

# 위내시경 영상에서의 위 병변 자동 검출 모델 개발을 위한 RetinaNet 기반 backbone 네트워크에 따른 학습 성능 비교

## Comparison of Training Performance According to RetinaNet-based Backbone Network for Development of Automatic Gastric Lesion Detection Model in Gastroscopy Images

이기표<sup>1</sup>, 김영재<sup>2</sup>, 박동균<sup>3</sup>, 김재승<sup>4</sup>, 김광기<sup>5\*</sup>

Gi Pyo Lee<sup>1</sup>, Young Jae Kim<sup>2</sup>, Dong Kyun Park<sup>3</sup>, Jae-Seoung Kim<sup>4</sup>, Kwang Gi Kim<sup>5\*</sup>

### 요약

본 연구에서는 위내시경 검사 시에 보조 시스템으로 활용할 수 있도록 RetinaNet 네트워크를 사용하여 위내시경 영상에서의 위 병변의 위치를 자동으로 검출하는 모델을 개발하였다. 위암은 한국이나 일본 등의 아시아권에서 대부분 발생한다. 그러나 위내시경 검사는 동시에 진단이나 치료할 수 있으며, 조기 발견 시 치료 성공확률이 매우 높다. 그러나 실시간으로 진행되는 검사 특성상 숙련도나 경험이 결과에 영향을 주며, 업무의 피로도 상승과 집중력 하락으로 인해 검사의 정확도가 낮아지게 된다. RetinaNet 기반의 backbone 네트워크로 ResNet50, ResNet152, EfficientNetB0, EfficientNetB4 네트워크를 사용하여 학습한 모델의 검출 성능을 확인하고, 각 모델 간의 성능을 비교하였다. RetinaNet 기반 backbone 네트워크별 모델들의 평균 민감도(FP/images)는 ResNet50 73.72%(0.0489), ResNet152 78.26%(0.0458), EfficientNetB0 79.67%(0.03268), EfficientNetB4 62.66%(0.0448)를 보였다. EfficientNetB0 네트워크는 가장 높은 민감도를 나타냈으나 FP/images가 매우 높게 나타나 두 성능치를 모두 만족하는 네트워크는 ResNet152였다.

핵심어 : 위내시경 영상, 위 병변, 검출, 딥러닝, 레티나넷

- 1 Department of Health Sciences and Technology, GAIHST, Gachon University, Incheon, Korea [Graduate Student]  
e-mail: pigyo123@gachon.ac.kr
  - 2 Department of Biomedical Engineering, College of IT Convergence, Gachon University, Seongnam, Korea [Associate professor]  
e-mail: youngjae@gachon.ac.kr (Corresponding author)
  - 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, Republic of 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)
- \* 이 논문은 2022년도 정부(과학기술정보통신부)의 재원으로 정보통신산업진흥원의 지원을 받아 수행된 연구임 (No.S0252-21-1001, AI정밀의료술루션(닥터앤서2.0)개발).

Received(December 7, 2022), Review Result(1st: December 22, 2022, 2nd: January 10, 2023), Accepted(April 12, 2023), Published(April 30, 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/>.

## Abstract

Gastric cancer occurs mostly in Asian countries such as Korea and Japan. Gastroscopy allows diagnosis and treatment of gastric cancer at the same time, and the probability of successful treatment is very high at early detection. However, due to the nature of the inspection which progresses in real time, proficiency and experience of the clinician affect the results, and the accuracy of the inspection can decrease due to increased work fatigue and decreased concentration. In this study, we developed a model that automatically detects the regions of gastric lesion in gastroscopic images using the RetinaNet network so that it can be used as an auxiliary system during gastroscopy. We confirmed the detection performance of models trained using ResNet50, ResNet152, EfficientNetB0, and EfficientNetB4 networks in a RetinaNet-based backbone network, and compared the performance between each model. The average sensitivities (FP/images) of RetinaNet-based backbone network-models were 73.72% (0.0489) for ResNet50, 78.26% (0.0458) for ResNet152, 79.67% (0.3268) for EfficientNetB0, and 79.67% (0.3268) for EfficientNetB4. The EfficientNetB0 network showed the highest sensitivity, but the FP/images were very high, so the network satisfying both performance values was ResNet152.

Keyword : Gastroscopy image, gastric lesion, detection, deep learning, RetinaNet

## 1. 서론

위암은 2018년 기준 전 세계에서 5번째로 많이 발생한 암이며, 아시아권에서 전체 발생 환자 중 3/4을 차지하는 질환이다. 2019년 대한민국 보건복지부 암 등록통계에 따르면 위암 발생자 수는 29,493명으로 전체 암 환자 중 11.6%를 차지하였으며, 갑상선암과 폐암 다음으로 국내에서 3번째로 많이 발생한 암종으로 나타났다 [1][2]. 위내시경 검사는 검사 도중 발견한 병변이 암 또는 궤양 등의 이상소견으로 의심되면 해당 부위에 대한 조직 검사를 즉시 수행할 수 있어 조기암 진단을 위해 효과적인 검사 방법이다. 이에 대한민국에서는 2002년도부터 국가 건강검진을 통해 40세 이상의 사람들에게 2년마다 위장 조영촬영 또는 위내시경 검사를 지원하고 있다 [3]. 그러나 실시간으로 검사를 진행하면서 이상소견 관찰이 이루어져야 하므로 검사를 수행하는 의사의 숙련도나 경험이 결과에 영향을 주기도 하며, 한 명의 의사가 수차례 검사를 진행하게 되면 업무의 피로도 상승과 집중력 하락으로 인해 위 병변을 발견하지 못하고 지나치는 경우도 발생하여 검사결과에 대한 정확도가 낮아지게 된다 [4][5].

이러한 의사의 주관적 판단에 의한 검사의 정확도를 향상하기 위해 다양한 컴퓨터 보조진단 시스템 개발이 이루어지고 있으며, 위내시경 검사뿐만 아니라 대장내시경 검사에서도 다양한 인공지능 및 영상처리 기술을 활용한 연구가 활발히 이루어지고 있다. Park 등은 위내시경 이미지에서 ResNet50, InceptionV3, Vgg16 모델과 전이학습을 활용하여 정상과 비정상 영상을 분류하는 모델을 개발하였으며, 전이학습을 사용한 ResNet50 모델 학습은 0.974의 정확도를 나타냈다 [5]. Lee 등은 위내시경 영상에서 정상과 위궤양 영상을 ResNet50 모델을 활용하여 개발하였으며, 90.16%의 정확도, 0.97의 Area under curve(AUC) 점수를 나타냈다 [6]. Eun 등은 대장내시경 환자의 동영상 데이

터에서 Fast R-CNN 모델을 적용하여 대장용종 유무에 대한 인식에 관한 연구를 진행하였다 [7]. 이들은 임상상의 대장용종 확인 결과와 제안한 방법을 통해 예측한 결과와의 비교를 통해 96.8%의 정확도를 보였다. 하지만 위에 제안된 방법은 정확한 병변 위치에 대한 검출이 아닌 4 분할된 영역 중 어느 영역에 존재한다는 결과만을 제시하여 실제 진단에 활용하기에는 한계점을 보인다.

본 연구에서는 실시간 위내시경 검사에서 활용할 수 있는 위병변 검출 모델 개발을 진행하였으며, 딥러닝 기술을 활용하여 RetinaNet 네트워크를 사용하여 위내시경 영상에서의 위 병변의 위치를 자동으로 검출하는 모델 학습을 진행하였다. RetinaNet 기반의 모델 학습 과정에서 피쳐 추출을 위해 다양한 네트워크를 backbone 네트워크로 사용할 수 있다. 다양한 분류 학습모델에서 피쳐 추출을 위해 활용되고 있는 ResNet50, ResNet152, EfficientNetB0, EfficientNetB4을 각각 사용하여 학습 모델 성능을 확인하고, 성능 결과를 비교하고자 한다.

## 2. 본론

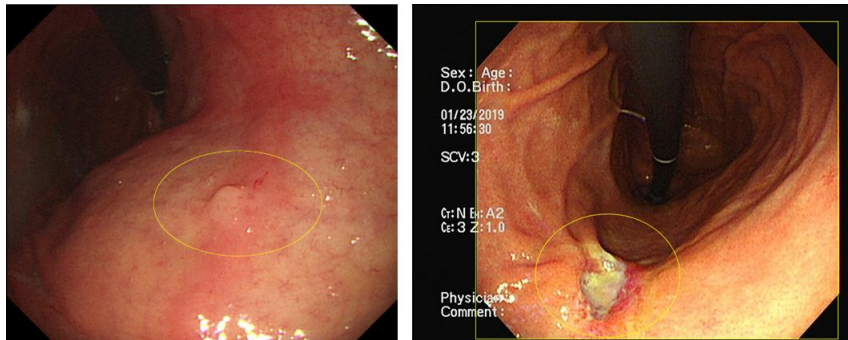
### 2.1 데이터

본 연구는 위내시경 영상에서의 병변 검출 학습모델 개발을 위해 가천대 길병원에서 위내시경 정지 영상 데이터를 수집하였다. 가천대 길병원을 방문하여 2019년 1월부터 2021년 12월까지 위내시경 검사를 받은 성인 환자를 대상으로 총 937명의 환자에 대한 위내시경 정지 영상 데이터를 수집하였다. 위내시경 검사 시에 암 또는 궤양으로 판단될 때 위내시경 영상을 캡처하여 이미지를 시스템에 저장하게 된다. 수집한 937명의 환자 중 위암을 판정받은 환자 461명, 위궤양을 판정받은 환자 330명, 정상 환자 146명에 대하여 저장된 정지 영상 이미지를 확인하였다. 위암이 나타난 영상 2,497장, 위궤양이 나타난 영상 1,988장, 정상 환자에게서 촬영한 영상 5,015장을 수집하여 총 9,500장의 위내시경 정지 영상을 수집하였다.

### 2.2 데이터 레이블링

본 연구에서는 위내시경 영상에서의 위암 또는 위궤양 검출 학습모델 개발을 위하여 위암 또는 위궤양 영역에 대한 레이블링 작업을 수행하였다. 환자의 검사에 대한 판독문 결과를 참고하여 수집한 위내시경 영상에서 발견된 위암 또는 위궤양 영역을 Image J (NIH, Bethesda, MD, USA) 프로그램을 사용하여 타원형(ellipse) 형태로 전문가가 수동으로 표시하여 수집하였다. 수집한 데이터 중 일부 데이터는 영상 주위에 검은 영역이 포함되거나 환자 ID, 이름 등 환자 정보를 표시하기 위한 영역 등 학습 과정에서 방해가 될 수 있는 불필요한 영역을 포함하고 있다. 이를 제거하기 위해 전처리 단계에서 위내시경 영상 영역만을 추출하여 사용하기 위한 사각형(box) 형태의 레이블링

작업을 동시에 수행하였다. [그림 1]은 위 병변과 영상 영역에 대한 레이블링 작업의 예시를 나타내고 있다.

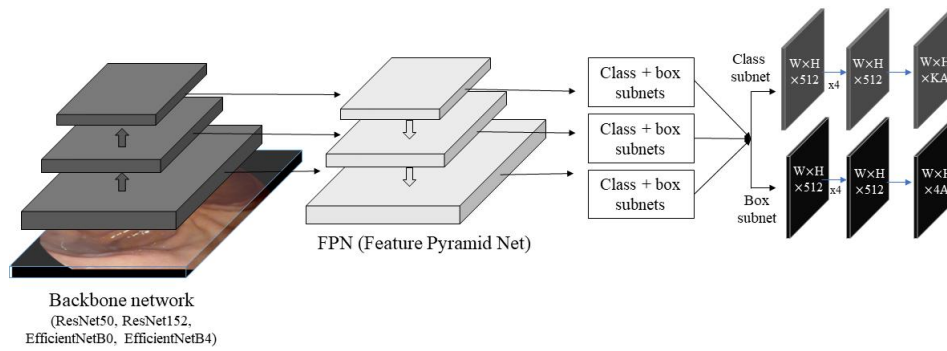


[그림 1] 위내시경 영상에서의 위 병변과 영상 영역에 대한 레이블링 작업 예시

[Fig. 1] Example of data labeling for gastric lesions and video regions in gastroscopy images

### 2.3 딥러닝 기반 위내시경 영상에서의 위 병변 검출 모델 개발

본 연구에서는 위 병변 검출 모델 학습을 위해 RetinaNet 모델을 사용하였으며 모델 학습 과정에서 특징 추출 단계에서 적용되는 backbone 모델을 다양하게 사용하여 학습모델의 성능을 비교하고자 한다. RetinaNet 모델은 backbone 네트워크와 두 개의 하위 네트워크로 구성되어 있다. backbone 네트워크는 입력된 이미지 전체에서 피쳐 맵(feature map)을 추출하기 위한 컨벌루션 네트워크(convolutional network)이며, 모델 아키텍처는 [그림 2]와 같이 구성하였다. 두 개의 하위 네트워크는 객체 분류(classification)를 수행하는 네트워크와 bounding box regression를 수행하는 네트워크로 이루어져 있다. bounding box regression은 backbone 모델을 통해 pyramid level 별로 추출된 피쳐 맵을 통해 검출한 영역에 대한 4개의 좌표값 ( $x, y, w, h$ )을 얻게 되며, 분류 네트워크는 검출 영역에 대한 클래스를 분류하도록 한다. 또한 Feature Pyramid Networks (FPN)를 적용하여 backbone 네트워크에서 추출한 다양한 scale의 피쳐 맵에 대한 정보를 활용하여 모델의 성능을 향상하였다 [8]. RetinaNet 모델은 DenseNet, EfficientNet [9], mobileNet, ResNet [10], Squeeze and excitation networks (SENet), Vgg 네트워크를 backbone으로 사용할 수 있다. 여러 backbone 네트워크 중 위내시경 분류 학습에서 높은 성능을 보이는 네트워크 모델이 피쳐 추출 단계에서 큰 효과를 보이기 때문에 선행된 연구 중 ResNet과 EfficientNet 네트워크를 비교군으로 선정 [5][6][11]하고 각 모델의 하위 모델 중 모델의 깊이와 크기 등을 고려하여 ResNet50, ResNet152, EfficientNetB0, EfficientNetB4 모델을 선정하고 학습모델의 성능을 비교하였다.



[그림 2] 딥러닝에 사용한 RetinaNet 모델 아키텍처

[Fig. 2] RetinaNet model architecture used for deep learning

딥러닝 학습을 위한 개발 환경은 Python (version 3.7.12)을 사용하였으며, 이미지 전처리 및 딥러닝 학습을 위하여 tensorflow (version 2.4.0), keras (version 2.4.3), keras\_resnet (version 0.2.0), OpenCV (version 4.5.0), read\_roi (version 1.6.0) 라이브러리를 사용하였다. 딥러닝 학습은 1개의 NVIDIA Tesla V100-SXM2-32GB를 사용하여 진행하였다. 모델 학습에 대한 하이퍼 파라미터는 학습률 (learning rate) 0.00001(1e-5), 에폭 (epochs) 300, 배치 사이즈 (batch size)는 EfficientNetB4 모델의 경우 8로 설정하였으며, 나머지 세 종류의 모델은 16으로 설정하여 학습을 진행하였다. 스텝 수(steps per epoch)는 (전체 학습 데이터 수 / 배치 사이즈 크기)로 계산되며, 이는 각 학습에 따라 계산되어 학습에 적용되었다. 또한, 학습에 사용되는 이미지의 크기는 512로 고정하여 학습하였다.

모델 학습은 4차 교차검증(4 fold cross-validation)으로 진행하였으며, 전체 데이터를 4개의 fold로 나누고 각각 다른 fold의 데이터를 테스트 데이터로 설정하여 모델의 학습 성능을 검증하였다. 모든 학습은 backbone 모델만을 제외하고 같은 조건으로 학습을 수행하였다. [표 1]은 각 fold 별 학습 및 테스트 데이터 수를 나타냈다.

[표 1] 4차 교차검증을 위해 구성한 각 fold 별 학습 및 테스트 데이터 세트 구성

[Table 1] Configuration of training and test data sets for each fold configured for 4 fold cross-validation

	Train		Test		Total	
	환자 수	영상 수	환자 수	영상 수	환자 수	영상 수
Fold 1	692	7,053	231	2,447	923	9,500
Fold 2	692	7,202	231	2,298	923	9,500
Fold 3	692	7,041	231	2,459	923	9,500
Fold 4	693	7,204	230	2,296	923	9,500

### 3. 결과

본 연구에서는 RetinaNet를 기반으로 4가지 종류의 ResNet50, ResNet152, EfficientNetB0, EfficientNetB4 네트워크를 backbone 모델로 사용하여 위내시경 영상에서의 위 병변을 검출하는 모델을 학습하고 검출 결과를 비교하였다. RetinaNet은 위 병변으로 예상되는 위치에 대한 box 정보와 box 영역에 대한 클래스 예측을 동시에 수행한다. 그중 box 정보에 대한 예측 결과를 통해 민감도(Sensitivity)와 FP/images 지표를 비교 성능 지표로 선택하여 각 fold별 모델 성능들의 평균을 계산하고 이를 비교하였다.

민감도 계산을 위한 혼동행렬(confusion matrix) 계산은 레이블링 작업을 통해 수집한 좌표값을 Ground Truth로 사용하고 학습모델이 예측한 box의 좌표값을 비교하여 두 결과 사이의 Intersection over Union (IoU) 값을 계산하고, 이를 통해 계산된 IoU 값이 설정한 경계값(threshold) 이상인 경우, 이를 진양성 (True Positive, TP)으로 계산한다. 또한, 학습모델이 검출하지 못한 위 병변은 위음성 (False Negative, FN), 학습모델이 검출하였으나 경계값 이하의 IoU를 나타낸 예측 결과는 위양성 (False Positive, FP)로 계산하였다. 각각의 backbone 모델과 교차검증을 통해 학습한 모델 결과를 혼동행렬을 계산하고 민감도와 FP/images 지표를 계산하여 [표 2]에 나타냈다.

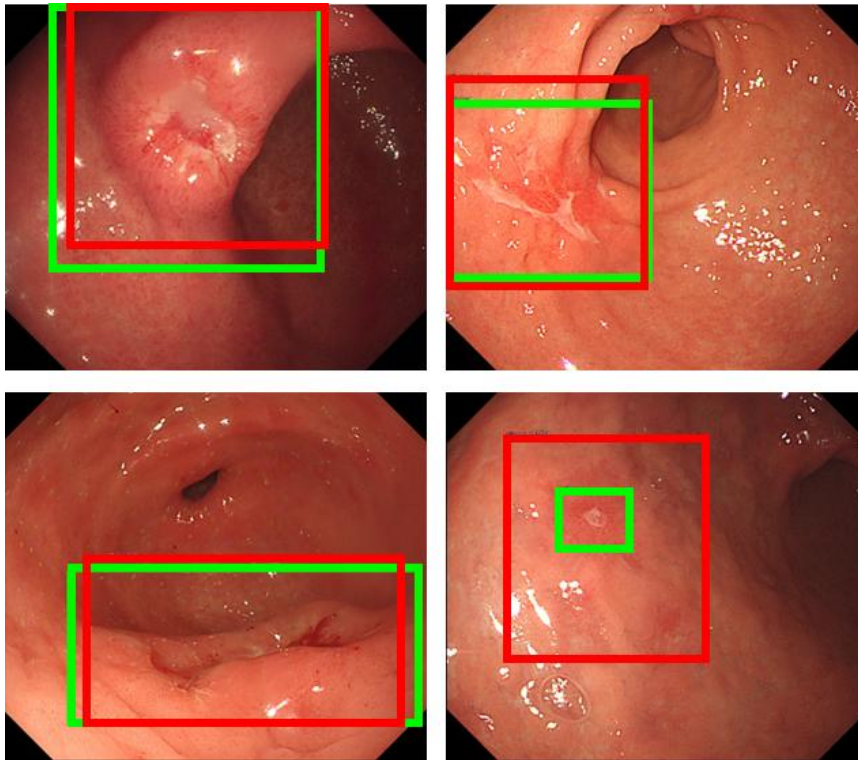
[표 2] RetinaNet 기반 backbone 네트워크별 모델 학습 성능의 교차검증 결과

[Table 2] Cross-validation results of model learning performance for RetinaNet-based backbone network

		TP	FN	FP	Sensitivity (%)	FP/images
ResNet50	Fold 1	864	290	98	74.87	0.0392
	Fold 2	744	440	60	62.84	0.0251
	Fold 3	853	233	149	78.55	0.0632
	Fold 4	834	227	153	78.61	0.0679
Total/Average		3295	1190	460	73.72	0.0489
ResNet152	Fold 1	889	265	79	77.04	0.0316
	Fold 2	948	236	149	80.07	0.0623
	Fold 3	856	230	87	78.82	0.0369
	Fold 4	818	243	118	77.10	0.0524
Total/Average		3511	974	433	78.26	0.0458
EfficientNetB0	Fold 1	915	239	347	79.29	0.1388
	Fold 2	888	216	418	80.43	0.1819
	Fold 3	849	237	520	78.18	0.2206
	Fold 4	857	204	651	80.77	0.2891
Total/Average		3509	896	1936	79.67	0.3268
EfficientNetB4	Fold 1	688	466	77	59.62	0.0308

	Fold 2	683	501	82	57.69	0.0343
	Fold 3	706	380	92	65.01	0.0390
	Fold 4	725	336	169	68.33	0.0750
Total/Average		2802	1683	420	62.66	0.0448

RetinaNet 기반 backbone 네트워크별 모델들의 평균 교차검증 성능은 ResNet50 73.72%, ResNet152 78.26%, EfficientNetB0 79.67%, EfficientNetB4 62.66%의 민감도를 보였으며 0.0489 (ResNet50), 0.0458 (ResNet152), 0.3268 (EfficientNetB0), 0.0448 (EfficientNetB4)의 FP/images를 보였다. 민감도 성능을 기준으로 판단하였을 때 EfficientNetB0 모델을 backbone으로 사용한 학습 결과가 가장 높게 나타났으나 다른 학습모델에 비해 FP/images가 높게 나타났다. 민감도와 FP/images 성능을 모두 고려하였을 때는 ResNet152 모델을 backbone 모델로 사용한 학습 결과가 우수하게 나타났다. [그림 3]은 EfficientNetB0 네트워크를 사용한 검출 학습모델의 예측 결과를 나타내고 있다.



[그림 3] 위내시경 영상에서의 위 병변 자동 검출 학습모델 예측 결과 예시 (backbone : EfficientNetB0, 초록 : Ground Truth, 빨강 : 모델 예측 결과)

[Fig. 3] Example of prediction results of the training model for automatic gastric lesion detection in gastroscopic images (backbone: EfficientNetB0, green: Ground Truth, red: model prediction results)

#### 4. 결론

본 연구는 위내시경 영상에서의 위 병변 검출을 위한 딥러닝 기반 학습모델을 개발하기 위한 검출 학습 네트워크 중 RetinaNet 네트워크를 선택하여 학습을 진행하였다. RetinaNet 네트워크는 피쳐 맵 추출을 위해 다양한 CNN 네트워크를 backbone으로 사용한다. 피쳐 맵 추출은 RetinaNet의 bounding box regression이나 classification 예측 결과에 직접적으로 사용되기 때문에 이를 위한 backbone 모델에 대한 선택이 검출 모델 성능에 중요하게 작용한다. 따라서 본 연구에서는 피쳐 추출을 위해 활용되는 다양한 모델 중 ResNet50, ResNet152, EfficientNetB0, EfficientNetB4 네트워크를 backbone으로 적용하여 학습한 모델의 성능을 비교한다.

위내시경 영상에서의 위 병변 검출을 위한 학습모델 개발을 위해 4차 교차검증을 진행하여 backbone 네트워크를 제외한 데이터 세트 구성과 하이퍼 파라미터 등의 조건들을 동일하게 학습하고 각 fold 별 학습모델의 성능을 계산하여 backbone 네트워크에 대한 평균 성능을 비교하였다. EfficientNetB0 네트워크를 사용한 학습의 민감도가 79.67%로 가장 높은 성능을 보였으며, ResNet152 네트워크가 78.26%로 두 번째로 높은 성능을 보였다. 하지만 EfficientNetB0 네트워크는 0.3268의 FP/images를 보이며 이는 다른 모델들에 비해 정확히 예측한 결과도 많지만 그만큼 오검출하는 결과도 많은 것을 확인할 수 있다. 따라서 민감도와 FP/images 수치를 모두 만족하는 backbone 모델은 78.26%의 민감도, 0.0458의 FP/images의 성능을 나타낸 ResNet152 네트워크라고 이야기할 수 있다.

본 연구에서는 위내시경 촬영 영상을 기반으로 위 병변을 자동으로 검출하는 딥러닝 기반의 학습모델을 개발하였다. 개발한 모델을 실시간 검사 과정에서 사용할 수 있도록 적용한다면 위내시경 검사를 수행하는 전문의가 숙련도 부족, 잦은 검사에 의한 피로도 증가에 따른 집중력 감소 등으로 인해 위 병변을 발견하지 못하고 지나치는 빈도를 획기적으로 낮출 수 있을 것이다. 위내시경 영상에서의 위 병변만을 검출하고자 하였으며 검출된 위 병변이 위암인지 위궤양인지에 대한 분석은 진행하지 않았다. 이는 실시간으로 위 병변을 검출하고자 할 때 검출에 성공하였는지에 대한 여부를 판단하는 것이 주된 목표이며 검출된 병변에 대한 종류를 예측하는 것은 추가적인 요소라고 생각하였기 때문이다. 따라서 향후 classification에 대한 모델 학습과 성능 향상을 위해 더 많은 데이터를 수집하여 학습을 수행하고 다양한 하이퍼 파라미터를 적용한다면 모델 성능을 고도화해나갈 수 있을 것으로 판단된다.



## Reference

- [1] K. P. Ko, "Epidemiology of gastric cancer in Korea", *Journal of the Korean Medical Association*, vol. 62, no. 8, August 2019, pp. 398-406, doi: 10.5124/jkma.2019.62.8.398.
- [2] Korean Statistical Information Service, "Cancer registration Statistics", kosis.kr, [https://kosis.kr/statHtml/statHtml.do?orgId=117&tblId=DT\\_117N\\_A00124&conn\\_path=12](https://kosis.kr/statHtml/statHtml.do?orgId=117&tblId=DT_117N_A00124&conn_path=12), (accessed November 25, 2022).
- [3] H. A. Park, S. Y. Nam, S. K. Lee, S. G. Kim, K. N. Shim, S. M. Park, S. Y. Lee, H. S. Shin, Y. M. Shin, K. M. Kim, K. J. Lee, T. Y. Lee, I. J. Choi, S. S. Hong, J. W. Kim, Y. J. Lee, S. Y. Kim, Y. Kim, W. C. Lee, I. K. Chung, "The Korean guideline for gastric cancer screening", *Journal of the Korean Medical Association*, vol. 58, no. 5, May 2015, pp. 373-384, doi: 10.5124/jkma.2015.58.5.373.
- [4] J. N. Lee, H. C. Cho, H. C. Cho, "A Study on Data Augmentation Methods Optimized for Gastric Cancer Classification in Gastroscopy Images", *The Transactions of the Korean Institute of Electrical Engineers*, vol. 70, no. 12, November 2021, pp. 2015-2021, doi: 10.5370/KIEE.2021.70.12.2015.
- [5] S. J. Park, Y. J. Kim, J. W. Chung, K. G. Kim, "Evaluation of Transfer Learning in Gastroscopy Image Classification using Convolutional Neural Network", *Journal of Biomedical Engineering Research*, vol. 39, October 2018, pp. 213-219, doi: 10.9718/JBER.2018.39.5.213.
- [6] G. Y. Yi, Y. J. Kim, K. G. Kim, "Classification model of normal and gastric ulcer Deep Learning on endoscopy Digital Image", *Journal of Digital Art Engineering & Multimedia*, vol. 6, no. 2, December 2019, pp. 133-140, doi: 10.29056/jdaem.2019.12.06.
- [7] S. J. Eun, E. Y. Jung, D. K. Kim, "Development of Colonoscopy Guide Algorithm based on Fast R-CNN", *Journal of Next-generation Convergence Information Services Technology*, vol. 8, no. 2, June 2019, pp. 213-222, doi: 10.29056/jncist.2019.06.09.
- [8] T. Y. Lin, P. Goyal, R. Girshick, K. He, P. Dollár, "Focal loss for dense object detection", *IEEE International Conference on Computer Vision (ICCV)*, October 22-29, 2017, Venice, Italy, pp. 2980-2988, doi: 10.1109/ICCV.2017.324.
- [9] K. He, X. Zhang, S. Ren, J. Sun, "Deep Residual Learning for Image Recognition", *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, June 27-30, 2016, Las Vegas, NV, USA, pp. 770-778, doi: 10.1109/CVPR.2016.90.
- [10] M. Tan, Q. V. Le, "EfficientNet: Rethinking Model Scaling for Convolutional Neural Networks", *36th International Conference on Machine Learning*, June 9-15, 2019, California, USA, pp. 6105-6114, doi: 10.48550/arXiv.1905.11946.
- [11] J. W. Park, Y. Kim, W. J. Kim, S. J. Nam, "Automatic Anatomical Classification Model of Esophagogastroduodenoscopy Images Using Deep Convolutional Neural Networks for Guiding Endoscopic Photodocumentation", *Journal of The Korea Society of Computer and Information*, vol. 26, no. 3, March 2021, pp. 19-28, doi: 10.9708/jksci.2021.26.03.019.