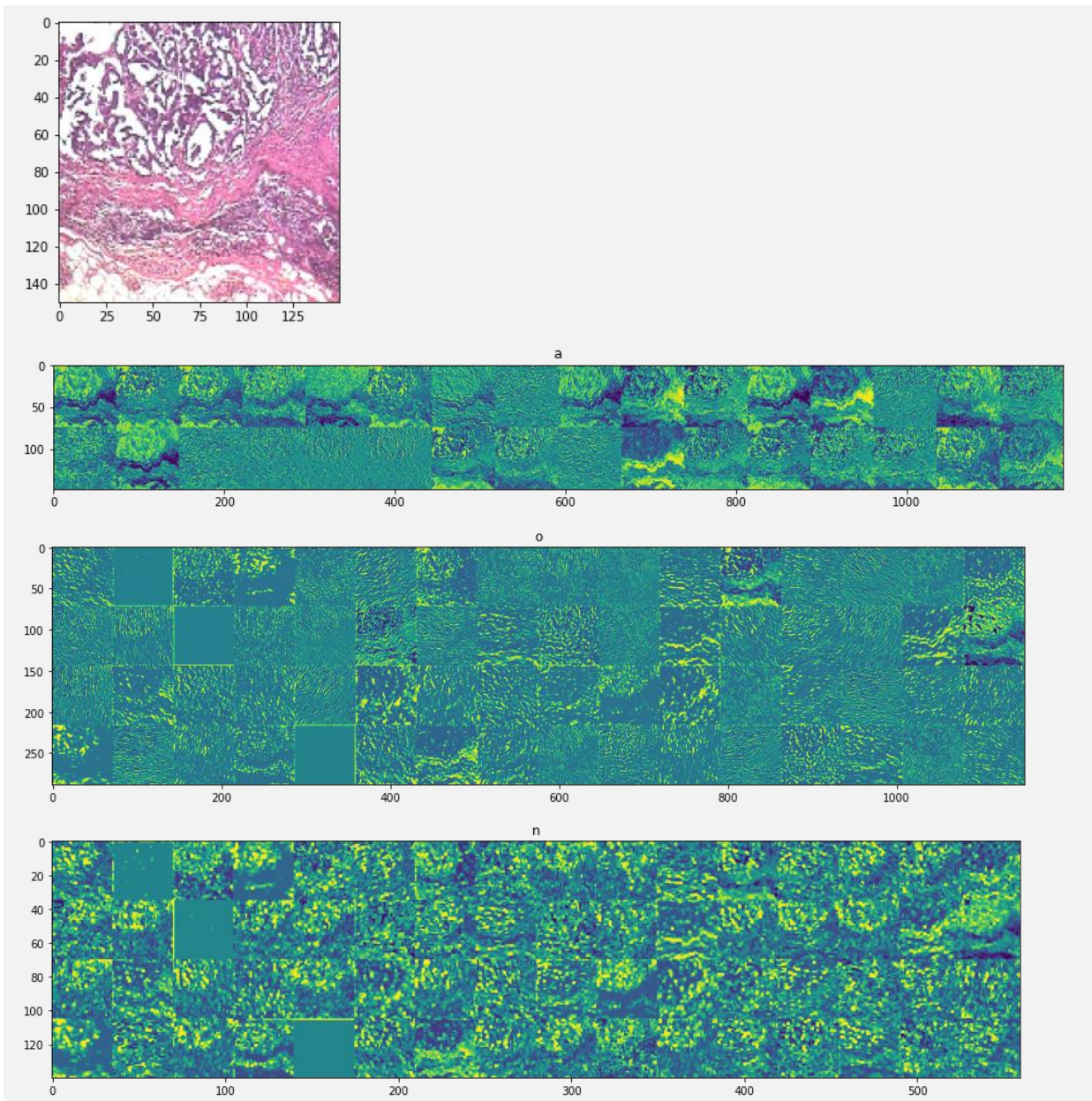
lobular_carcinoma



mucinous_carcinoma



papillary_carcinoma

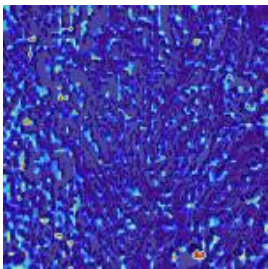
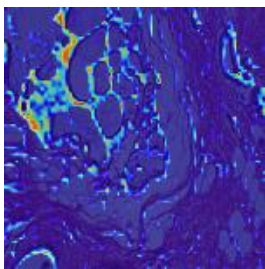
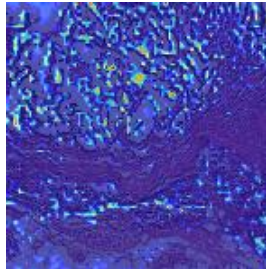
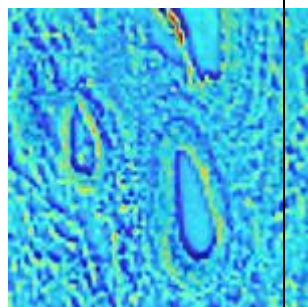
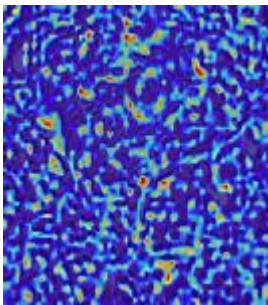
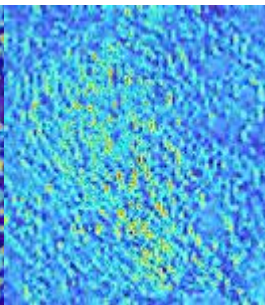
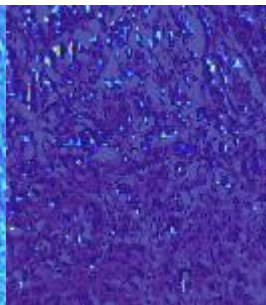
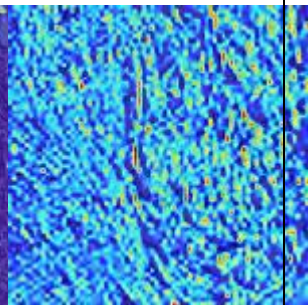
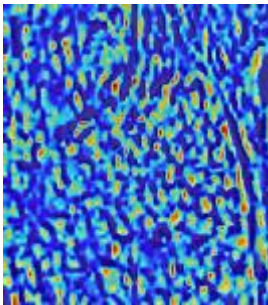
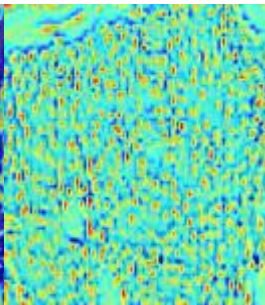
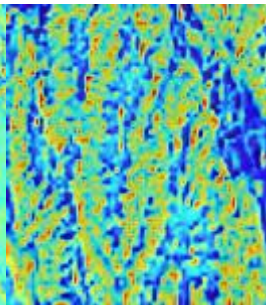
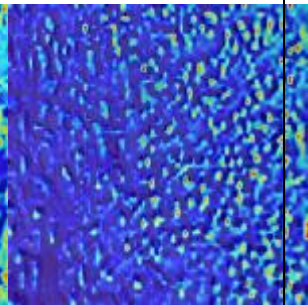


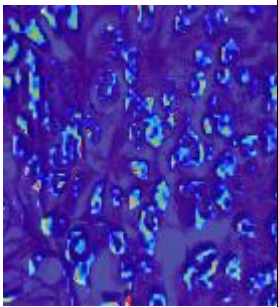
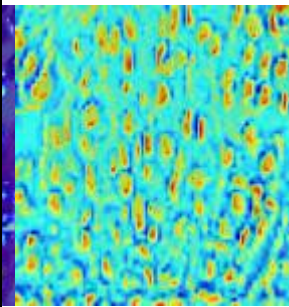
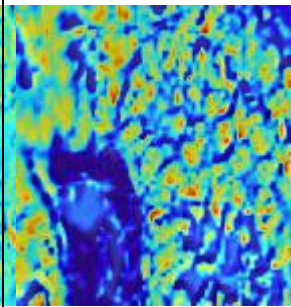
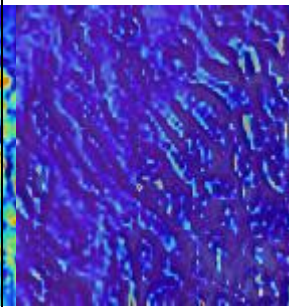
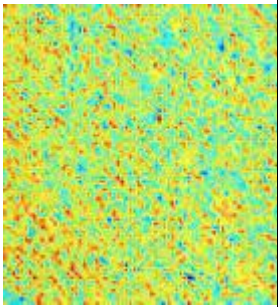
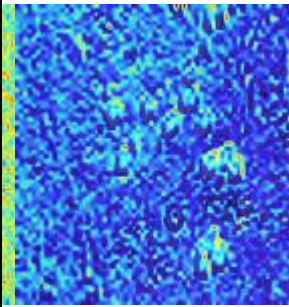
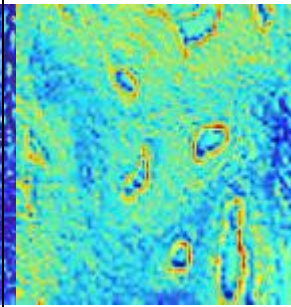
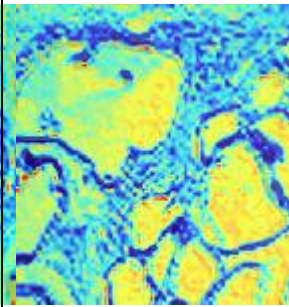
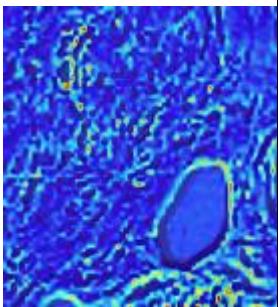
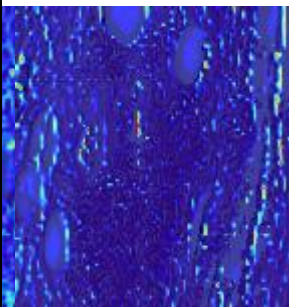
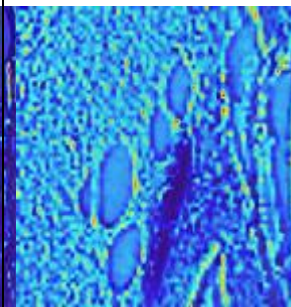
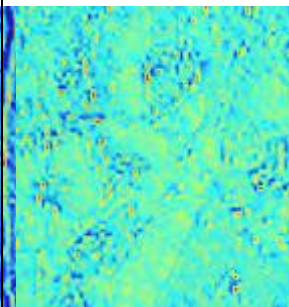
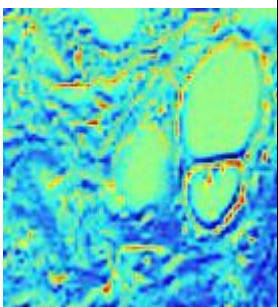
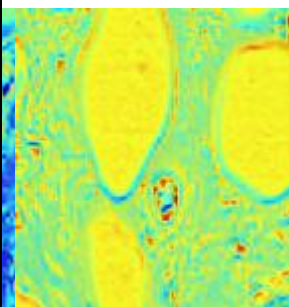
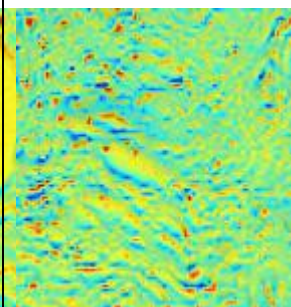
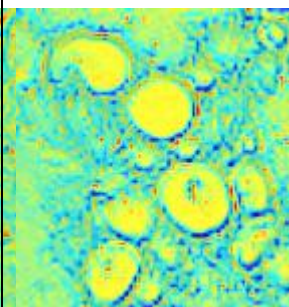
위의 cam 시각화에서 본 저자들이 학습시킨 11 개의 레이어의 액티베이션 맵이나 출력이다.

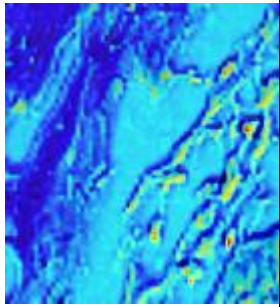
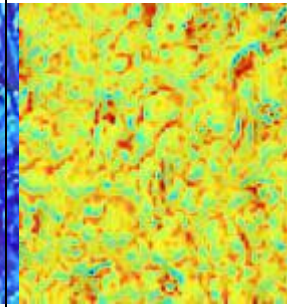
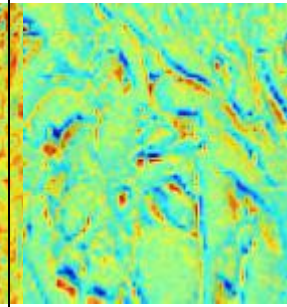
첫 번째 레이어 (conv2d_1)의 activation map 과 output 은 원래 인풋과 유사하게 가져가서 입력 이미지에 여러 필터리의 가중치를 볼 수 있었다. 레이어가 추가될 때마다 activation map 은 깊이가 증가함에 따라 레이어 activation map 은 시각적으로 더 추상적이 되어 단순히 모서리와 곡선이 아닌 클래스의 더 구체적인 특징을 감지할 수 있게 되는 것을 볼 수 있었다. 위에서 남색으로 되어있는 빈 레이어는 희소한 activation map 을 가지고 있는 것을 확인할 수 있었는데

이러한 공백 activation map 은 해당 레이어에서 사용된 필터가 주어진 입력에서 일치하는 패턴을
 찾지 못했을 때 발생하는 것으로 일부 공백 activation map 도 확인할 수 있었다.

4.4 Grad-cam 결과

악 성	ductal_carcinoma	mucinous_carcinoma	papillary_carcinoma	lobular_carcinoma
40				
100				
200				

400				
양 성	adenosis	fibroadenoma	phyllodes_tumor	tubular_adenoma
40				
100				
200				

400				
-----	---	---	--	---

Grad-CAM 을 이용하여 확인해 본 결과 세포핵이 염색이 잘 되다 보니 세포핵에 gradient 가 높게 나오는 것을 확인하였다. 세포핵의 모양이 불분명하거나 모양이 일정하지 않은 것을 보고 판단하는 조직병리학 소견에 따라 모델이 판단하는 중요한 세포핵을 보여줄 수 있다.

5 Conclusion & Recommendations

레이어의 깊이기 깊어질수록 성능은 높아지지만 free parameter 수가 증가하고, 유방암 데이터 셋은 배율에 따라서 1800 개에서 2000 개정도 되는데 각 class 의 데이터 셋도 적다는 것을 확인할 수 있었다. 그래서 데이터 학습량이 작으면 과적합 문제가 발생하고 기울기 감소문제가 생겼다는 것을 확인했다. 그래서 문제를 해결하기 위해서 인셉션모듈을 추가하여 하였을 때 정확도 향상을 확인해볼 수 있었다. 배율에 따라서는 유의미한 상관관계를 확인할 수 없었다. 하지만 세포핵이나 rest of cells 의 경계를 강화하기 위한 기준값을 넘으면 최대값 아니면 0 임계처리, 기준값을 넘으면 원래값 아니면 0, 흑백변환, 평균 적응임계처리, 가우시안블러 적응임계처리 방법을 사용했을 때는 평균 적응임계처리, 가우시안블러 적응임계처리 방법을 통해 이미지 적응임계처리를 한 후에 더 정확도가 높은 것을 확인할 수 있었다. 이는 임계처리의 경우 이미지 전체에 하나의 기준값을 적용해서 일정한 영역내의 근접해있는 픽셀값들을 고려하지 못하여 유방암 조직검사를 시행할 때 염색을 할 때 염색 정도에 따라서 오히려 특징이 부각이 안되는 경우가 생겨서 본 논문에서는 근접해있는 픽셀값들을 이용해서 해당 영역에 적용할 기준값을 자체적으로 계산하는 적응임계처리 방식이 더 좋다고 해석한다. 또한 CAM 을 사용한 activation

map 시각화를 통해 이미지에 여러 필터의 가중치를 볼 수 있었다. Grad-CAM 을 이용하여 확인해 본 결과 세포핵이 염색이 잘 되다 보니 세포핵에 gradient 가 높게 나오는 것을 확인할 수 있었다. 세포핵의 모양이 불분명하거나 모양이 일정하지 않은 것을 보고 판단하는 조직병리학에 따라 모델이 판단한 세포핵을 보여줄 수 있다. 하지만 일부 사진이 전체적으로 해석하기 어려운 부분이 있어 이 부분이 한계점인 것 같다.

6 References

- 유방 초음파 영상에서의 영상학적 특징들을 이용한 유방암 진단, 권보준(2013)
- Breast cancer detection by leveraging Machine Learning, Anji Reddy Vaka, Badal Soni, Sudheer Reddy K. (2020)
- Computer Methods and Programs in Biomedicine, Kalpana George, Praveen Sankaran, Paul Joseph K (2020)
- Wenzhong, L., Huanlan, L., Caijian, H., & Liangjun, Z. (2020). Classifications of Breast Cancer Images by Deep Learning. medRxiv.
- Loukas, C., Kostopoulos, S., Tanoglidi, A., Glotsos, D., Sfikas, C., & Cavouras, D. (2013). Breast cancer characterization based on image classification of tissue sections visualized under low magnification. Computational and mathematical methods in medicine, 2013.
- Alom, M. Z., Yakopcic, C., Nasrin, M. S., Taha, T. M., & Asari, V. K. (2019). Breast cancer classification from histopathological images with inception recurrent residual convolutional neural network. Journal of digital imaging, 32(4), 605-617.
- Kumar, R., Srivastava, R., & Srivastava, S. (2015). Detection and classification of cancer from microscopic biopsy images using clinically significant and biologically interpretable features.

Journal of medical engineering, 2015.

-Alzubaidi, L., Al-Shamma, O., Fadhel, M. A., Farhan, L., Zhang, J., & Duan, Y. (2020). Optimizing the Performance of Breast Cancer Classification by Employing the Same Domain Transfer Learning from Hybrid Deep Convolutional Neural Network Model. *Electronics*, 9(3), 445.

-Sun, Y., Xu, Z., Strell, C., Moro, C. F., Wörnberg, F., Dong, L., & Zhang, Q. (2018, December). Detection of Breast Tumour Tissue Regions in Histopathological Images using Convolutional Neural Networks. In *2018 IEEE International Conference on Image Processing, Applications and Systems (IPAS)* (pp. 98-103). IEEE.

- Breast Cancer Image Classification via Multi-Network Features and Dual-Network Orthogonal Low-Rank Learning, YONGJUN WANG 외 8명 (2019)

- Multi spectral classification and recognition of breast cancer and pneumonia, Aditya KAKDE, Nitin ARORA, Durgansh SHARMA, Subhash Chander SHARMA (2020)

- Prediction of Breast Cancer, Comparative Review of Machine Learning Techniques, and Their Analysis, NOREEN FATIMA, LI LIU, SHA HONG, HAROON AHMED (2020)

-Tumor Grade Fact Sheet. (n.d.). Retrieved from cancer.gov/about-cancer/diagnosis-staging/prognosis/tumor-grade-fact-sheet