

다중 슬라이스 다중 라벨 기반의 공간적 맥락 정보 학습을 통한 MRE 영상에서 장 염증성 질환 검출

이한상^{1°} 이서현^{1°} 서니은² 임준석² 홍헬렌^{1*}

¹서울여자대학교 소프트웨어융합학과

²연세대학교 의과대학 영상의학교실

° 공동 1저자

* 교신저자

hlhong@swu.ac.kr

Detection of Inflammatory Bowel Diseases in MRE Images Using Multi-Slice Multi-Label Spatial Context Learning

Hansang Lee^{1°} Seohyun Lee^{1°} Nieun Seo² Joonseok Lim² Helen Hong^{1*}

¹Department of Software Convergence, Seoul Women's University

²Department of Radiology, Yonsei University College of Medicine

° Co-first authors

* Corresponding author

요약

크론병(Crohn's disease)은 만성 염증성 장 질환으로 크론병의 정확한 진단과 중증도 판정을 위해 MRE 영상에서 염증성 부위를 정확하게 검출하는 것이 필수적이다. 기존 단일 슬라이스 기반 MRE 분석은 연속적인 공간적 맥락 정보를 반영하지 못해 탐지 성능이 제한적이다. 본 연구는 다중 슬라이스의 맥락 정보를 YOLOv4 모델에 통합하는 다중 슬라이스 다중 라벨(MSML) 방법을 제안한다. 각 염증 영역에 대해 연속된 세 개의 슬라이스에서 중복 발생 횟수를 라벨로 부여하여 공간적 맥락을 강화한다. 실험 결과, 제안방법은 단일 슬라이스 기반 모델 대비 높은 mAP(69.27%)와 탐지 일관성을 보였으며, 위양성 검출이 감소됨을 확인하였다.

Abstract

Crohn's disease is a chronic inflammatory bowel disease, and accurate detection of inflammatory regions in MRE images is essential for its diagnosis and severity assessment. Traditional single-slice-based MRE analysis fails to capture the spatial context across consecutive slices, limiting detection performance. This study proposes a Multi-Slice Multi-Label (MSML) method that integrates spatial context into the YOLOv4 model. By assigning labels based on the occurrence frequency of inflammatory regions across three consecutive slices, the method enhances spatial awareness. Experimental results demonstrate that the proposed MSML approach achieves higher mAP (69.27%) and improved detection consistency compared to single-slice-based models, while also reducing false positive detections.

키워드: 의료 영상 분석, 객체 검출, 딥러닝, 크론병, 염증성 장 질환

Keywords: Medical image analysis, Object detection, Deep learning, Crohn's disease, Inflammatory bowel disease

*corresponding author: Helen Hong / Department of Software Convergence, Seoul Women's University (hlhong@swu.ac.kr)

1. 서론

크론병(Crohn's disease)은 소화관의 모든 부위에서 발생할 수 있는 만성 염증성 장질환으로, 장벽의 염증과 손상을 유발하여 심각한 합병증을 유발할 수 있다 [1]. 현재 크론병 진단에 가장 널리 사용되는 방법은 대장내시경으로, 장 내부를 시각적으로 확인하고 조직 검사를 시행할 수 있다. 그러나 대장내시경은 침습적 검사로 환자에게 불편함을 주며, 장 점막 표면의 작은 병변을 탐지하는 데 한계가 있다 [2]. 이에 반해 자기공명 장조영술(Magnetic Resonance Enterography, MRE)은 비침습적 검사, 크론병의 활성도를 측정할 수 있는 효과적인 대안으로, 병변의 위치, 크기, 깊이 등을 시각화하여 조기 발견과 적시 치료에 유용하다 [3]. 그러나 MRE 영상을 해석하여 크론병을 진단하는 과정은 시간이 많이 소요되며, 고도의 전문성이 필요하다. 숙련된 임상가의 경우 MRE를 통한 진단 정확도가 최대 89%에 이르지만, 비숙련자의 경우 정확도가 크게 떨어져 임상 결과에 편차가 발생할 수 있다 [4]. 이러한 문제를 해결하기 위해, 딥러닝 기반의 자동화된 MRE 영상 분석이 요구된다. 이를 통해 진단 시간을 단축하고, 진단의 정확성과 일관성을 향상시켜 크론병 관리에 중요한 역할을 할 수 있다.

기존 연구들은 MRE 영상의 품질을 개선하고, 딥러닝 기반 방법론을 통해 진단 정확도를 높이는 다양한 접근법을 제시하였다. Shand et al. [4]은 딥러닝 기반 다중 인스턴스 CNN 모델을 사용하여 MRE 데이터에서 라벨 오염에 강인한 크론병 병변의 자동 분류를 수행하였다. Son et al. [5]은 딥러닝 기반 복원(Deep Learning Reconstruction, DLR) 기법을 통해 MRE 영상의 노이즈와 움직임 아티팩트를 줄이고 크론병의 염증을 분류하고 평가하였다. 또한, Lian et al. [6]은 비선형 방정식 모델 기반의 영상 복원을 통해 MRE 영상에서 Gibbs 아티팩트를 제거하고 고주파 세부 정보를 복원함으로써 크론병과 궤양성 대장염의 병변 분포 및 활성도를 평가하였다. 기존의 딥러닝 기반 장 염증 검출 모델들은 주로 MRE 슬라이스 한 장씩을 독립적으로 학습하고 예측하는 방식을 사용해 왔다. 이러한 접근 방식은 MRE 영상이 인체의 연속된 단면으로 구성되며, 각 슬라이스 간에 공간적 관계를 포함하고 있다는 점에서 이러한 접근 방식은 한계가 있다. 특히, 단일 슬라이스 학습 방식은 연속된 슬라이스에서 동일한 위치에 나타나는 병변을 통합적으로 처리하지 못해 병변 검출이 간헐적으로 실패하는 문제가 발생할 수 있다. 예를 들어, Lee et al. [7, 8]은 연속된 슬라이스에서 동일 병변의 위치를 놓치는 사례를 제시하며, 이를 해결하기 위해 공간적 문맥을 반영하는 접근법의 중요성을 강조하였다.

YOLO(You Only Look Once)는 실시간 객체 탐지를 위해 설계된 딥러닝 모델로, 빠른 처리 속도와 높은 정확도를 바탕으로 다양한 탐지 작업에서 활용되고 있다 [9]. 최근에는 YOLO가 의료 영상 분석에도 폭넓게 적용되어 여러 종류의 병변 및 구조를 탐지하는 데 사용되고 있다. Chen et al. [10]은 YOLOv3 기반 알고리즘을 자기공명영상(MRI)에서 경동맥 플라크를 자동으로 분

Table 1: Statistics of inflammatory bowel diseases dataset.

	Train	Validation	Test	Total
# of patients				
All	104	52	52	208
sMARIA = 0	14	7	7	28
sMARIA $\in (0, 8)$	53	26	26	105
sMARIA ≥ 8	37	19	19	75
# of slices				
All	8038	4474	4030	16542
Avg. # of boxes				
All	38.1	33.1	40.8	37.5
sMARIA = 0	0	0	0	0
sMARIA $\in (0, 8)$	26.3	25.7	37.7	29.0
sMARIA ≥ 8	69.5	55.4	60.1	63.5

할하고, 뇌졸중 위험을 평가하는 데 활용하여 94.81%의 높은 정확도를 보여주었다. Shand et al. [4]은 컴퓨터단층촬영(CT) 영상에서 폐 결절을 탐지하기 위해 개선된 YOLOv5를 활용하여 mAP 86.4%의 높은 성능을 달성하고, 결절 유형을 CNN-SVM을 통해 효과적으로 분류하였다. Li et al. [11]은 YOLOv8 알고리즘을 사용하여 무릎 관절 MRI에서 슬개골 불안정성과 탈구를 탐지하였으며, 초급 방사선과 전문의와 유사한 정확도(83%)와 더 빠른 해석 시간을 기록하였다. 그러나 MRE 영상에서 장 염증성 질환을 탐지하기 위해 YOLO가 적용된 연구는 아직 수행되지 않았다.

본 논문에서는 MRE 영상의 연속 슬라이스 간 공간적 맥락 정보를 YOLOv4 모델 [9]에 통합하여 기존 단일 슬라이스 방식의 한계를 극복하는 새로운 방법을 제안한다. 제안된 방법은 연속된 세 개의 슬라이스를 모델의 입력으로 결합하고, 각 염증 영역의 슬라이스 간 중복 발생 횟수를 라벨로 활용하여 공간적 맥락 정보를 강화한다. 이를 통해 YOLOv4 모델의 빠르고 효율적인 검출 성능을 유지하면서도 연속적인 병변 검출 누락 문제를 해결하고 탐지 일관성을 향상시킬 수 있다. 실험 결과, 제안된 방법은 슬라이스 간 맥락 정보를 효과적으로 학습하여 병변 검출의 일관성을 개선하였으며, 기존 단일 슬라이스 기반 및 단일 라벨 방식에 비해 위양성 검출 또한 감소함을 확인하였다.

2. 제안 방법

2.1 데이터 수집 및 전처리

본 연구에서는 2016년 3월부터 2018년 12월까지 신촌세브란스 병원에서 222명의 환자로부터 MR 소장 조영술(MRE) 영상을 수집하였다. 이 중 장 염증 위치에 대한 추가 검토가 필요한 14명의 데이터를 제외한 208명의 환자 데이터를 최종적으로 사용하였다. 환자 데이터는 장 염증성 질환의 중증도를 나타내는 sMARIA 점수 기준으로 정상(0), 경증(0-8 사이), 중증(8 이상)에 대해 28:105:75로 다양한 상태의 환자를 포함하고 있다. 모든 MRE 영상은 정맥기(portal phase)에서 촬영된 조영증강 영상으로, 해

상도는 512×512 화소이며, 화소 크기는 각각 0.78mm×0.98mm 이고, 절편 두께는 3–4mm 범위이다. 장 염증의 위치는 세 명의 임상 의사 바운딩 박스 형태로 수동 분할하여 레이블링하였다. 수동 레이블링을 통해 16,542장의 영상에 대해 총 7,800개의 바운딩 박스 좌표 데이터를 획득하였다. 데이터셋은 학습, 검증, 테스트 데이터셋으로 계층적 분할(stratified split)하여 각 데이터셋 내에서 중증도 별 환자 수의 비율을 유지하였다. 이는 훈련 데이터와 평가 데이터 간 분포 차이를 최소화하여 모델의 성능을 신뢰할 수 있도록 하기 위함이다. 표 1은 계층적 분할 결과를 요약한 것이다.

$$\mu = \frac{1}{N} \sum_{i=1}^N I(x_i, y_i), \quad \sigma = \sqrt{\frac{1}{N} \sum_{i=1}^N (I(x_i, y_i) - \mu)^2} \quad (1)$$

$$I_{\text{normalized}}(x, y) = \frac{I(x, y) - \min}{\sigma} \quad (2)$$

$$I_{\text{scaled}}(x, y) = \frac{I_{\text{normalized}}(x, y) - \min}{\max - \min} \times 255 \quad (3)$$

수집된 MRE 영상에는 가우시안 밝기값 정규화를 적용하였다. 먼저, 수식 (1)과 같이 영상의 평균 밝기값(μ)과 표준편차(σ)를 계산하였다. 이후 각 픽셀의 밝기값을 수식 (2)의 정규화 공식에 따라 변환하고, 마지막으로 수식 (3)을 이용하여 정규화된 값을 0–255 사이의 값으로 선형 변환하였다. 이때, 최소값(min)과 최대값(max)은 각각 정규화된 픽셀 값($I_{\text{normalized}}$)의 최소값과 최대값을 사용하였다. 가우시안 밝기값 정규화를 통해 영상 내 픽셀 밝기값 분포를 개선하고 배경 잡음을 감소시켜, 염증 부위의 시각적 특성을 강조하였다 [12].

2.2 다중 슬라이스 다중 라벨 학습을 통한 장염증 검출

제안 방법은 다중 슬라이스 다중 라벨 데이터(Multi-Slice Multi-Label, MSML) 생성과 YOLOv4 모델 학습의 두 단계로 구성된다. 첫 번째 단계에서는 주어진 MRE 영상과 장 염증 부위의 바운딩 박스에 대하여 MSML 데이터를 생성한다. MSML 데이터는 연속된 세 장의 슬라이스의 영상에서 염증 영역의 중복 발생 횟수 ($L=1, 2, 3$)를 라벨로 부여하여 공간적 맥락 정보를 강화한다. 두 번째 단계에서는 생성된 MSML 데이터를 활용하여 YOLO v4 모델을 학습시키고, 학습된 모델을 통해 염증 검출 결과를 도출한다.

2.2.1 다중 슬라이스 다중 라벨(MSML) 데이터 생성

그림 1은 MSML 데이터 생성의 관점에서 제안방법과 기존 데이터를 비교한 개요도이다. 단일 슬라이스 단일 라벨(Single-Slice Single-Label, SSSL) 방식은 기본적인 입력 데이터 생성 방식으로, 주어진 단일 슬라이스 영상과 해당 슬라이스의 바운딩 박스 마스크를 학습 데이터로 사용한다. 그러나 이 방식은 슬라이스 간

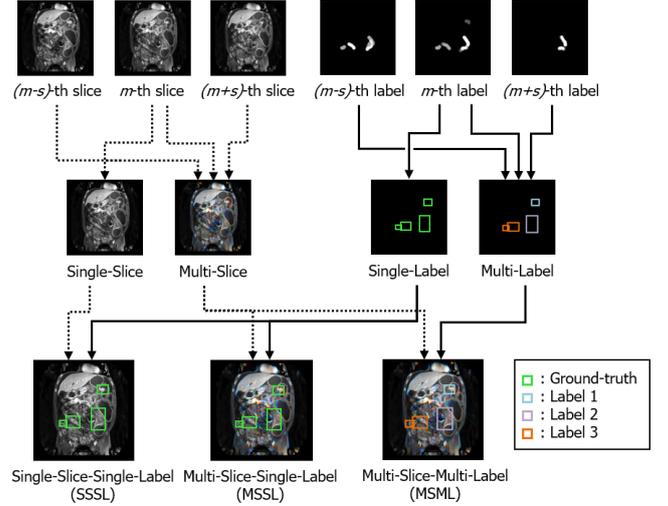


Figure 1: Overview of various input data generation methods.

공간적 연관성을 반영하지 못하며, 연속된 슬라이스에서 동일한 염증이 나타나더라도 이를 통합적으로 학습하지 못하는 한계가 있다. 이러한 한계는 염증의 정확한 위치와 크기를 연속적으로 파악하는 데 어려움을 초래한다.

다중 슬라이스 단일 라벨(Multi-Slice Single-Label, MSSL) 방식은 SSSL 방식의 단점을 보완하기 위해, m 번째 슬라이스를 중심으로 이전($m-s$) 및 이후($m+s$) 슬라이스를 각각 다중 채널로 결합한 영상을 입력 데이터로 사용하며, m 번째 슬라이스의 바운딩 박스 마스크를 라벨로 활용한다. 이때 s 는 슬라이스 간 거리를 나타내며, 공간적 맥락 정보를 조절하는 역할을 한다. 이를 통해 연속된 슬라이스 간 공간적 정보를 일부 반영할 수 있다. 그러나 MSSL 방식은 여전히 단일 슬라이스의 마스크 정보에 의존하기 때문에 연속된 슬라이스에서 염증의 위치, 크기, 모양 변화가 충분히 학습에 반영되지 못한다. 예를 들어, 여러 슬라이스에 걸친 염증이더라도 중간 슬라이스의 정보만 학습되는 경우, 모델이 전체 염증 특성을 완전히 이해하지 못할 가능성이 있다.

다중 슬라이스 다중 라벨(Multi-Slice Multi-Label, MSML) 방식은 이러한 한계를 극복하기 위해, MSSL 방식에 연속된 슬라이스의 마스크 정보를 추가적으로 활용한다. 학습 중인 m 번째 슬라이스를 기준으로, 이전($m-s$) 및 이후($m+s$) 슬라이스의 동일 위치에서 나타나는 염증의 발생 횟수를 확인하고, 라벨을 부여한다. 예를 들어, 동일 위치에 염증이 세 슬라이스에 모두 나타날 경우 라벨 3, 두 슬라이스에 나타날 경우 라벨 2, 한 슬라이스에만 나타날 경우 라벨 1로 지정한다. 이러한 라벨링 방식은 연속된 슬라이스 간 염증의 위치와 크기 변화를 학습할 수 있도록 설계되어, 모델이 공간적 연관성과 연속성을 더 효과적으로 반영할 수 있다. 표 2는 $s=1$ 일 때 MSML 방식으로 생성된 바운딩 박스의 라벨 분포를 나타낸다. 지정된 라벨들을 살펴보면 라벨 1, 2, 3의 바운딩 박스 개수가 대략 1:7:9로 분포하는 것을 확인할 수 있다. 이는 연속된 슬라이스에서 나타나는 염증의 빈도를 반영하며, 클래스 불균형(class imbalance)을 유발한다. 이로

Table 2: Statistics of MSML bounding boxes in our dataset.

# of boxes	Train	Validation	Test	Total
Total	3966	1721	2123	7810
MSML (Label 1)	254	77	140	471
MSML (Label 2)	1639	675	846	3160
MSML (Label 3)	2073	969	1137	4179

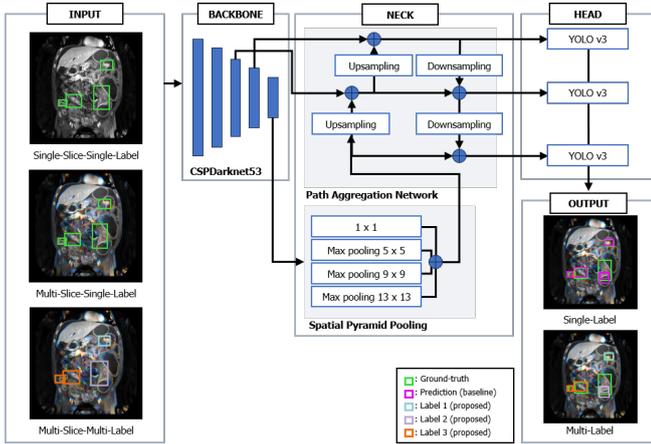


Figure 2: Architecture of YOLOv4 model.

인해 모델이 연속된 슬라이스에서 많이 나타나는 염증을 더욱 집중적으로 학습하도록 유도한다. 생성된 바운딩 박스 좌표와 라벨 정보는 연속된 슬라이스 기반의 다중 채널 영상과 함께 YOLOv4 모델의 학습 데이터로 사용된다. MSML 방식은 연속된 슬라이스 간 염증이 동일한 위치에서 나타나는 빈도에 따라 라벨을 부여하여, 공간적 맥락과 염증의 변화 양상을 동시에 학습하여 연속된 슬라이스에서 염증이 일관되게 검출되도록 유도한다.

2.2.2 YOLOv4 모델 학습

그림 2는 장 염증 질환 부위 검출을 위한 YOLOv4 모델의 구조를 나타낸다. 본 연구에서 사용된 YOLOv4 모델은 CSPNet (Cross Stage Partial Network) [13]을 기반으로 한 CSPDarknet53을 백본 (Backbone)으로 채택하여 특징 추출을 수행한다. 백본에서 추출된 특징 맵은 SPP (Spatial Pyramid Pooling) [14] 모듈과 결합되어 다양한 해상도로 분석이 가능하며, 크기가 다른 염증 부위를 정밀하게 검출할 수 있도록 한다. 또한, PAN (Path Aggregation Network) [15]는 다단계 특징 조합을 통해 작은 염증에서 큰 염증까지 다양한 크기의 병변을 정확히 탐지하도록 설계되었다. 마지막으로, YOLOv3 [9] 헤드는 다중 스케일 특징 맵을 활용하여 염증 부위를 빠르고 정확하게 검출한다. 모델의 입력은 SSSL, MSSL, MSML 방식으로 생성된 데이터로 구성되며, 각 방식에 따라 출력 형식이 달라진다. 먼저 SSSL 방법의 출력으로는 단일 슬라이스 영상에서 하나의 라벨로 예측된 염증 부위의 바운딩 박스 좌표 정보가 반환된다. MSSL 방법의 출력은 세 장의 연속된 슬라이스가 결합된 다중 슬라이스 영상에서 예측된 m 번째 슬라이스

이상의 염증 부위의 바운딩 박스 좌표가 동일한 라벨 정보와 함께 반환된다. 마지막으로 MSML 방법은 다중 슬라이스 영상에서 예측된 염증 부위와 그에 해당하는 라벨이 바운딩 박스 좌표와 함께 반환된다.

3. 실험 및 결과

3.1 실험 환경 및 계획

실험은 데이터 전처리와 성능 평가를 위해 주피터 노트북 7.0.8 환경에서 파이썬 3.11.7 버전을, 모델 학습은 구글 코랩 환경에서 파이썬 3.10.12와 Darknet 프레임워크 [16]를 사용하여 진행되었다. GPU 환경은 CUDA 12.2와 cuDNN 8.9.6을 기반으로 구성되었으며, MS COCO 데이터로 사전 훈련된 yolov4-tiny.conv.29 모델을 초기 가중치로 사용하였다. 모델을 최적화하기 위해 조정된 하이퍼파라미터는 모멘텀(momentum) 값을 0.95로, 기울기 감소율 (weight decay)은 0.001로 설정하였다. 색 포화도(saturation)와 노출도(exposure)는 각각 1.0으로 고정하여 입력된 원본 이미지의 색상 및 밝기에 변화가 없도록 하였으며, 학습률(learning rate)은 0.0015로 설정하였다. 클래스 수, 좌표 정보, 오브젝트 유무 등을 예측하는 파라미터인 filters는 SSSL 방법과 MSSL 방법에서는 1개의 클래스에 대해 18개의 필터를 사용하였고, MSML 방법에서는 3개의 클래스에 대해 24개의 필터를 사용하였다. 최대 배치 수 (max batches)는 6000, burn in 단계는 1000으로 설정되었으며, 이 단계에서는 학습 초기의 과도한 변동을 방지하기 위해 학습률이 점진적으로 증가한다. 비최대 억제 기법(nms)은 DIoU nms [17]를 적용하였다.

실험에서는 제안된 방법의 효과를 검증하기 위해 SSSL(baseline) 방법에 대해서 제안 방법인 MSSL과 MSML 방법을 각각 $s=1, 2, 3$ 일 때의 실험 결과를 정성적으로 비교하고, 정량적 평가를 통해 염증 검출 성능을 종합적으로 분석하였다. 성능 평가는 수정된 mAP(mean Average Precision)와 정밀도(precision), 재현율(recall)을 활용하였다. 기존 mAP는 IoU(Intersection over Union) 임계값을 초과하는 가장 높은 신뢰도 점수를 가진 하나의 바운딩 박스만 참 긍정(True Positive)으로 간주한다. 반면 수정된 mAP는 IoU 임계값을 초과하는 모든 바운딩 박스를 참 긍정으로 처리하여 연속적인 슬라이스 간 검출의 일관성을 더 잘 반영한다. 이를 통해 수정된 mAP는 연속된 슬라이스 정보를 포함하는 MSML 방식의 성능을 보다 공정하고 정확하게 평가할 수 있다. 또한 수정된 mAP와 정밀도, 재현율은 전체 염증 부위 (All ground truth, All GT) 및 라벨 3 염증 부위 (Label 3 GT) 두 가지 관점에서 평가되었다. All GT 방식은 기존과 같이 전체 염증 부위에 대한 검출 정확도를 평가하는 방식으로 MSML의 경우 라벨 1, 2, 3에 관계없이 모두 단일 라벨 바운딩 박스로 간주하여 평가하였다. Label 3 GT 방식의 경우 중증도 판정에 있어 중요한 것으로 평가되는 크기가 큰 염증들에 대한 검출 정확도에 초점을 맞춘 방식으로 정답 데이터와 MSML

Table 3: Comparison of bowel inflammation detection results between the baseline and the proposed methods for All GT.

All GT Methods	mAP (%)		Precision (%)		Recall (%)	
	@10	@50	@10	@50	@10	@50
SSSL (Baseline)	68.83	62.27	66.17	58.16	70.54	65.46
MSSL ($s = 1$)	63.60	56.48	59.84	51.31	65.84	60.33
MSSL ($s = 2$)	65.30	58.03	61.84	53.77	67.44	61.42
MSSL ($s = 3$)	66.95	60.90	64.56	57.19	68.71	63.77
MSML ($s = 1$)	69.27	63.06	67.10	59.68	70.90	65.85
MSML ($s = 2$)	67.45	61.52	65.56	58.23	69.41	64.45
MSML ($s = 3$)	67.65	62.11	65.93	59.27	69.18	64.71

Table 4: Comparison of bowel inflammation detection results between the baseline and the proposed methods for Label 3 GT.

Label 3 GT Methods	mAP (%)		Precision (%)		Recall (%)	
	@10	@50	@10	@50	@10	@50
SSSL (Baseline)	68.15	62.49	65.15	58.29	69.85	65.45
MSSL ($s = 1$)	62.06	55.82	58.15	50.9	64.05	59.24
MSSL ($s = 2$)	62.67	57.28	59.13	53.3	64.6	60.13
MSSL ($s = 3$)	63.22	59.41	60.66	56.39	64.79	61.65
MSML ($s = 1$)	71.01	66.21	69.45	63.29	72.36	68.24
MSML ($s = 2$)	65.45	61.08	63.13	57.89	67.06	63.39
MSML ($s = 3$)	74.18	71.38	73.86	70.25	75.07	72.59

검출결과 모두 Label 3에 대한 바운딩 박스만을 대상으로 평가 및 비교를 수행하였다. 또한 SSSL과 MSML의 검출 결과 영상을 직접 비교하고 정성적으로 분석함으로써 MSML 방식이 연속된 슬라이스 간 공간적 맥락을 통해 슬라이스 간 검출 일관성을 실제로 높이고 위양성 검출을 줄였는지 검증하였다.

3.2 실험 결과

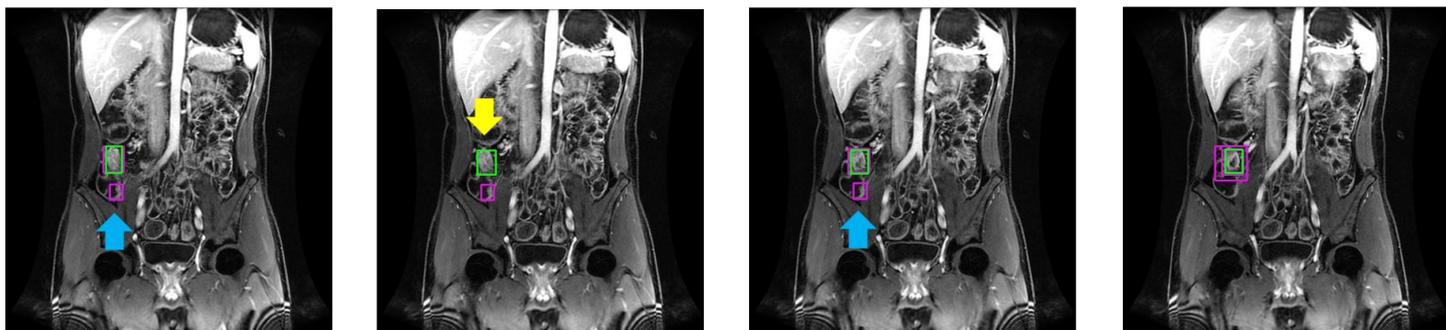
표 3는 전체 염증 부위(All GT)에 대해 제안 방법 및 비교 방법의 장 염증 검출 성능을 비교한 결과를 나타낸다. 기존의 SSSL 방식은 단일 슬라이스를 독립적으로 처리하여 mAP@10 68.83%, Recall@10 70.54%의 안정적인 성능을 보였으나, 공간적 맥락을 반영하지 못해 연속된 슬라이스 간 검출 일관성을 강화하는 데 한계가 있다. MSSL은 연속된 슬라이스 정보를 추가하여 공간적 맥락을 반영하려 했으나, 중간 슬라이스의 마스크 정보에만 의존한 단일 라벨 구조로 인해 SSSL보다 낮은 성능(mAP@10: 63.60%–66.95%)을 기록하며 연속된 슬라이스에서의 염증 검출 특성을 효과적으로 학습하지 못하는 한계를 보였다. 반면, 제안된 MSML은 연속된 슬라이스 간의 염증 발생 빈도를 라벨링하여 공간적 맥락 학습을 통해 SSSL과 MSSL 대비 우수한 성능(mAP@10: 69.27%, Recall@10: 70.90%)을 기록하였다. 특히 MSML에서 $s = 1$ 일 때 가장 높은 성능을 보였는데, 이는 연속된 슬라이스 간 거리가 너무 멀어지면 공간적 맥락이 희석되고 모델의 학습에 방해가 될 수 있으며, 인접한 슬라이스 간의 관계를 학습할 때 전체적으로 가장 효과적인 공간적 맥락이 학습됨을

확인하였다.

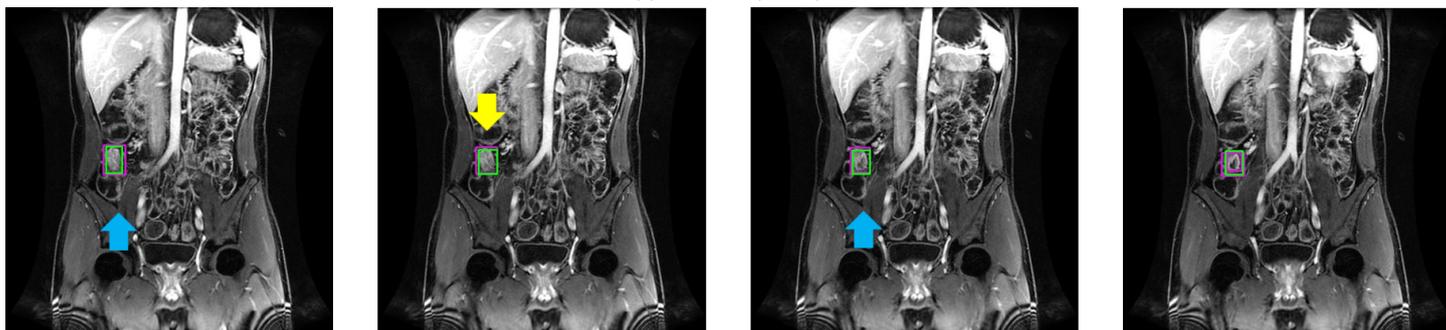
표 4는 라벨 3 염증 부위(Label 3 GT)에 대해 제안 방법 및 비교 방법의 장 염증 검출 성능을 비교한 결과를 나타낸다. 기존의 SSSL 방식은 mAP@10 68.15%, Recall@10 69.85%를 기록하며 공간적 맥락을 반영하지 못해 중증도가 큰 염증을 검출하는 데 한계가 있었다. MSSL은 연속된 슬라이스를 활용했음에도 SSSL 대비 낮은 성능(mAP@10: 62.06%–63.22%, Recall@10: 64.05%–64.79%)을 기록하였는데, 이는 All GT 결과에서와 마찬가지로 중간 슬라이스 중심의 학습 방식이 큰 염증 부위의 연속적 특징을 충분히 반영하지 못했기 때문으로 분석된다. 반면, MSML은 연속된 슬라이스의 염증 정보와 발생 빈도를 함께 학습함으로써 SSSL과 MSSL 대비 크게 개선된 성능을 보였다 (mAP@10: 71.01%–74.18%, Recall@10: 72.36%–75.07%). 특히, Label 3 GT에서는 $s = 3$ 일 때 MSML 방식이 가장 높은 성능을 기록했는데, 이는 크기가 큰 염증 부위의 공간적 맥락을 멀리 떨어진 슬라이스까지 학습할 때 검출 성능이 더욱 강화됨을 의미한다. 이는 All GT 결과에서 $s = 1$ 이 가장 좋은 성능을 기록한 것과 대조적이며, All GT에서는 작은 크기의 염증도 포함된 데이터 특성상 인접 슬라이스($s = 1$)가 더 효과적인 반면, Label 3 GT에서는 중증도가 큰 염증의 경우 넓은 맥락 정보($s = 3$)가 더 효과적임을 보여 준다. 이를 통해 제안된 MSML이 전체적인 염증 검출 정확도를 높일 뿐만 아니라 염증 크기와 중증도에 따라 다른 맥락 정보를 학습할 수 있는 유연한 방식임을 확인할 수 있다.

그림 3, 4는 서로 다른 두 환자의 연속한 네 장의 슬라이스 영상

□ : Ground Truth
 □ : Prediction
 ↓ : FP Reduction
 ↓ : Slice Consistency



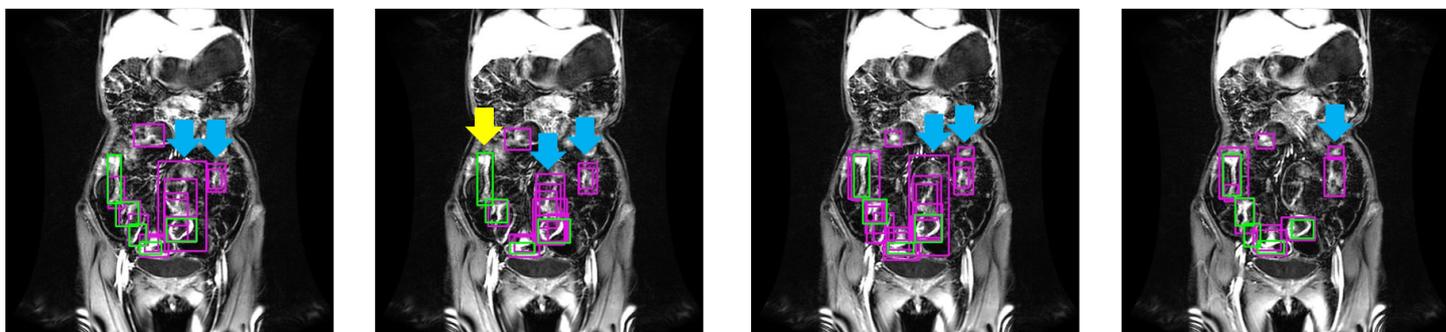
(a) Baseline (SSSL)



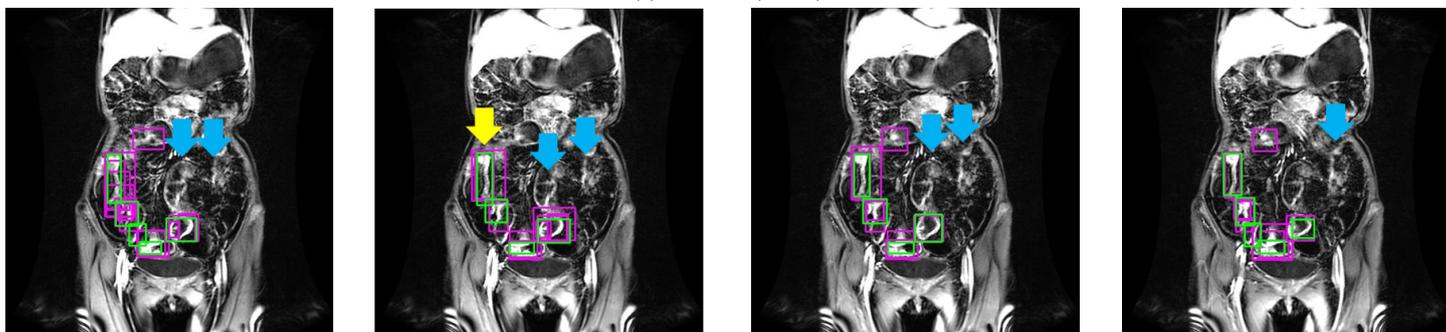
(b) MSML ($s = 1$)

Figure 3: Examples of bowel inflammation detection results comparing slice-wise consistency between the baseline and the proposed method.

□ : Ground Truth
 □ : Prediction
 ↓ : FP Reduction
 ↓ : Slice Consistency



(a) Baseline (SSSL)



(b) MSML ($s = 1$)

Figure 4: Examples of bowel inflammation detection results comparing false positive reduction between the baseline and the proposed method.

에서 기존 방법(SSSL)과 제안된 MSML ($s = 1$)의 장염증 검출 결과를 나타낸 것이다. 그림에서 파란색 화살표는 제안방법에 의해 위양성 검출이 감소된 부분을 나타낸다. 두 환자의 SSSL 결과에서 파란색 화살표로 표시된 위양성 검출을 보면 장 벽이 미세하게 두꺼워진 영역에 대하여 장염증과의 영상적 특징 유사성으로 인해 다수의 위양성 검출이 이루어진 것을 볼 수 있다. 반면, 제안된 MSML은 이러한 위양성 검출을 크게 줄이고, 염증 부위에 집중하여 검출 정확도를 높인 것을 확인할 수 있다. 이는 MSML 방식이 연속된 슬라이스의 정보를 활용하여 실제 염증 부위와 유사 밝기값의 잡음을 구별할 수 있는 학습 능력을 효과적으로 향상시켰음을 보여준다.

그림 3, 4에서 노란색 화살표는 제안방법에 의해 슬라이스 일관성이 개선된 부분을 나타낸다. 그림 3의 노란색 화살표를 보면 SSSL 결과에서는 앞뒤 슬라이스에서는 염증이 바르게 검출된 반면 두 번째 슬라이스에서는 염증이 검출되지 않는 비일관성을 보였으나, 제안된 MSML은 같은 염증 부위를 모든 슬라이스에서 일관되게 검출하는 것을 확인할 수 있다. 그림 4에서도 마찬가지로 노란색 화살표 위치에서 SSSL은 슬라이스에 따라 염증에 비일관적으로 검출된 반면 MSML은 일관되게 염증을 검출해내는 것을 보였다. 이는 MSML 방식이 연속된 슬라이스 간의 공간적 맥락을 학습하여 동일한 염증을 일관되게 탐지할 수 있는 강점을 가졌음을 보여준다.

4. 결론

본 연구에서는 만성 염증성 장 질환의 진단 및 중증도 판정을 위해 MRE 영상에서 염증 부위를 정확히 검출할 수 있는 다중 슬라이스 다중 라벨 (MSML) 방법을 제안하였다. 제안 방법은 YOLOv4 모델에 연속된 세 개의 슬라이스를 결합하여 공간적 맥락 정보를 반영하고, 염증 영역의 슬라이스 간 중복 발생 횟수 ($L=1, 2, 3$)를 라벨로 부여함으로써 기존 단일 슬라이스 기반 방법의 한계를 극복하고자 하였다. 실험 결과, 제안 방법은 단일 슬라이스 기반 모델 대비 mAP, 정밀도, 재현율 등 성능 지표에서 모두 향상된 결과를 보였으며, 공간적 맥락 거리 $s = 1$ 일 때 가장 높은 성능을 보였다. 특히 위양성 검출이 감소하고 연속된 슬라이스에서 탐지 일관성이 개선됨을 확인하여 제안된 방법이 효과적으로 작동함을 검증하였다.

감사의 글

본 연구는 정부(과학기술정보통신부)의 재원으로 한국연구재단의 지원(No. RS-2024-00336063 및 No. RS-2023-00207947)과 서울여자대학교 학술연구비의 지원(2025)을 받아 수행된 연구임.

References

- [1] K. Cushing and P. D. R. Higgins, "Management of crohn disease: A review," *JAMA*, vol. 325, no. 1, pp. 69–80, Jan 2021.
- [2] L. Alfarone, A. Dal Buono, V. Craviotto, A. Zilli, G. Fiorino, F. Furfaro, F. D'Amico, S. Danese, and M. Allocca, "Cross-sectional imaging instead of colonoscopy in inflammatory bowel diseases: Lights and shadows," *Journal of Clinical Medicine*, vol. 11, no. 2, p. 353, Jan 2022.
- [3] O. Brem, D. Elisha, E. Konen, M. Amitai, and E. Klang, "Deep learning in magnetic resonance enterography for crohn's disease assessment: a systematic review," *Abdominal Radiology*, pp. 1–7, 2024.
- [4] M. Shand, J. Manderfield, S. Surbhi, C. McLafferty, Y. Sharma, S. Sengupta, P. Fernandes, D. Koroulakis, S. Syed, and D. E. Brown, "Identifying pediatric crohn's disease using deep learning to classify magnetic resonance enterography (mre) images," in *2021 Systems and Information Engineering Design Symposium (SIEDS)*, 2021, pp. 1–6.
- [5] J. Son, Y. Lee, H. Lee, J. Lee, H. Kim, and M. Lebel, "Lava hypersense and deep-learning reconstruction for near-isotropic (3d) enhanced magnetic resonance enterography in patients with crohn's disease: utility in noise reduction and image quality improvement," *Diagn Interv Radiol*, vol. 29, no. 3, pp. 437–449, May 2023.
- [6] G. Lian, Y. Peng, J. He, J. Yi, Y. Yin, X. Liu, and F. Zeng, "Diagnosis and prognosis of epidemic inflammatory bowel disease under convolutional neural network algorithm and non-linear equation model," *Results in Physics*, vol. 22, p. 103912, 2021.
- [7] S. Lee, H. Lee, and H. Hong, "Detection of chronic inflammatory disease incorporating spatial information of continuous slices based on yolov4," in *Proceedings of the 2024 Spring Conference of the Korean Multimedia Society*, vol. 27, no. 1, 2024.
- [8] H. Lee, S. Lee, N. Seo, J. S. Lim, and H. Hong, "Detection of inflammatory bowel diseases in mre images with spatially context-aware yolov4," in *the Radiological Society of North America (RSNA) Annual Meeting, 2024*, 2024.
- [9] J. Redmon and A. Farhadi, "Yolov3: An incremental improvement," *Computer Vision and Pattern Recognition*, vol. 1804, pp. 1–6, 2018.

- [10] Y. F. Chen, Z. J. Chen, Y. Y. Lin, Z. Q. Lin, C. N. Chen, M. L. Yang, J. Y. Zhang, Y. Z. Li, Y. Wang, and Y. H. Huang, "Stroke risk study based on deep learning-based magnetic resonance imaging carotid plaque automatic segmentation algorithm," *Frontiers in Cardiovascular Medicine*, vol. 10, p. 1101765, 2023.
- [11] T. Li, N. M. Gharaibeh, S. Jia, Z. Qinaer, S. Aihemaiti, A. HaNaTe, and G. Wu, "Yolov8 algorithm-aided detection of patellar instability or dislocation on knee joint mri images," *Acta Radiologica (Stockholm, Sweden : 1987)*, 2024.
- [12] S. Das, J. Saikia, S. Das, and N. Goni, "A comparative study of different noise filtering techniques in digital images," *International Journal of Engineering Research and General Science*, vol. 3, no. 5, pp. 180–191, 2015.
- [13] C. Y. Wang, H. Y. M., Y. H. Wu, P.-Y. Chen, J.-W. Hsieh, and I.-H. Yeh, "Cspnet: A new backbone that can enhance learning capability of cnn," in *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition workshops*, 2020, pp. 390–391.
- [14] K. He, X. Zhang, S. Ren, and J. Sun, "Spatial pyramid pooling in deep convolutional networks for visual recognition," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 37, no. 9, pp. 1904–1916, 2015.
- [15] S. Liu, L. Qi, H. Qin, J. Shi, and J. Jia, "Path aggregation network for instance segmentation," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2018, pp. 8759–8768.
- [16] AlexeyAB, "Darknet: Open source neural networks in c," <https://github.com/AlexeyAB/darknet/releases/>, accessed: 2024-12-27.
- [17] Z. Zheng, P. Wang, W. Liu, J. Li, R. Ye, and D. Ren, "Distance-iou loss: Faster and better learning for bounding box regression," in *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 34, no. 07, 2020, pp. 12 993–13 000.
- [18] A. M. Harale, V. K. Bairagi, E. Boonchieng, and M. R. Bachute, "Nodules detection in lungs ct images using improved yolov5 and classification of types of nodules by cnn-svm," *IEEE Access*, vol. 12, pp. 140 456–140 471, 2024.
- [19] J. Sobek, J. R. Medina Inojosa, B. J. Medina Inojosa, S. M. Rassoulinejad-Mousavi, G. M. Conte, F. Lopez-Jimenez, and B. J. Erickson, "Medyolo: A medical image object detection framework," *Journal of Imaging Informatics in Medicine*, vol. 37, no. 6, pp. 3208–3216, 2024.

〈 저자 소개 〉



- 이 한 상
- 2011년 2월 한국과학기술원 전기 및 전자공학과 졸업(학사)
- 2013년 2월 한국과학기술원 전기 및 전자공학과 졸업(석사)
- 2019년 2월 한국과학기술원 전기 및 전자공학과 졸업(박사)
- 2019년 3월~2025년 2월 한국과학기술원 정보전자연구소 연수연구원
- 2025년 3월~현재 서울여자대학교 미래산업융합연구소 연구교수
- 관심분야 : 인공지능, 딥러닝, 의료영상분석
- <https://orcid.org/0000-0003-0917-1589>



- 홍 헬 렌
- 1994년 2월 이화여자대학교 전자계산학과 졸업(학사)
- 1996년 2월 이화여자대학교 전자계산학과 졸업(석사)
- 2001년 8월 이화여자대학교 컴퓨터학과 졸업(박사)
- 2001년 9월~2003년 7월 서울대학교 컴퓨터공학부 BK 조교수
- 2006년 3월~현재 서울여자대학교 소프트웨어융합학과 교수
- 관심분야 : 의료 인공지능, 딥러닝, 영상처리 및 분석
- <https://orcid.org/0000-0001-5044-7909>



- 이 서 현
- 2025년 2월 서울여자대학교 소프트웨어융합학과 졸업(학사)
- 관심분야 : 인공지능, 딥러닝, 기계학습, 생성모델
- <https://orcid.org/0009-0004-2339-7570>



- 서 니 은
- 2008년 02월 울산대학교 의학과 졸업(학사)
- 2012년 02월 울산대학교 의학과 졸업(석사)
- 2016년 02월 울산대학교 의학과 졸업(박사)
- 2015년 03월~현재 연세대학교 의과대학 영상의학과 교수
- 관심분야 : 복부영상의학, 위장관질환, 자기공명영상, 영상분석
- <https://orcid.org/0000-0001-8745-6454>



- 임 준 석
- 1994년 02월 연세대학교 의과대학 의학과 졸업(학사)
- 1998년 02월 연세대학교 의과대학 진단방사선과 졸업(석사)
- 2007년 02월 연세대학교 의과대학 진단방사선과학 졸업(박사)
- 2005년 03월~현재 연세대학교 의과대학 영상의학교실 교수
- 관심분야 : 복부영상의학, 디지털헬스, 인공지능
- <https://orcid.org/0000-0002-0334-5042>