

인공지능을 활용한 대장 내시경 영상에서의 용종 검출 모델 연구

A Study on a Polyp Detection Model in Colonoscopy Images Using Artificial Intelligence

이소현¹, 김영재², 박동균³, 김재승⁴, 김광기^{5*}

So Hyeon Lee¹, Young Jae Kim², Dong Kyun Park³, Jae Seung Kim⁴, Kwang Gi Kim^{5*}

요 약

본 연구에서는 딥러닝 기술을 활용해 대장 내시경 정지 영상에서 용종의 위치를 검출하는 모델을 제안하고자 한다. 실제 대장 내시경 영상로만 이루어진 데이터로 학습을 진행하여 용종 검출 모델을 구축하였고, 용종이 존재하지 않는 정상 데이터 6,264장과 용종이 존재하는 비정상 데이터 4,445장을 학습 및 성능 검증을 위한 평가용 데이터 셋으로 구성했다. 용종의 위치 검출을 위해 객체 감지 모델 중 하나인 RetinaNet을 기반으로 한 용종 위치 검출 모델을 개발하였다. 9,639장의 학습 데이터를 사용하여 모델을 개발하였으며, 1,070장의 성능 평가용 데이터를 사용하여 모델의 성능을 확인하였다. K-fold 교차검증을 통해 모든 집합에 대해 검증을 하였고 그 결과 학습된 모델의 성능은 평균적으로 96.2%의 민감도(Sensitivity)와 0.227의 FPPI(Fasles Positives Per Image)를 나타냈다. 데이터 전처리 및 후처리 적용과 동영상 데이터 수집 및 추가 학습을 통한 학습 모델의 성능 고도화를 통해 다양한 방면으로 임상자에게 도움을 줄 수 있다.

핵심어 : 대장암, 내시경, 딥러닝, 객체 검출, 컴퓨터 보조 진단

Abstract

In this study, we propose a model for detecting the position of polyps in colonoscopy still images using deep learning technology. The polyp detection model was constructed by performing learning only with data consisting of actual colonoscopy images, and 6,264 normal data without polyps and 4,445 abnormal data with polyps were composed of evaluation data sets for learning and performance verification. For polyp position detection, we developed a polyp of position detection model based on RetinaNet, one of the object

1 Department of Biomedical Engineering Medical Center, College of Medicine, Gachon University, Incheon, Korea [Researcher]
e-mail: 99lsh0319@gmail.com

2 Department of Biomedical Engineering, College of IT Convergence, Gachon University, Seongnam, Korea [Associate professor]
e-mail: youngjae@gachon.ac.kr

3 Health IT Research center, Gachon University Gil Medical Center, Incheon, Korea [Professor]
e-mail: pdk66@gilhospital.com

4 Health IT Research center, Gachon University Gil Medical Center, Incheon, Korea [Researcher]
e-mail: mpmkjs@gilhospital.com

5 Department of Biomedical Engineering Medical Center, College of Medicine, Gachon University, Incheon, Korea [Professor]
e-mail: kimkg@gachon.ac.kr (Corresponding author)

* 본 연구는 보건복지부의 재원으로 한국보건의료정보원의 의료데이터 중심병원 지원사업 지원에 의하여 이루어진 것임

Received(December 7, 2022), Review Result(1st: January 5, 2023), Accepted(February 10, 2023), Published(February 28, 2023)



© 2023 The Authors. Published by NCISS.
This is an open access article licensed under the Creative Commons Attribution-NonCommercial 4.0 International License.
To view a copy of this license, visit <http://creativecommons.org/licenses/by-nc/4.0/>.

detection models. The model was developed using 9,639 learning data, and the performance of the model was confirmed using 1,070 performance evaluation data. K-fold cross-validation was used to validate all sets, and the results showed that the learned model's performance averaged 96.2% sensitivity and 0.227 False Positives Per Image (FPPI). It can help clinicians in various ways by applying data preprocessing and postprocessing, and upgrading the performance of the learning model through video data collection and additional learning.

Keyword : colon cancer, endoscopy, deep learning, object detection, Computer Aided Diagnosis(CAD)

1. 서론

대장암은 국가암검진 사업 대상 암 중인 6대 암 중 하나이며 국립 암 센터 2019년 통계에 따르면 남녀 전체에서 발병률 남녀 전체에서 4위를 차지한 암이다 [1]. 대부분의 대장암은 선종에서 발생하며 용종에서 암적 변화를 일으키는 데까지 장시간이 소요되고 대장암 초기에는 눈에 띄는 증상이 없어 정기적인 검사를 받는게 중요하다. 대장내시경 검사는 전 대장의 병변을 진단하고 경우에 따라서는 병변 제거 및 생검을 할 수 있으며 [2] 현재 진단의 정확도를 높이기 위해 내시경 이미지의 고화질화 및 FPS(Frame Per Second) 증가되었다. 내시경 기술의 발전에 따라 내시경 검사를 받는 환자의 수가 증가로 인해 임상상의 진찰 시간과 피로도가 증가하였다. 이로 인해 용종 간과율이 증가할 수 있고, 현재 간과율이 22~28%로 보고되고 있다 [3].

대장 내시경 검사의 정확성을 높이고 진찰 시간과 피로도를 줄여 간과율을 낮추기 위해 컴퓨터 보조 진단(Computer-Aided Diagnosis, CAD) 기술이 연구가 이루어지고 있다.

인공지능이 적용되기 이전인 2003년에 Stavros A Karkanis 등은 용종의 색상이나 모양을 기반으로 특징을 추출하여 LDA(Linear Discriminant Analysis)를 이용해 용종을 분류 [4]하는 연구가 진행되었지만 용종과 비슷하게 보이는 것에 대한 구분이 어려워 신뢰도가 낮은 문제점이 있었다. 최근 인공지능 기술을 접목한 보조 진단 연구에 대한 예시로 2019년 Xu Zhang등은 위내시경에서 Single Shot Multibox Decton(SSD) 네트워크를 이용한 용종 검출 방법을 제안하였고, 708개 데이터를 대상으로 학습 후 50건에 대해 검증한 결과, 최대 90.9mAP를 보였다 [5]. 2021년 이정남 등은 You Only Look Once(YOLOv4) 네트워크를 이용한 용종 탐지 시스템 연구에서는 이미지 증강 기법인 자동 증강(AutoAugment) 기법을 사용하여 얻은 데이터를 포함하여 12,728장을 학습 데이터 사용 후, 61장의 데이터에 대해 검증하여 87.9mAP의 결과를 도출해냈다 [6]. 또한, 2021년도 윤원중 등은 YOLOv4 네트워크와 Fast RCNN(Fast Region-based Convolutional Network method) 네트워크를 이용해 실시간으로 대장용종을 검출 방법을 제안하는 연구에서는 6720장의 이미지로 학습하여 얻은 모델을 52장의 검증하여 YOLOv4 모델의 성능은 58%의 재현율(Rcall)과 88%의 정밀도(Precision)를 보였고 Fast RCNN은 60%의 재현율과 81.8%의 정밀도를 도출해냈다 [7].

기존의 연구들은 공통적으로 비교적 적은 수의 데이터를 수집 후 자동 증강 기법이나 이미지 회전(Rotation), 반전(Flip)등 이미지 증강 방법을 통해 부족한 데이터를 확보하였다. 이러한 고식적

데이터 증강 방법은 모델 성능 향상에 영향을 미칠 수 있으나 크지 않으며, 일반화 문제를 해결하는 데에는 한계가 있다. 따라서 본 연구에서는 많은 수의 임상 데이터를 수집 후 실제 데이터로만 데이터셋을 구성하고 인공지능 기술을 기반으로 정확도 높은 대장 내시경에서 자동으로 용종을 인식하고 검출할 수 있는 모델은 개발하고 검증하고자 한다.

2. 본론

2.1 개발환경

본 연구에서는 시스템 환경이 4개의 NVIDIA TITAN Xp (NVIDIA, Santa Clara, California, USA) 그래픽 처리 장치와 Intel Xeon E5-1650 v4의 CPU, 125GB RAM으로 갖추어져 있는 시스템을 이용하였다.

딥러닝 학습을 위해 Ubuntu 18.06 운영 체제에서 Python(version 3.6.12) 및 OpenCV(version 3.4.2), Tensorflow(version 1.14.0), Keras(version 2.3.1) 등의 라이브러리를 사용하여 모델을 학습시켰으며 pandas 라이브러리를 사용해 성능에 대한 결과를 추출했다.

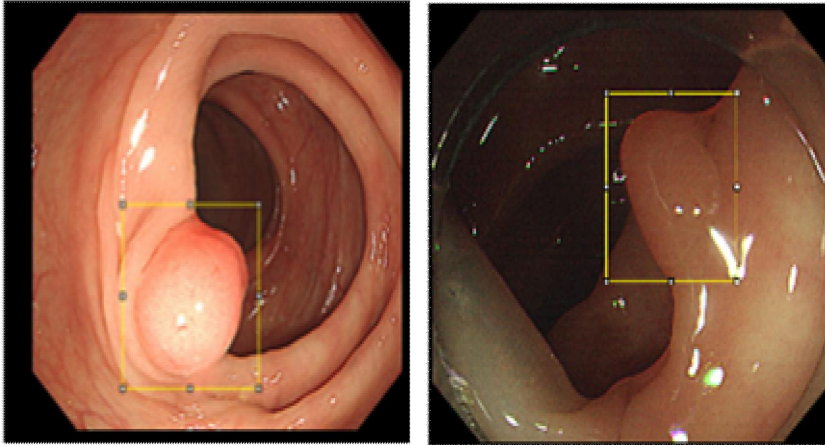
2.2 데이터

본 연구는 가천대 길병원 소화기내과에서 대장내시경 검사를 받은 환자의 대장내시경 정지 이미지 10,709장을 수집하였다. 수집된 대장내시경 이미지는 정상데이터 6,264건, 용종이 존재하는 데이터 4,445건으로 구성되어 있다.

대장 내시경에서의 용종 검출 모델 개발을 위해 수집된 영상의 크기는 992pixel x 862pixel에서 1280pixel x 1024pixel 사이로 영상의 형식은 jpg, 영상의 비트수는 1pixel당 24bit다. 또한, 학습 데이터셋은 수집된 데이터를 바탕으로 9:1 비율로 학습 데이터와 성능 검증을 위한 평가용 데이터를 나눠 9,639장의 학습 데이터와 1,070장의 성능 검증을 위한 평가용 데이터로 사용하였다.

2.3 데이터 레이블링

본 연구에서는 딥러닝 모델의 학습 및 검증을 위해 용종의 위치에 대한 레이블링을 통해 정답 데이터(Ground Truth, GT)를 확보하였다. 이를 위해 자체 개발한 레이블링 소프트웨어를 사용하여 소화기내과 전문의가 직접 용종의 위치에 대한 박스(Box) 형태의 관심 영역(Region of Interest, ROI)을 그려서 레이블링 데이터를 수집하였다. 이때 관심 영역은 용종이 잘리지 않고 전체 형태를 포함할 수 있도록 그렸으며, [그림 1]은 본 연구에서 수집된 용종 위치에 대한 관심 영역의 예를 보여준다.



[그림 1] 관심 영역에 대한 레이블링 데이터

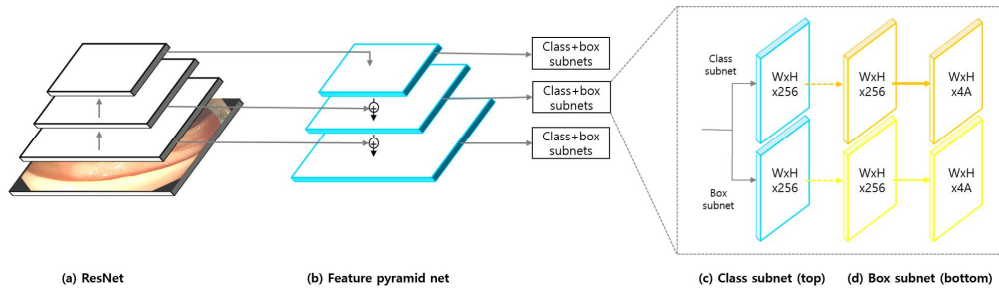
[Fig. 1] Labeling data for Regions of interest

2.4 전처리

대장 내시경에서의 용종 검출 모델 개발을 위해 다양한 크기를 가지는 영상의 높이와 너비를 512pixel x 512pixel로 동일하게 입력 데이터의 크기를 조정하였다. 영상의 높이와 너비를 줄이는 과정에서 데이터 소실을 막기 위해 원본 데이터의 비율을 유지하면서 영상의 크기를 조정하는 제로 패딩(Zero Padding) 후 크기 조정(Resize)을 했다. 또한, 원본 데이터를 입력 데이터의 크기를 조정하는 데에 쓰인 보간 방법으로는 역방향사상으로 계산된 실수값 위치에 가장 가까운 입력 화소를 출력 화소로 할당하는 방법인 최근접 이웃 보간법(Nearest Neighbor Interpolation)을 사용하였다 [8].

2.5 용종 위치 검출 모델 개발

대장의 용종 검출을 위해 사용한 모델은 객체 검출 모델 중 하나인 RetinaNet 모델을 사용하였다. RetinaNet 모델은 One-stage detector로 속도가 빠르다는 장점이 있으며 백본넷(backbone network)과 2개의 서브넷(subnet)이 있는 구조다. RetinaNet 모델은 ResNet 기반으로 입력 이미지에서 다양한 크기의 특징 피라미드를 추출한 후 2개의 서브넷을 통해 첫 번째 서브넷은 객체분류(classification)를 하고 두 번째 서브넷은 bounding box regression을 수행한다 [9]. 또한, One-stage detector 학습 시 발생하는 클래스 불균형 문제를 해결하기 위해 Focal loss를 사용했으며 모델의 학습에 사용된 모델의 구조는 [그림 2]와 같으며 딥러닝 학습 파라미터는 배치 사이즈(batch size)는 12, 학습률(learning rate)은 0.001, 학습횟수(epoch)는 100으로 설정하여 학습을 진행하였다.



[그림 2] 학습에 사용된 모델인 RetinaNet 네트워크 구조 (a) ResNet 구조 (b) 특징 피라미드 (c) 분류를 위한 서브넷 (d) 객체 검출을 위한 서브넷

[Fig. 2] The model used for learning: RetinaNet network structure (a) ResNet structure (b) feature pyramid (c) subnet for classification (d) subnet for object detection

3. 결과

대장 내시경 영상에서의 용종 위치를 검출해 내는 모델의 성능 평가를 위해 1,070장의 성능 평가용 데이터를 사용하였다. 모델의 성능 검증을 위한 평가용 데이터 1,070장 중 627장은 정상 데이터, 446장은 용종이 1개 이상 있는 데이터로 구성하였다. 본 연구에서는 모델의 성능 검증을 위해 전체 데이터셋의 모든 샘플들을 한 번씩 테스트 될 기회를 갖도록 하는 방법인 K-fold 교차검증을 사용하였다 [10]. 이 방식을 사용한 이유는 학습 데이터와 성능 평가용 데이터로 여러 번 나눈 것의 평균적인 성능을 계산하면 한 번만 나누어 학습하는 것에 비해 일반화된 성능을 얻을 수 있기 때문이다.

본 연구에서는 K값을 10으로 설정하여 전체 데이터 10,709장의 데이터를 10개로 분할한 뒤 9개를 학습용 데이터 셋으로, 1개를 성능 평가용 데이터셋으로 사용하여 10번 반복을 통해 모든 집합에 대해 검증을 하였다.

또한, 용종이 있는 위치를 올바르게 탐지해낸 경우를 진양성(True Positive, TP), 용종이 없는 부분을 탐지한 경우에는 위양성(False Positive, FP), 용종을 검출하지 못한 경우를 위음성(False Negative, FN)이라고 하며 이를 이용하여 민감도와 FPPI(Fasles Positives Per Image)로 용종 위치 검출 모델에 대한 성능 평가를 진행하였다. 민감도(Sensitivity)는 식(1)처럼 계산하였고, FPPI는 이미지 1장당 평균 위양성의 개수를 의미하며 식(2)와 같이 계산하였다.

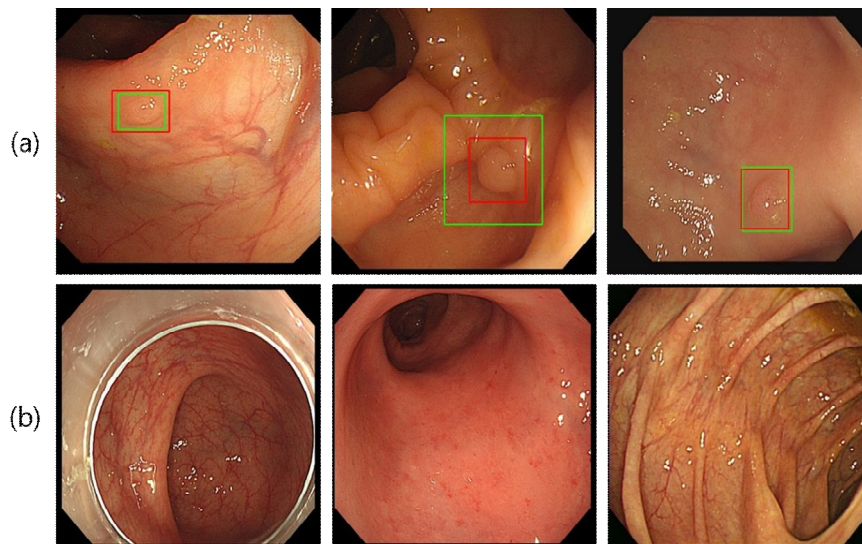
$$\text{민감도(Sensitivity)} = \frac{TP}{TP + FN} \quad (1)$$

$$FPPI = \frac{FP}{\text{Number of images}} \quad (2)$$

[표 1] 10-fold 교차검증에 대한 각 집합의 결과값 및 평균값

[Table 1] Results and average values of each set for 10-fold cross-validation

CV	TP	FN	FP	Sensitivity (95% CI)	FPPI
1	427	16	245	0.964 (0.946 - 0.981)	0.232
2	420	23	128	0.948 (0.938 - 0.957)	0.121
3	427	19	167	0.957 (0.944 - 0.969)	0.158
4	428	16	314	0.964 (0.942 - 0.985)	0.298
5	437	22	279	0.952 (0.932 - 0.971)	0.265
6	428	11	233	0.975 (0.958 - 0.991)	0.221
7	426	14	284	0.968 (0.948 - 0.987)	0.269
8	429	11	301	0.975 (0.954 - 0.995)	0.285
9	424	25	190	0.944 (0.929 - 0.958)	0.181
10	433	12	253	0.973 (0.955 - 0.990)	0.240
Total	4,279	169	2,394	0.962 (0.956 - 0.967)	0.227



[그림 3] 용종의 위치 검출 학습모델의 예측 결과 이미지. (a) 용종이 있는 데이터에서의 모델 예측 결과 (b) 정상 데이터에서의 모델 예측 결과. 빨간색 box는 실제 용종의 위치 및 영역, 초록색 box는 모델이 예측한 용종의 위치 및 영역

[Fig. 3] Prediction result image of polyp location detection learning model. (a) Model prediction results in data with polyps (b) Model predicted results on normal data. The red box is the location and area of the actual polyp, and the green box is the location and area of the polyp predicted by the model.

확률값(probability)이 0.1 이상인 결과값과 IoU(Intersection over Union)는 0.3 이상인 결과값에 대한 것들만 검증을 진행해 높은 민감도를 얻고자 하였다. 그 결과 평균적으로 0.962의 민감도와 0.227의 FPPI 값을 얻었으며 10개의 각 집합에 대한 결과값을 [표 1]과 같이 정리하였다. [표 1]을 보면 위음성에 비해 위양성의 개수가 많이 나타나는 것을 볼 수 있었고 이미지 1장당 1개 이상의 오검출되는 것을 확인할 수 있다.

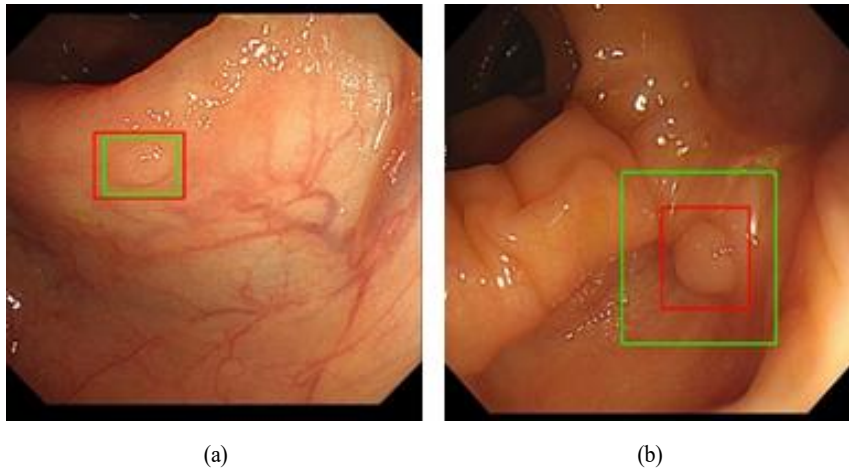
[그림 3]는 모델이 용종이라고 검출해 낸 결과로 실제 용종의 위치를 빨간색 Box로, 모델이 예측한 용종의 위치를 초록색 Box로 나타내었다.

4. 고찰

본 연구에서는 대장 내시경에서 딥러닝을 이용하여 용종의 위치를 검출해 내는 학습 모델을 제안하였다. 용종의 위치를 검출하기 위해 사용한 RetinaNet 모델은 성능이 우수한 객체검출기 중 하나로 손꼽히며 One Stage Detector 방식으로 빠른 검출 시간이 장점인 모델이다. 용종 위치 검출 모델은 10-fold 교차검증을 이용해 모든 데이터에 대한 모델 성능 평가를 진행하여 특정 성능 평가용 데이터에서 훈련 모델이 잘 맞거나 맞지 않는 문제를 방지하였다. 육안으로 구별하기 힘든 작은 용종도 찾아내기 위해 확률값(probability)을 0.1로 IoU 값을 0.3으로 설정하였다. 학습 모델에 대해 성능 평가를 총 1,070장의 성능 평가용 데이터로 진행하여 평균적으로 96.2%로 높은 민감도와 0.227의 FPPI의 성능을 도출해냈다.

[표 1]을 보면 FN의 수에 비해 FP의 수가 많은 것을 볼 수 있는데 이는 용종으로 의심되는 부분을 임상의가 확인하여 간과율을 낮추고자 학습 모델이 10% 이상의 확률값을 가지고 용종의 위치를 예측하였고 예측한 값에 대해 성능 평가를 진행하면서 위양성의 수가 많이 검출되는 결과가 초래했다고 판단된다. 또한, IoU값을 0.3으로 설정하여 모델이 예측한 영역과 정답 데이터 영역이 중복되는 영역의 크기가 0.3이상일 경우 정답으로 처리했기 때문에 위음성의 수가 적게 검출되었다고 판단된다.

특정 데이터에서는 실제 용종의 위치 영역보다 넓은 영역을 모델이 예측한 경우가 발생했는데 이는 대장의 특성상 모든 구간의 표면이 고르지 않기 때문에 표면이 매끈한 곳에서 발생하는 용종은 육안으로도 확인하기 쉬운 반면 그렇지 않은 부분인 대장이 접히는 부분에서 발생하는 용종은 육안으로도 확인하기 어렵다. 따라서 [그림 4]의 (a)와 같이 표면이 매끈한 곳에서 발생한 용종에 대해서는 관심 영역으로 설정한 영역과 거의 동일하게 모델이 검출하지만, 그렇지 않은 부분인 대장의 접힌 부분과 인접하게 발생하는 용종에 대해서는 [그림 4]의 (b)와 같이 실제 용종의 위치보다 넓은 영역으로 검출해 내는 것을 볼 수 있었다.



[그림 4] 대장 표면에 따른 모델 검출 결과. 빨간색 box는 실제 용종의 위치 및 영역, 초록색 box는 모델이 예측한 용종의 위치 및 영역

[Fig. 4] Result of model detection depending on colon surface. The red box is the location and area of the actual polyp, and the green box is the location and area of the polyp predicted by the model

5. 결론

본 연구에서는 대장 내시경 진단 시 얻은 정지 영상으로 용종 검출 모델을 개발한 것으로 대장 내시경을 진행하면서 실시간으로 용종을 검출하는 데에 제한점이 있다. 따라서 본 연구에서 개발한 모델을 동영상 데이터에 적용하거나 동영상 데이터를 이용한 학습 모델 개발 연구가 진행된다면 경험이 적은 임상이나 판단하기 힘든 용종에 대해 결정하는데 많은 도움을 줄 수 있을 것으로 기대된다. 또한, 본 연구를 진행할 때 학습 모델의 성능 향상을 위해 전처리나 후처리 과정 없이 데이터의 입력 크기만 동일하게 조정 후 학습을 진행하였다. 추후 대장 내시경 영상에서의 용종 검출이 어려운 요인인 대장의 굴곡면에 대해 영상처리 [11]를 전처리로 진행 후 학습이 이뤄지고 위양성 데이터에 대한 후처리를 적용하여 학습이 이뤄지면, 대장이 접힌 부분에 있는 용종을 검출해내지 못하거나 형태학적 특성상으로 돌출되어 있는 부분들을 용종으로 검출하는 경우가 줄어들고 용종을 검출하는 모델 성능이 향상될 수 있을 것으로 판단된다.

Reference

- [1] K. W. Jung, Y. J. Won, M. J. Kang, H. J. Kong, J. S. Im, H. G. Seo, "Prediction of Cancer Incidence and Mortality in Korea, 2022", *Cancer Res Treat*, vol. 54, no. 2, March 2022, pp. 345-351, doi: 10.4143/crt.2022.179.
- [2] S. Y. Jeong, "Early diagnosis of colorectal cancer: Recommendations for early screening for colorectal cancer", *Korean Journal of Gastrointestinal Endoscopy*, vol. 24, no. 6, January 2002, pp 317-320.
- [3] A. M. Leufkens, M. G. H. van Oijen, F. P. Vleggaar, P. D. Siersema, "Factors influencing the miss rate of polyps in a back-to-back colonoscopy study," *Endoscopy*, vol. 44, no. 5, March 2012, pp. 470-475, doi: 10.1055/s-0031-1291666.
- [4] S. A. Karkanis, D. K. Iakovidis, D. E. Maroulis, D. A. Karras, M. Tzivras, "Computer-aided tumor detection in endoscopic video using color wavelet features", *IEEE transactions on information technology in biomedicine*, vol. 7, no. 3, September 2003, pp. 141-152, doi: 10.1109/TITB.2003.813794.
- [5] X. Zhang, F. Chen, T. Yu, J. An, Z. Huang, J. Liu, W. Hu, L. Wang, H. Duan, J. Si, "Real-time gastric polyp detection using convolutional neural networks", *PloS one*, vol. 14, no. 3, March 2019, doi: 10.1371/journal.pone.0214133.
- [6] J. N. Lee, H. C. Cho, "Automated Polyp Detection System in Colonoscopy using Object DetectionAlgorithm based on Deep Learning", *KIEE*, vol. 70, no. 1, January 2021, pp. 152-157, doi: 10.5370/KIEE.2021.70.1.152.
- [7] W. J. Yun, S. H. Ha, Y. I. Choi, Y. H. Shin, "Deep learning based polyp detection study for polyp diagnostic assistance", *IEIE*, June 30-July 2, 2021, Jeju, Republic of Korea, pp. 1539-1542.
- [8] R. Olivier, H. Cao, "Nearest neighbor value interpolation", *International Journal of Advanced Computer Science and Applications(IJACSA)*, vol. 3, no. 4, November 2021, pp. 25-30, doi: 10.14569/IJACSA.2012.030405.
- [9] T. Lin, P. Goyal, R. Girshick, K. He, P. Dollar, "Focal loss for dense object detection", *International Conference on Computer Vision (ICCV)*, October 22-29, 2017, Venice, Italy, pp. 2980-2988.
- [10] D. Anguita, L. Ghelardoni, A. Ghio, L. Oneto, S. Ridella, "The 'K'in K-fold cross validation", *European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning (ESANN)*. April 25-27, 2012, Bruges, Belgium, pp. 441-446.
- [11] J. H. Lee, B. S. Shin, "An Adaptive Unfolding Method for Visualizing Wrinkled Colon Surface", *Journal of Korea Multimedia Society*, vol. 9, no. 9, September 2006, pp. 1160-1172.