

# 심층신경망을 이용한 조직검사 영상에서의 대장암 병변 분류<sup>+++</sup>

<sup>1)</sup>김상현, <sup>2)</sup>고현민, <sup>1)</sup>이찬수, <sup>1)</sup>이병대

<sup>1)</sup>경기대학교 컴퓨터공학부, <sup>2)</sup>창원경상대학교병원 병리과

e-mail : kws02352@naver.com, shekoh@hanmail.net,

ee.chansu@gmail.com, blee@kgu.ac.kr

## Classification of colorectal cancers in colon biopsy images using deep neural networks

<sup>1)</sup>Sang-Hyun Kim, <sup>2)</sup>Hyun-Min Koh, <sup>1)</sup>Chan-su Lee, <sup>1)</sup>Byoung-Dai Lee

<sup>1)</sup>Division of Computer Science & Engineering, Kyonggi University

Kyonggi, South Korea

<sup>2)</sup>Department of Pathology, Gyeongsang National University Changwon Hospital  
Changwon, South Korea

### 요 약

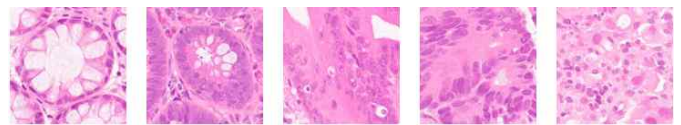
최근 우리나라의 사망 원인 1위는 암이며, 그중에서도 대장암 발생률이 급격하게 증가하고 있다. 대장암 대부분은 선암이며, 선암은 그 전암병변인 선종으로부터 발생하는 것으로 알려졌다. 선암 및 선종은 주로 조직검사를 통한 병리학적 관독으로 진단되는데, 이 과정이 매우 노동 집약적이어서 최근에 딥러닝을 이용한 병리관독에 관한 연구가 활발해지고 있다. 따라서 본 논문에서는 대장내시경을 통해 채취된 대장조직의 이미지를 이용하여 딥러닝을 통해 선종 및 선암을 판독하는 2단계 분류 방법을 제안한다.

### 1. 서 론

최근 10년 사이 우리나라의 대장암 발생률은 급격하게 증가하고 있다. 2018년 통계청에서 공개한 '2017년 사망통계원인' 보고서에 따르면 한국인 사망 원인 1위는 암으로 나타났으며, 그중에서 대장암은 폐암, 간암에 이어 세 번째로 많이 발생하는 것으로 보고되었다. 선암은 대장암 대부분을 차지하며, 그 전암병변인 선종으로부터 발생하는 것으로 알려졌다.

선암 및 선종은 대개 대장내시경을 이용한 조직검사를 통해 채취된 검사대상물을 가공 및 염색하여 유리 슬라이드로 제작한 후, 병리 의사에 의해 진단된다. 병리 의사는 제작된 유리 슬라이드에 포함된 조직을 현미경으로 관독한 후 진단하게 되는데, 정상 대장 및 선종의 경우 세포 배열 및 증식 정도의 차이점을 육안으로 상대적으로 쉽게 구분 가능한 반면, 선암은 상피세포의 융합 정도가 그림 1과 같이 연속성을 가지기 때문에 고분화형 (well differentiated), 중분화형 (moderately differentiated), 저분화형 (poorly differentiated) 유형으로 구분하는데 더 많은 관독 시간이 소요될 뿐만 아니라 매우 노동집약적이고 때에 따라서는 관독 자의 의견이 다른 때도 있다. 최근 이러한 문제점을 보완하기 위해 딥러닝을 이용한 병리관독에 관한 연구가 활발해지고 있다.

본 연구에서는 대장내시경을 통해 채취된 대장 조직의 이미지를 대상으로 정상, 선종 및 3가지 선암의 유형을 자동 분류하기 위한 심층신경망 기법을 제안한다. 특히 본 연구에서는 1단계에서는 정상, 선종, 선암을 분류한 후 선암인 경우에만 다시 2단계 분류 네트워크를 통해 선암을 재분류하는 2단계 분류 방법을 제안한다.



(그림 1) 정상, 선종, 고분화형, 중분화형, 저분화형의 모양

### 2. 관련 연구

유럽의 한 대학에서는 정상 대장, 관 용모샘종, 선암을 자동 분류하기 위해 VGG16[1] 모델을 이용하여 전이 학습을 수행하였다. 전체 27장의 슬라이드 이미지로부터 109개의 정상, 관 용모샘종, 선암 영역을 추출하고, 해당 영역으로부터 13,500개의 이미지 패치를 생성하여 학습 데이터 세트를 구축하였다. 실험 결과를 통해 전이 학습의 유효성을 증명하였으며, 특히 상대적으로 단순한 CNN 모델을 이용하더라도 일정 수준의 정확도를 가지고 정상 대장, 선종 및 선암 분류가 가능하다는 것을 보였다. [2] 이러한 실험 결과는 본 논문에서 제안하는 다중 분류 네트워크의 1단계 분류 과정에 대한 타당성을 입증한 것으로 볼 수 있다.

국내 한 대학과 기업이 공동 연구를 통해 정상 대장, 관샘종, 선암을 자동 분류하기 위해 2단계 분류 방법을 제안하였다. [3] 1단계에서는 입력 슬라이드 이미지를 작은 크기의 패치 영역으로 나눈 후, 개별 패치를 대상으로 Inception V3 기반의 CNN을 통해 정상, 관샘종, 또는 선암으로 분류한다. 개별 패치에 대한 분류 작업이 완료되면, 2단계에서는 패치들의 클래스 분포를 바탕으로 Random Forest 기법을 적용하여 대장암 종류를 최종 판별하게 된다. 학습을 위해 클래스별 30개의 샘플 슬라이드 이미지를 사용하였으며, 전체 548개의 슬라이드 이미지를 이용하여 테스트를 수행하였다. [4]

### 3. 데이터셋 & 메소드

#### 3.1 데이터셋

대장조직 이미지는 2016년 3월 1일에서 2019년 4월 30일까지 창원 경상대학교 병원에서 대장내시경을 통해 채취된 대장조직 총 850건을

+ 본 연구는 경기도의 경기도 지역협력연구센터 사업의 일환으로 수행하였음. [2019-0217, 영상 및 네트워크 기반 지능정보 제조 서비스 연구]

++ This study was approved by the Institutional Review Board of Gyeongsang National University Hospital with a waiver for informed consent(2019-05-008).

이용해 제작된 유리 슬라이드를 스캔하여 사용하였다. 또한, 본 연구에서는 경험이 있는 병리 의사의 판독을 통해 대장조직을 정상, 선종, 선암으로 크게 3가지로 분류하였으며, 선암의 경우에는 고분화형, 중분화형, 저분화형으로 분류하여 데이터 레이블링을 수행하였다.

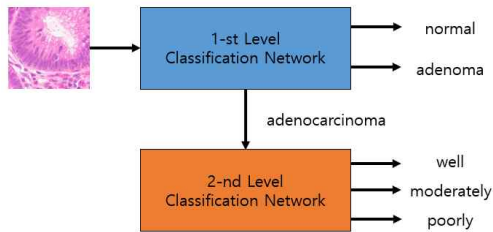
학습 데이터를 구성하기 위해 병리 의사에 의해 정상, 선종 혹은 선암으로 판별된 부분의 슬라이드 이미지를 40 배율로 확대하여 해당 병변의 특징이 포함된 영역을 400x400 크기 이미지 패치로 생성하였다. 또한 생성된 이미지 패치를 대상으로 크롭핑, 회전, 등의 기본적인 데이터 확대 기법(Augmentation)을 적용하여 표 1과 같이 데이터 세트를 구성하였다.

(표 1) 데이터 세트

	정상	선종	선암		
			고분화형	중분화형	저분화형
원본	348	353	360	342	348
패치수	25,111	25,452	25,920	24,624	25,096

### 3.2 메소드

정상과 선종, 선암의 경우 세포 배열 및 증식 정도의 차이점을 육안으로 상대적 쉽게 구분이 가능하다. 그러나, 선암의 세부 유형인 고분화형, 중분화형, 저분화형은 상피세포의 융합 정도가 연속성을 가지기 때문에 정확히 분류하는데 많은 어려움이 있을 뿐만 아니라 임상 의 간 개인차도 존재한다. 따라서, 본 논문연구에서는 클래스 간 특징이 상대적으로 뚜렷한 정상과 선종, 선암을 1단계 분류기에서 분류하며, 선암으로 판단된 경우에는 2단계 분류기를 통해 고분화형, 중분화형, 저분화형으로 그림 2와 같이 구분한다.

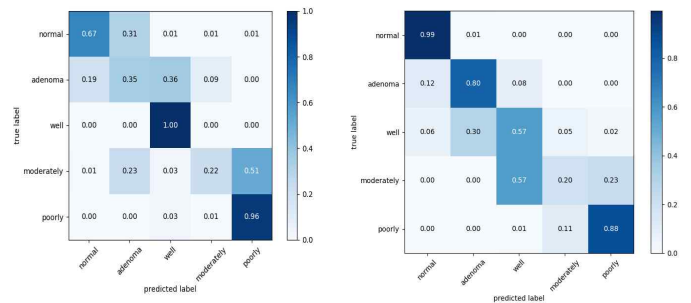


(그림 2) 2단계 분류

### 4. 실험 및 결과

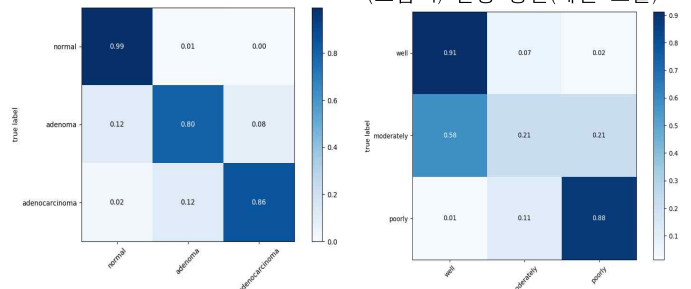
본 연구에서 1단계와 2단계 분류기를 위해 121개 층으로 구성된 Dense-net[5]을 사용하여 구현하였다. 학습에 사용된 데이터 세트는 전체 데이터 세트 수의 80%를 사용하였으며, 테스트 세트는 학습에 사용되지 않은 정상, 선종, 고분화형, 중분화형, 저분화형 유형에 대해 각 150장으로 총 750장의 데이터로 구성하였다. 학습의 빠른 수렴을 위해서 이미지-넷에 대해 사전 훈련된 네트워크[6]의 매개변수를 사용하여 네트워크를 초기화하였다. 배치 크기는 32로 설정하였다. 학습률은  $1e^{-4}$ 과  $1e^{-5}$ 로 하였으며, 최적화를 위해서 Adam을 사용하여 학습을 진행하였다.

본 논문에서 제시하는 2단계 분류 방법과의 성능 비교를 위해 1단계 과정에서 대장암의 5가지 유형을 모두 분류할 수 있는 기본 모델을 함께 구현하였다. 기본 모델은 제안된 모델에서 사용된 것과 동일한 121개층으로 구성된 Dense-net을 사용하였으며 출력층의 수를 5개로 설정하여 학습을 수행하였다. 기본 모델 및 2단계 분류 모델의 혼동 행렬(Confusion Matrix)은 그림 3과 4와 같다. 또한 기본 모델 및 2단계 분류 모델의 mAP은 각각 0.6과 0.62으로 본 연구에서 제안하는 2단계 분류 모델이 성능상의 우위를 보였다. 그러나, 그림 3과 4에서 보여지는 바와 같이 1단계 분류기의 성능은 우수한 반면, 2단계 분류기는 상대적으로 성능이 떨어졌으며, 이를 통해 연속성을 가지는 선암의 세부 유형을 분류하는데 한계가 존재함을 확인할 수 있었다.



(그림 3) 혼동 행렬(기본 모델)

(그림 4) 혼동 행렬(제안 모델)



(그림 5) 제안 모델의 1단계 분류기 성능

(그림 6) 제안 모델의 2단계 분류기 성능

### 5. 결론 및 향후 발전 계획

본 논문에서는 선암의 융합 정도가 연속성을 가지기 때문에 정상, 선종, 선암을 1단계에서 분류를 한 후, 선암으로 판별된 경우에 고분화형, 중분화형, 저분화형으로 다시 3가지로 나누는 2단계 분류 방법을 제안하였다.

향후에는 선암의 분류 대상 클래스 간 연속성으로 인해 경계 영역에서 구분이 어려운 문제를 해결하기 위한 방법의 일환으로 Ranking-CNN[7]을 사용하여 연속성을 가진 데이터의 분류에 대한 성능 개선 방안을 연구하고자 한다.

### 참고문헌

- [1]K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition." Proc. of ICLR2015, 2015.
- [2]F. Ponzio, et. al., "Colorectal Cancer Classification using Deep Convolutional Networks: An Experimental Study," Proc. of BIOSTEC2018, January, 2018.
- [3]C. Nam, et. al., "Lung nodule segmentation with convolutional neural network trained by simple diameter information," Proc. of MIDL2018, July, 2018.
- [4]C. Szegedy, et. al., "Going deeper with convolutions," Proc. of CVPR2015. 2015.
- [5]Gao Huang, et .al. "Densely connected Convolutional Networks," Proc. of CVPR2017, 2017.
- [6]<https://github.com/tensorflow/models/tree/master/research/slim>.
- [7]Shixing Chen, et. al. "Using ranking-cnn for age estimation," Proc. of CVPR2017, 2017.