




딥러닝기반 대장내시경 용종 검출에서 작은 용종의 검출 성능분석 및 성능향상 연구

Small Polyp Detection Performance Analysis and Performance Improvement Study in Deep Learning-based Colonoscopy Polyp Detection

하성화*, 신영학**† 

Seong-Hwa Ha and Younghak Shin†

*목포대학교 컴퓨터공학과 학사과정, **목포대학교 컴퓨터공학과 교수

*Bachelor's Course, Department of Computer Engineering, Mokpo National University

**Professor, Department of Computer Engineering, Mokpo National University

요 약

딥러닝 기술의 발전으로 이미지 객체 검출 기술이 여러 분야에서 적용되고 있다. 하지만, 작은 객체 검출 기술의 한계 때문에 기술이 적용된 서비스를 도입하는 데 어려움을 겪고 있다. 특히 사람의 생명과 직결된 의료 영상 분야에서 작은 객체 검출의 어려움은 큰 문제점인 동시에 해결해야 할 숙제이다. 작은 객체 검출의 어려움은 주로 학습 데이터의 편향으로 인해 발생한다. 학습 데이터에는 작은 객체보다는 큰 객체가 상대적으로 많이 포함되어 있기 때문이다. 본 연구에서는 이러한 편향을 줄이기 위해 대장 내시경 이미지의 크기를 축소하여 학습하는 방법과 작은 객체가 포함된 데이터의 양을 늘려 학습하는 오버 샘플링, 이미지 내에 작은 객체 부분을 복사하여 다른 이미지에 붙여 넣는 copy-pasting 방법을 제안한다. 각각의 알고리즘에 대해 대량의 대장 내시경 비디오 데이터 셋으로 비교 평가하였다. 실험결과, 기존의 데이터 셋으로 테스트 하였을 때 0~6cm² 크기의 용종 772개 중 55개만 검출에 성공 하였지만, 이미지 크기를 축소하여 학습하는 방법을 적용하였을 때 772개 중 222개를 검출에 성공하였다.

키워드 : 객체검출, 데이터편향, 대장내시경, 오버샘플링, 딥러닝

Abstract

With the development of deep learning technology, image object detection is being applied in various fields. However, due to the limitations of small object detection, it is difficult to introduce a service to which the technology is applied. In particular, in the field of medical imaging directly related to human life, the difficulty of detecting small objects is a big problem and homework to be solved. It is mainly caused by the bias of the training data. This is because the training data contains relatively more large objects than small objects. In this study, to reduce this bias, the training method by reducing the size of the colonoscopy image, oversampling training by increasing the amount of data containing small objects, and we propose a copy-pasting method. For each method, a large amount of colonoscopy video dataset is used for comparative evaluation. As a result of the experiment, when tested with the existing data set, only 55 of 772 polyps with a size of 0 to 6 cm² are successfully detected, but when the training method by reducing the image size is applied, 222 out of 772 polyps are successfully detected.

Key Words : Object Detection, Data Bias, Colonoscopy, Oversampling, Deep Learning

Received: Dec. 13, 2021
Revised : Feb. 16, 2022
Accepted: Feb. 16, 2022
†Corresponding author
(younghak@mokpo.ac.kr)

본 논문은 2021년도 한국연구재단 연구
자지원사업(2020R1G1A1100926)에서
지원하여 연구하였음.



This is an Open-Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (<http://creativecommons.org/licenses/by-nc/3.0>) which permits unrestricted noncommercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

1. 서 론

2019년 통계청 ‘암으로 인한 사망률’ 자료에 따르면 인구 10만 명당 17.5명이

대장암으로 사망했는데 폐암(36.2명), 간암(20.6명)에 이어 세 번째로 높은 수치이다 [1]. 또한, 2018년 국립암센터 국가암등록사업 연례 보고서에 따르면 대장암은 조기에 발견하면 완치율 90% 이상 완치가 가능하고, 5년 생존율도 75% 정도로 높다 [2]. 따라서, 대장암을 조기에 예방하기 위해 대장 내시경은 필수 요소이다.

대장내시경 검사는 의료전문가가 모니터로 대장 내부를 직접 관찰하고, 암으로 발전할 가능성이 있는 용종은 검사 중 제거할 수 있어 대장암 조기 진단과 예방에 효과적이다. 대장암의 씨앗인 용종을 제거하면 대장암 발생률이 70~80% 감소하고, 사망률 또한 50% 감소한다. 하지만, 크기가 작은 조기 용종은 대장내시경 검사 중 놓칠 가능성이 높다.

논문[3]에서는 500여건 이상의 대장내시경 검사 경험이 있는 각 2명의 의사가 218건의 용종 절제술을 시행했을 때 절제된 용종의 개수는 총 362개였고, 이 중 29건(17.9%)에서 첫 내시경 검사에서 발견하지 못한 용종이 있었고 개수는 51개(14.1%)였다. 이 51개를 용종의 크기별로 나누면 1cm보다 작은 용종 200개 중 22(11.0%)개를 맞추지 못했고, 1~2cm의 용종은 135개 중 25개(18.5%) 2cm보다 큰 용종에 대해서는 27개 중 4(14.8%)개를 맞추지 못하였다. 즉, 작은 용종에 대한 검출능력이 상대적으로 떨어진다.

최근 딥러닝을 이용한 대장 내시경 용종 검출 연구가 진행되고 있다. 논문[4]에서는 용종 탐지를 위한 객체 인식 알고리즘 종류로 YOLOv4를 사용하여 용종 검출률을 높이기 위한 연구를 진행하였다. 검출률을 높이기 위한 방법으로 AutoAugment 데이터 증대 기법을 사용하였다. 그 결과 증대 후 mAP가 증대 전보다 최대 27.44% 높게 나왔다.

논문[5]는 위장관 위치추적에 초점을 둔 연구로 위장관 랜드마크를 식별할 수 있는 CNN 모델을 설계하였고, 이를 이용하여 위장관 교차점을 추정하였다. 실험결과 위장에서 소장으로 변화하는 위장관 교차점은 88%환자가 의심 구역 안에 들어왔으며, 소장에서 대장으로 변화하는 교차점의 경우 100% 환자가 교차점 의심 구역 안에 들어온 것을 확인하였다.

[6]의 연구에서는 대장암의 정확한 발견을 위한 가이드 알고리즘으로, 카메라의 벡터 정보와 용종 정보를 Faster R-CNN을 통해 학습하여 내시경 카메라의 적절한 속도를 가이드해주는 연구를 진행하였고, 정확도 향상을 확인하였다. 이와 같은 객체 검출 기술을 대장 내시경 용종 검출에 적용하면 의사의 진단을 보조하고 놓치는 용종을 줄일 수 있을 것이다.

하지만 기존 딥러닝 기반 용종 검출 연구에서는 용종의 크기 구분 없이 전체 용종 이미지에 대한 성능 분석만을 진행하고 있다. 큰 용종의 경우 대

부분 쉽게 검출이 되지만 작은 용종의 경우 놓치는 경우가 많고 이에 따라 시간이 흐른 뒤 용종이 커지고 대장암으로 발전할 가능성이 커지게 된다.

본 논문에서는 딥러닝 기반 대장 용종 검출에서 작은 용종에 대한 분류 성능을 분석하고, 작은 용종의 검출 성능 향상을 위한 연구를 진행한다. 대장 내시경 학습 데이터 셋에 존재하는 용종 크기의 편향을 줄이기 위해 이미지 증강기법을 통한 학습 이미지의 크기 축소 기법, 작은 객체가 포함된 이미지의 수를 오버샘플링(oversampling)으로 늘리는 기법, 작은 용종 객체를 복사하여 다른 이미지에 붙여 넣는 copy-pasting 기법을 적용한다. 각각의 방법을 객체 검출 모델 학습에 적용한 뒤 대량의 대장 내시경 비디오에 대해 검출 성능을 비교 평가하였다. 객체 검출 모델로는 딥러닝 기반의 Faster R-CNN 기법을 사용하였고, 두 가지 CNN 특징추출기를 이용하여 검출 성능을 비교하였다.

본 논문의 구성은 2장에서 실험에 사용되는 데이터에 대해 설명하고, 3장에서는 제안하는 이미지 편향을 줄이기 기법에 대해 설명한다. 4장에서는 실험 환경 및 각 기법에 대한 용종 검출 실험 결과에 대해서 기술한다. 마지막으로 5장에서는 본 논문의 결론과 향후 연구 방향에 대하여 기술한다.

2. 실험데이터

본 연구에서 사용한 학습 데이터는, MICCAI(Medical Image Computing and Computer-Assisted Intervention) 2015 대장내시경 자동용종검출 챌린지에서 사용되었던 CVC-ClinicDB 데이터이며 연구목적으로 오픈되어 있다 [7]. 데이터 구성은 대장 내시경 이미지 612장과 그에 해당하는 용종 바운딩 박스(bounding box) 레이블 마스크 이미지 612장이다. 학습에 사용하기 전 이미지 증강기법 중 rotation 기법과 수평 및 수직 flip을 적용하여 총 3360장으로 증강하였다. 이미지의 크기는 가로 312, 세로 266 픽셀이다.

모델의 성능 테스트를 위한 테스트 데이터 셋으로는 CVC-Clinic VideoDB가 사용되었으며 총 18개의 내시경 비디오로 구성되어 있다. 테스트 비디오는 총 11954장의 이미지로 구성되어 있으며 가로 384, 세로 288픽셀이며 한 이미지 프레임에 최대 1개의 용종이 포함되어 있으며 11954장 중 용종이 포함된 이미지는 10025장이다. 그림1은 학습에 사용된 원본 이미지와 해당 이미지의 바운딩 박스(bounding box) 레이블 마스크 이미지 예시이다.

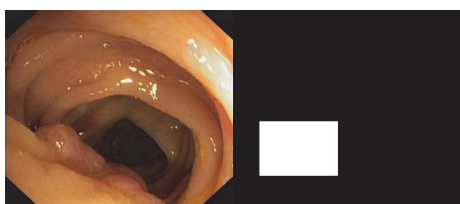


그림 1. 학습이미지 및 객체 바운딩 박스 레이블 마스크
Fig. 1. Examples of wet images and object bounding box label masks

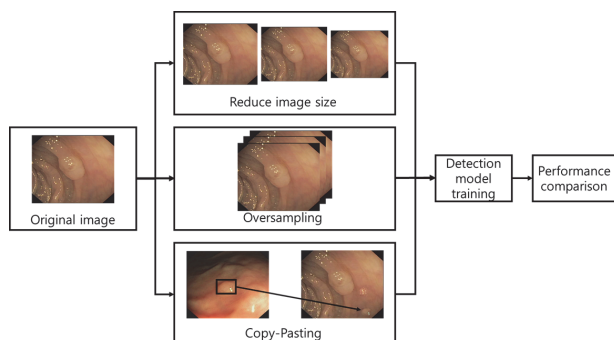


그림 2. 작은 용종검출 성능향상을 위한 3가지 기법 적용
Fig. 2. Application of three techniques to improve the performance of small polyp detection

3. 제안하는 방법

3.1 딥러닝기반 용종검출 방법

본 연구에서 용종 검출을 위한 딥러닝 기반 객체 검출 모델로서 Faster R-CNN(region-based CNN)을 사용하였다. Faster R-CNN [8]은 기존의 R-CNN의 개량된 모델이다. R-CNN에서 객체 후보 영역을 생성하는 데에 사용된 별도의 선택적 검색 알고리즘을 이용하지 않고, 특징 맵을 추출하는 CNN의 마지막 층을 객체 후보 영역을 생성하는 영역 제안 네트워크(region proposal network)로 사용하였다. 이처럼 특징 맵을 추출하는 CNN 과정과 후보 영역을 생성하는 과정을 하나의 네트워크로 구성함으로써, 훈련시간을 감소시키고, 정확도 또한 향상시켰다.

본 연구에서는 Faster R-CNN 검출 방법의 특징추출기로 InceptionV2, ResNet101 두 가지 특징추출기를 사용하였고, MS COCO 데이터로 사전 학습된 모델을 사용하였다.

3.2 작은 용종 검출의 어려움

이미지 객체 검출에서 작은 객체 검출의 어려움은 학습에 사용되는 이미지 데이터의 편향으로 인해 주로 발생한다. 학습에 사용되는 이미지들은 작은 물체보다 크고 눈에 띄는 명확한 물체가 상대

적으로 더 많기 때문이다.

표 1은 본 연구의 학습에 사용되는 용종 이미지를 용종의 크기별로 분류한 표이다. 용종 크기는 바운딩 박스 레이블 마스크의 넓이로 측정하였다. 실제로 학습에 사용하는 이미지들의 용종 크기가 대부분 6cm²이상인 큰 크기의 용종이었다. 이처럼 학습 데이터 셋에 작은 크기의 용종이 큰 용종보다 훨씬 적기 때문에 이를 이용해 학습된 용종 검출 모델이 작은 크기의 용종을 많이 놓치게 된다.

표 1. 용종 크기별 학습 이미지 수
Table 1. Number of training images by polyp size

polyp size(cm ²)	number of images
0~2	2
2~4	37
4~6	75
6~	446

아래 표 2는 주어진 3360장의 학습 이미지 데이터로 학습하였을 때 테스트 데이터 셋에 대해서 용종 크기별로 검출 모델이 제대로 검출해낸 용종의 개수이다. 예상대로 두 모델 모두 작은 용종에 대해서 검출률이 매우 낮은 것을 볼 수 있다.

표 2. 테스트데이터 셋에 대해 용종 크기별로 놓치는 용종의 수
Table 2. The number of polyps missing by polyp size in the test dataset

polyp size(cm ²)	InceptionV2 (correct/all)	ResNet101 (correct/all)
0~2	0/32	0/32
2~4	0/237	18/237
4~6	14/503	37/503
all	4803/10025	5469/10025

그림 2는 본 연구에서 제안한 주어진 학습 데이터 셋의 용종 크기별 데이터 편향을 줄여 작은 객체에 대한 성능을 향상시키기 위한 세 가지 기법에 대한 도식도이다. 이미지 크기를 축소하여 전체 데이터 셋의 객체의 크기를 줄이는 이미지 축소, 작은 크기의 용종이 포함된 이미지의 수를 늘리는 오버샘플링, 학습 이미지 내의 작은 객체를 복사하여 다른 이미지에 붙여 넣는 copy-pasting 기법을 통해 작은 용종의 수를 상대적으로 늘려 학습한 뒤, 각각의 성능을 두 가지 검출 모델에 대해 비교 평가한다.

3.3 이미지 크기 축소

그림 2의 이미지 축소 부분과 같이 학습에 사용되는 전체 이미지의 크기를 줄여서 학습에 사용하였다. 이미지의 크기를 줄이면 이미지 속에 포함된 객체의 크기 또한 작아져 데이터 셋의 모든 용종

의 크기를 줄이는 효과가 있다. 원본 학습 이미지의 크기를 각각 10%, 20%, 30%로 줄여서 학습에 사용하였다.

표 3. 이미지 축소 후 용종 크기별 학습 이미지 수
Table 3. Number of training images by polyp size after image reduction (10%, 20%, 30%)

polyp size(cm ²)	Original	10% ↓	20% ↓	30% ↓
0~2	2	5	19	41
2~4	37	66	107	145
4~6	75	95	141	163
6~	446	394	283	200

표 3은 학습 원본 이미지를 축소하였을 때 용종 크기별 개수를 나타낸 것이다. 원본 이미지와 비교할 때, 이미지의 크기를 축소할수록 작은 용종이 포함된 이미지의 수가 증가하는 것을 확인할 수 있다. 최종 성능 평가에는 10%, 20%, 30% 줄인 각각의 데이터 셋과 이 세 가지를 모두 더한 총 네 가지 데이터 셋을 학습에 사용하였다.

3.4 오버샘플링 (Oversampling)

오버샘플링은 학습 데이터 셋에서 상대적으로 작은 크기의 용종이 포함된 이미지를 학습에 더 많이 사용하는 방법이다. 우리는 간단하게 작은 용종이 포함된 이미지들을 여러 개 복제한 뒤 기존의 학습 데이터 셋에 추가하여 학습에 사용하였다. 기존 학습 데이터 셋 3360장에서 3cm²보다 작은 용종이 포함된 177장의 이미지를 두 번 복사해서 더한 3714장, 세 번 복사해서 더한 3891장, 네 번 복사해서 더한 4068장, 총 네 가지 데이터 셋을 학습에 사용하여 성능을 비교 평가하였다.

3.5 Copy-Pasting

Copy-pasting 기법에서는 기존 원본 학습 이미지에서 0~2, 2~4, 4~6cm² 크기 별로 추출한 용종을 1~3개씩 랜덤으로 다른 이미지에 붙여 넣었다. 용종을 copy 할 때는 용종들의 segmentation 맵을 사용하여 용종 부분만 copy하였고, 이를 다른 이미지의 임의의 위치에 붙여 넣었다. 이때, 붙여 넣는 이미지 내의 이미 존재하는 용종과 복사된 용종들이 겹치지 않도록 각 객체에 대한 IoU(Intersection over Union)를 계산하였고, 계산한 값이 0보다 크다면 다른 위치에 붙여지도록 하였다. 그림 3은 copy-pasting으로 만들어진 예시 이미지이다. 이미지 생성 후 이미지에 대한 바운딩 박스 레이블 마스크를 오른쪽과 같이 추출하였다.

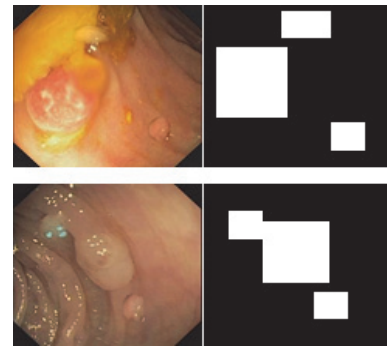


그림 3. Copy-pasting으로 생성된 예시 이미지와 bounding box 레이블 마스크

Fig. 3. Example images created by Copy-pasting and binding box label masks

4. 실험 결과

4.1 실험 환경

본 논문에서 진행한 딥러닝 기반 용종 객체 검출 실험은 공정한 비교를 위해 동일한 PC 환경에서 진행하였다. Anaconda 가상 환경(Python 3.7)에서 실험하였으며 딥러닝 기반 객체 검출 모델 학습을 위한 딥러닝 프레임워크로는 Tensorflow object detection api (Faster R-CNN)를 사용하였다. 기타 이미지 축소, 오버샘플링, copy-pasting 등의 알고리즘은 Python을 이용하여 직접 구현하였다. 실험에 사용된 GPU는 NVIDIA RTX 2080 SUPER (8GB 메모리)를 사용하였다. 모델 학습은 10ep을 진행하였고, InceptionV2 모델의 경우 1ep당 약 8분, ResNet101 모델은 1ep당 20분이 소요되었다.

4.2 용종 검출 성능비교

본 논문에서의 용종 검출 성능 평가를 위하여 TP(true positive), FP(false positive), FN(false negative)을 다음과 같이 정의하였다.

TP: 모델의 검출결과 바운딩 박스와 실제정답 바운딩 박스 사이의 IoU값이 50%이상인 경우

FP: 모델의 검출결과 바운딩 박스와 실제정답 바운딩 박스 사이의 IoU값이 50%미만인 경우

FN: 이미지 내에 용종 정답 바운딩 박스가 존재하지만, 모델의 검출 결과에는 없는 경우

위에서 정의한 TP, FP, FN을 이용하여 Recall(Re), Accuracy(Acc), F1-score(F1)를 최종 검출 성능 평가 지표로 사용하였다.

$$Recall = \frac{TP}{TP + FN} \quad (1)$$

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad (2)$$

$$F1-score = \frac{Pre \times Rec}{Pre + Rec} \quad (3)$$

4.3 원본 학습 이미지 결과

먼저 성능 비교를 위해 원본 학습 이미지로 학습한 결과를 확인해 보았다. 표 4는 원본 데이터셋에 이미지 증강이 적용된 기본 3360이미지만으로 학습한 뒤 테스트 비디오로 평가한 결과이다. 테스트 결과를 용종의 크기별로 나타내어 성능을 비교하였다. Faster R-CNN 기반의 두 모델(ResNet101, Inception v2)의 전체 F1-score는 각각 0.681, 0.606으로 측정되었다. 또한, 테스트 데이터셋의 용종의 크기가 작을수록 낮은 검출률을 보였다. 두 모델 모두 작은 크기인 $0 \sim 2_{cm^2}$ 크기의 용종은 하나도 검출해 내지 못하였다.

4.4 이미지 크기 축소 기법 결과

표 5는 원본 이미지로 학습하였을 때와 학습 이미지를 10%, 20%, 30% 축소하였을 때 그리고 세 가지 축소한 이미지를 모두 합하여 학습하였을 때에 대한 성능을 비교하여 나타낸 표이다.

이미지 크기를 축소시킴으로서 전체 객체의 크기가 줄어 작은 객체에 대한 성능 향상을 확인할 수 있었다. 두 가지 검출 모델에 대해 가장 좋은 성능은 10%, 20%, 30% 축소한 세 가지 데이터셋을 모두 더하여 사용한 것으로 원본 데이터셋보다 모든 구간에서 F1-score 성능이 증가하였고, 특

히 Faster R-CNN inception_v2 모델의 경우 전체 F1-score가 원본 학습 이미지로만 평가했을 때와 비교하여 상대적으로 28% 증가하였다.

표 4. 원본 데이터셋에 대한 테스트비디오 검출 결과
Table 4. Test video detection result for the original data set

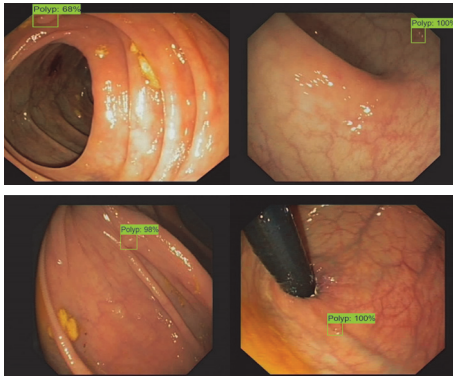
Object size (cm ²)	Eval. Index	Faster R-CNN (resnet 101)	Faster R-CNN (inception_v2)
0 ~ 2	Re/TP	0 / 0	0 / 0
	Acc	0.774	0.86
	F1	0	0
2 ~ 4	Re/TP	0.075 / 18	0 / 0
	Acc	0.006	0
	F1	0.113	0
4 ~ 6	Re/TP	0.073 / 37	0.027 / 14
	Acc	0.057	0.024
	F1	0.109	0.047
all	Re/TP	0.585 / 5469	0.479 / 4803
	Acc	0.523	0.51
	F1	0.631	0.606

그림 4는 원본 이미지 데이터로 학습하였을 때는 검출하지 못했던 작은 용종을 10%, 20%, 30% 축소한 세 가지 데이터셋을 모두 사용한 축소 기법 적용 후 검출해낸 이미지 예시이다.

표 5. 이미지 크기 축소 후 테스트비디오 검출결과
Table 5. Test video detection result after reducing the image size

Modal Name		Faster R-CNN(resnet 101)					Faster R-CNN(inception_v2)				
Object size (cm ²)	Eval. index	Original	10% ↓	20% ↓	30% ↓	10%, 20%, 30% ↓	Original	10% ↓	20% ↓	30% ↓	10%, 20%, 30% ↓
0 ~ 2	Re / TP	0 / 0	0 / 0	0 / 0	0 / 0	0 / 0	0 / 0	0 / 0	0 / 0	0 / 0	0.003/1
	Acc	0.774	0.778	0.836	0.824	0.79	0.86	0.835	0.82	0.866	0.778
	F1	0	0	0	0	0	0	0	0	0	0.004
2 ~ 4	Re / TP	0.075 / 18	0.101 / 24	0.13 / 31	0.135 / 32	0.126 / 30	0 / 0	0.004 / 1	0.042 / 10	0.08 / 19	0.194/46
	Acc	0.006	0.091	0.126	0.113	0.121	0	0.004	0.04	0.078	0.186
	F1	0.113	0.167	0.223	0.203	0.217	0	0.008	0.078	0.145	0.315
4 ~ 6	Re / TP	0.073 / 37	0.121 / 61	0.216 / 109	0.28 / 141	0.381 / 192	0.027 / 14	0.083 / 42	0.107 / 54	0.196 / 99	0.483/243
	Acc	0.057	0.105	0.2	0.25	0.351	0.024	0.072	0.096	0.181	0.438
	F1	0.109	0.191	0.333	0.4	0.519	0.047	0.135	0.175	0.307	0.609
all	Re / TP	0.585 / 5469	0.631 / 6332	0.653 / 6547	0.681 / 6829	0.715 / 7173	0.479 / 4803	0.584 / 5862	0.603 / 6049	0.624 / 6257	0.744/6892
	Acc	0.523	0.611	0.651	0.657	0.699	0.51	0.594	0.606	0.641	0.68
	F1	0.631	0.716	0.749	0.756	0.793	0.606	0.696	0.708	0.737	0.814

그림 4. 이미지 축소기법 적용 후 작은 용종 검출 성공 이미지
Fig. 4. After applying the image reduction technique, a small polyp detection successful image.



4.5 Oversampling 기법 결과

표 6은 기존 데이터 셋 3360장에 3cm²보다 작은 용종이 포함된 이미지 177장을 뽑아 oversampling 기법을 적용하여 평가한 결과이다. 원본 이미지 데이터 셋과 3cm²보다 작은 용종 이미지 177장을 두 번, 세 번, 네 번 추가한 데이터 셋에 대하여 학습한 뒤 테스트 비디오에 대해서 비교 평가하였다. 177장을 2번 복사한 데이터 셋에서 가장 좋은 성능을 보여주었지만, 원본 데이터 셋과 비교했을 때

성능 차이가 거의 나지 않았다. 2~4cm² 크기에서 검출률이 소폭 향상했지만, 전체적인 검출률은 오히려 줄어들었다. 이는 원본 이미지에 작은 용종 이미지 자체가 매우 적어서 이를 오버샘플링하여도 성능 증가 효과 보다는 오히려 과적합이 발생한 것으로 판단된다.

4.6 Copy-Pasting 기법 결과

표 7은 학습 이미지 내의 작은 용종 부분만을 복사하여 다른 이미지에 붙여 넣는 copy-pasting 기법에 대한 결과이다. 원본 이미지 데이터 셋과 0~2cm², 2~4cm², 4~6cm² 각 세 가지 크기 범위의 용종을 copy-pasting 하여 만든 이미지 데이터 셋에 대하여 비교 평가하였다. 작은 사이즈 용종 검출 검출률이 증가하는 경우도 있지만, 전체 사이즈에 대한 검출 결과, 두 검출 모델 모두에서 원본 데이터 셋과 비교했을 때 오히려 성능이 줄어드는 결과가 나왔다. 이는 단순 copy-pasting 기법으로 생성한 이미지의 용종 부분이 실제 학습 이미지와 다르게 부자연스럽기 때문에 테스트 비디오에 존재하는 실제 용종 검출에는 도움이 되지 못하는 것으로 판단된다.

표 6. Oversampling 테스트비디오 검출 결과
Table 6. Test video detection result for oversampling

Modal Name		Faster R-CNN(resnet 101)				Faster R-CNN(inception_v2)			
Object size(cm ²)	Eval. index	Original	2×	3×	4×	Original	2×	3×	4×
0 ~ 2	Re / TP	0 / 0	0 / 0	0 / 0	0 / 0	0 / 0	0 / 0	0 / 0	0 / 0
	Acc	0.774	0.832	0.773	0.759	0.86	0.773	0.831	0.6
	F1	0	0	0	0	0	0	0	0
2 ~ 4	Re / TP	0.075 / 18	0.131 / 19	0.059 / 14	0.042 / 10	0 / 0	0.008 / 2	0.008 / 2	0.008 / 2
	Acc	0.006	0.113	0.044	0.029	0	0.006	0.007	0.006
	F1	0.113	0.204	0.085	0.057	0	0.012	0.015	0.012
4 ~ 6	Re / TP	0.073 / 37	0.113 / 51	0.093 / 47	0.095 / 48	0.027 / 14	0.063 / 32	0.075 / 38	0.117 / 59
	Acc	0.057	0.1	0.079	0.067	0.024	0.045	0.062	0.084
	F1	0.109	0.182	0.146	0.126	0.047	0.087	0.118	0.155
all	Re / TP	0.585 / 5469	0.551 / 5527	0.488 / 4894	0.523 / 5253	0.479 / 4803	0.496 / 4980	0.457 / 4589	0.546 / 5482
	Acc	0.523	0.544	0.467	0.48	0.51	0.469	0.475	0.485
	F1	0.631	0.648	0.572	0.59	0.606	0.575	0.571	0.608

표 7. Copy-Pasting 테스트비디오 검출 결과
Table 7. Test video detection result for Copy-Pasting

모델명		Faster R-CNN(resnet 101)				Faster R-CNN(inception_v2)			
Object size(cm ²)	Eval. index	Original	0~2cm ² copy	2~4cm ² copy	4~6cm ² copy	Original	0~2cm ² copy	2~4cm ² copy	4~6cm ² copy
0 ~ 2	Re / TP	0 / 0	0 / 0	0 / 0	0 / 0	0 / 0	0 / 0	0 / 0	0 / 0
	Acc	0.774	0.859	0.852	0.876	0.86	0.872	0.876	0.893
	F1	0	0	0	0	0	0	0	0
2 ~ 4	Re / TP	0.075 / 18	0 / 0	0 / 0	0.08 / 19	0 / 0	0 / 0	0 / 0	0 / 0
	Acc	0.006	0	0	0.08	0	0	0	0
	F1	0.113	0	0	0.148	0	0	0	0
4 ~ 6	Re / TP	0.073 / 37	0.003 / 2	0.005 / 3	0.047 / 24	0.027 / 14	0.005 / 3	0.001 / 1	0.017 / 9
	Acc	0.057	0.003	0.005	0.047	0.024	0.005	0.019	0.017
	F1	0.109	0.007	0.011	0.09	0.047	0.011	0.038	0.033
all	Re / TP	0.585 / 5469	0.39 / 3911	0.386 / 3872	0.495 / 4963	0.479 / 4803	0.339 / 3403	0.374 / 3755	0.433 / 4349
	Acc	0.523	0.443	0.444	0.551	0.51	0.41	0.445	0.497
	F1	0.631	0.527	0.527	0.646	0.606	0.48	0.524	0.585

5. 결론 및 향후 연구

본 연구에서 우리는 딥러닝 기반 대장 내시경 용종 검출에서 용종 크기별로 검출 성능에 대하여 분석하였고, 그 결과 크기가 작은 용종에 대한 검출 성능이 현저히 낮다는 것을 발견하였다. 또한, 이를 극복하기 위한 세 가지 기법을 제안하였고 각각에 대해 대량의 대장 내시경 비디오 데이터 셋을 이용하여 용종 검출 성능을 비교평가 하였다. 그 결과 학습이미지 중에서 작은 크기의 용종 이미지에 대한 오버샘플링(Oversampling) 기법과 작은 용종 객체 부분 copy-pasting 기법에서는 성능 개선을 확인하지 못하였다. 반면에, 간단한 기법인 학습 이미지 사이즈 축소 방법에서 두 가지 검출 모델 모두에서 성능 향상을 확인하였다. 특히 이미지 축소 시 다양한 크기로 축소한 것을 함께 사용 한 경우가 가장 큰 성능 향상이 있는 것을 확인하였다. 향후 연구에서는 단순 copy-pasting이 아닌 이미지 생성 모델 기반의 실제와 같은 작은 용종 객체 생성을 통한 성능 평가를 진행할 예정이다.

Conflict of Interest

저자는 본 논문에 관련된 어떠한 잠재적인 이해 상충도 없음을 선언한다.

References

- [1] Korea National Statistical Office (KOSTAT), "Cancer incidence and mortality", Available: [https:// www. index. go. kr/ potal/ main/ EachDtlPageDetail. do?idx_cd=2770](https://www.index.go.kr/potal/main/EachDtlPageDetail.do?idx_cd=2770), 2020, [Accessed: February 22, 2022]
- [2] National Cancer Center, "National Cancer Registration Project Annual Report", [https://ncc. re. kr/ cancer StatsView. ncc?bbsnum=558&searchKey=total&searchValue=&pageNum=1](https://ncc.re.kr/cancerStatsView.ncc?bbsnum=558&searchKey=total&searchValue=&pageNum=1), 2018, [Accessed: February 22, 2022]
- [3] Choon Sik Chung, Jin Seok Yoon, Yong Geul Joh, Yoon Jung Cha, Kyung Jo Kim, Eui Ryun Park, Seon Han Kim, and Dong Kun Lee, "Colonoscopic Missing Rate of Colorectal Polyps," *Journal of the Korean Society of Gastrointestinal Endoscopy*, no. 28. pp. 179-182, 2004.
- [4] Hyeonwoong Jang, Changnam Lim, Yesuel Park, Gwangjae Lee and Junwon Lee, "Estimating Gastrointestinal Transition Location Using CNN-based Gastrointestinal Landmark Classifier," *Journal of the KIPS Trans. Softw. and Data Eng.*, vol. 9, no. 3, pp. 101-108, 2020.
- [5] Sungjong Eun, Eunyong Jung and Dongkyun Park "Development of Colonoscopy Guide Algorithm based on Fast R-CNN," *Journal of Next-genescoren Convergence Information Services Technology*, vol. 8, no. 2, pp. 213-222, 2019.
- [6] Hyunjong Cho and Jungnam Lee, "Colonoscopy polyp

detection system using deep learning-based object detection algorithms,” *the paper of the Society of Electrical Sciences*, vol. 70, no. 1, pp. 152-157, 2021.

- [7] Jorge Bernal and Aymeric Histace, “Endoscopic Vision Challenge,” <https://polyp.grand-challenge.org/>, 2015, [Accessed: February 22, 2022]
- [8] Shaoqing Ren, Kaiming He, Ross Girshick and Jian Sun, “Faster R-CNN: Towards real-time object detection with region proposal networks,” in *Proc. Adv. Neural Inf. Process. Syst.*, Montreal, Canada, pp. 91-99, Dec. 2015.

저 자 소 개



하성화(Sung-hwa Ha)

2022년 : 목포대학교 컴퓨터공학과 공학사
예정

관심분야 : Artificial Intelligence, Deep Learning,
Smart Farm
E-mail : sl63342@ce.mokpo.ac.kr



신영학(Youn-ghak Shin)

2009년 : 광운대학교 전자통신공학과
공학사
2011년 : 광주과학기술원 정보통신공학부
공학석사
2016년 : 광주과학기술원 정보통신공학부
공학박사
2018년~2020년 : LGCNS
AI빅데이터연구소
책임연구원
2020년~현재 : 목포대학교 컴퓨터공학과
교수

관심분야 : Artificial Intelligence, Deep Learning,
Smart Farm
ORCID : 0000-0003-2213-5232
E-mail : younghak@mokpo.ac.kr