

Estimating Gastrointestinal Transition Location Using CNN-based Gastrointestinal Landmark Classifier

Hyeon Woong Jang[†] · Chang Nam Lim[†] · Ye-Suel Park^{††} ·
Gwang Jae Lee^{†††} · Jung-Won Lee^{††††}

ABSTRACT

Since the performance of deep learning techniques has recently been proven in the field of image processing, there are many attempts to perform classification, analysis, and detection of images using such techniques in various fields. Among them, the expectation of medical image analysis software, which can serve as a medical diagnostic assistant, is increasing. In this study, we are attention to the capsule endoscope image, which has a large data set and takes a long time to judge. The purpose of this paper is to distinguish the gastrointestinal landmarks and to estimate the gastrointestinal transition location that are common to all patients in the judging of capsule endoscopy and take a lot of time. To do this, we designed CNN-based Classifier that can identify gastrointestinal landmarks, and used it to estimate the gastrointestinal transition location by filtering the results. Then, we estimate gastrointestinal transition location about seven of eight patients entered the suspected gastrointestinal transition area. In the case of change from the stomach to the small intestine(pylorus), and change from the small intestine to the large intestine(ileocecal valve), we can check all eight patients were found to be in the suspected gastrointestinal transition area. we can found suspected gastrointestinal transition area in the range of 100 frames, and if the reader plays images at 10 frames per second, the gastrointestinal transition could be found in 10 seconds.

Keywords : Capsule Endoscopy(CE), Convolutional Neural Network(CNN), Gastrointestinal Location Tracking

CNN 기반 위장관 랜드마크 분류기를 이용한 위장관 교차점 추정

장 현 웅[†] · 임 창 남[†] · 박 예 슬^{††} · 이 광 재^{†††} · 이 정 원^{††††}

요 약

최근의 영상 처리 분야는 딥러닝 기법들의 성능이 입증됨에 따라 다양한 분야에서 이와 같은 기법들을 활용해 영상에 대한 분류, 분석, 검출 등을 수행하려는 시도가 활발하다. 그중에서도 의료 진단 보조 역할을 할 수 있는 의료 영상 분석 소프트웨어에 대한 기대가 증가하고 있는데, 본 연구에서는 데이터 셋이 방대하고 판단에 시간이 오래 걸리는 캡슐내시경 영상에 주목하였다. 본 논문의 목적은 캡슐내시경 영상의 판독에서 모든 환자에 대해 공통으로 수행되고, 판독하는 데 많은 시간을 차지하는 위장관 랜드마크를 구별하고 위장관 교차점을 추정하는 것이다. 이를 위해, 위장관 랜드마크를 식별할 수 있는 CNN 학습 모델을 설계하였으며, 이를 이용하여 결괏값을 필터링해 위장관 교차점을 추정하였다. 무작위로 환자 데이터를 샘플링한 모델을 이용해서 나온 결과를 필터링 후에 위장관 교차점을 추정하였을 때, 88% 환자는 위장에서 소장으로 변화하는 위장관 교차점(유문판) 의심 구역 안에 들어왔으며, 소장에서 대장으로 변화하는 위장관 교차점(회맹판)의 경우 100% 환자가 위장관 교차점 의심 구역 안에 들어온 것을 확인할 수 있었다. 100프레임 범위로 위장관 교차점 의심 구역을 찾을 수 있었으며, 판독자가 초당 10프레임의 속도로 판독을 진행한다면 10초안에 위장관 교차점을 찾아낼 수 있다.

키워드 : 캡슐내시경, 컨볼루션 신경망, 위장관 위치 추적

※ 이 논문은 정부(과학기술정보통신부)의 재원으로 한국연구재단의 지원을 받아 수행된 연구임(No. 2019R1H1A2101112).

※ 이 논문은 2019년도 한국정보처리학회 추계학술발표대회에서 '캡슐내시경의 위치추적을 위한 CNN 기반 위장관 랜드마크 분류기 설계'의 제목으로 발표된 논문을 확장한 것이다.

† 준 회원 : 아주대학교 전자공학과 석사과정

†† 준 회원 : 아주대학교 전자공학과 박사과정

††† 비 회원 : 아주대학교 의과대학 소화기내과 교수

†††† 종신회원 : 아주대학교 전자공학과 교수

Manuscript Received : December 31, 2019

First Revision : January 29, 2020

Accepted : February 6, 2020

* Corresponding Author : Jung-Won Lee(jungwony@ajou.ac.kr)

1. 서 론

최근의 영상 처리 분야에서는 딥러닝과 머신러닝 기법이 우수한 성능을 보임에 따라, 영상에 대한 분류, 분석, 검출 등의 목적을 가지고 다양한 분야에서 활용되고 있다. 그중에서도 최근 주목받고 있는 분야로는 의료 분야가 있다. 의사의 인력이 점점 부족해지고 기술에 발전에 따라 다양한 의료 영상들이 등장하면서 의료 영상을 진단을 보조할 수 있으면서 진단율을 높여줄 수 있는 연구들이 활발해지고 있다[1]. 국내에

서는 대표적으로 Vuno의 BoneAge[2], Lunit의 Insight[3] 등과 같은 기계학습 기반의 의료 진단 보조 소프트웨어도 많이 개발되고 있다. 목적에 따라서 컴퓨터단층촬영영상(CT: Computed Tomography), 자기공명영상(MR: Magnetic Resonance), X-ray 등 다양한 의료 영상들이 존재하는데, 본 연구에서는 소장 촬영을 목적으로 오랜 시간 동안 방대한 데이터 셋을 가지며, 많은 진단 시간을 필요로 하는 캡슐내시경 영상에 주목하였다.

캡슐내시경은 식도에서 항문까지 긴 위장관을 촬영하기 위한 장비로, 특별한 마취 없이 행해질 수 있고 다른 내시경과는 다르게 이물감이거나 거부감이 덜하기 때문에 점차 사용량이 늘어나고 있다. 캡슐내시경의 가장 큰 장점은 대장 내시경으로는 판독이 어려운 소장을 촬영할 수 있다는 것이다. 이때 캡슐내시경은 약 8시간에서 10시간의 측정 시간을 가지고 1초에 약 2장의 사진을 찍는다.

캡슐내시경은 식도, 위장, 소장 그리고 대장을 차례로 이동하면서 촬영하게 되며 이들을 위장관 랜드마크라 부른다. 위장관 랜드마크는 위장관 교차점을 기준으로 나뉘게 되며, 식도와 위장의 교차점은 식도 위 경계(식도 위 이음부), 위장과 소장의 교차점은 유문판, 소장과 대장의 교차점은 회맹판막 이라 한다. 의사 혹은 판독관은 위장관 교차점을 기반으로 위장관 랜드마크를 관찰하는 데 영상에서 해당 지점이 환자마다 다르다. Fig. 1은 캡슐내시경 영상의 구성을 간략하게 보여준다.

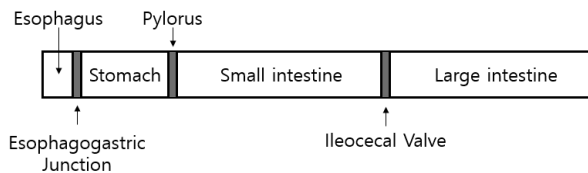


Fig. 1. Sequence of Wireless Capsule Endoscopy Video

현재, 캡슐내시경은 다음과 같은 문제점을 지니고 있다. 첫째, 캡슐내시경의 촬영 과정에서 많은 양의 데이터가 쌓이게 되며 이를 판독하는 의사는 보통 1시간에서 2시간의 판독 시간이 필요하다. 캡슐내시경 영상의 데이터가 크기 때문에 판독 시간이 오래 걸리게 되며, 특히 수많은 이미지 중 병변을 가지고 있는 이미지를 찾기 어려워지고, 촬영의 주목적이 소장 촬영이므로 이를 위한 시간이 걸리게 된다.

둘째, 병원 현장에서 캡슐내시경 시술에 어려움을 겪고 있는 정체(retention) 현상이 있다. 정체 현상이란 캡슐이 특정 위치에 정체되어 모든 구간을 통과하지 못하는 현상을 의미한다. 정체 현상이 발생하면 실시간 추적을 통해 정체 현상을 해결하기 어려운 상황이며, 최악의 경우 소장을 촬영하지 못하고 배터리가 소진되어 촬영에 실패할 수 있다.

이러한 문제점들을 해결하기 위해서 다양한 연구들이 진행되고 있다. 대표적으로 영상의 판독 시간을 줄여주고, 영상 진단을 보조하기 위한 연구들이 있다. 이러한 연구들은 병변을 찾거나, 위장관 랜드마크를 구별하거나, 위장관 교차점을

추정하는 것에 초점을 맞추었다. [4-6]의 연구는 캡슐내시경 영상에서 다양한 병변들을 진단 보조를 목적으로 한 연구들이며, [7-9]의 연구는 캡슐내시경 영상 내에서의 노이즈 영상 처리 및 중복 영상 처리, [10-14]의 연구는 위장관 랜드마크 분류 및 위장관 교차점을 추정하는 연구들이 진행되어왔다. 그러나 위장관 교차점 추정의 경우에는 식도 위 경계에 대해서는 좋은 성능을 보였지만, 회맹판의 경우에는 오차가 큰 문제점을 가지고 있다. 따라서 본 논문에서는 이를 보완하기 위한 위장관 분류 모델 기반의 교차점 추정 기법을 제안한다.

제안하는 기법은 크게 두 가지 모델로 구성되어 있다. 첫 번째는 위장관 랜드마크 분류기 모델, 두 번째는 분류기의 결과를 이용한 위장관 교차점 추정 모델이다. 위장관 랜드마크 분류기는 이미지의 특성을 학습할 수 있는 CNN 학습모델을 이용하여 캡슐내시경 영상을 위장관 랜드마크 별로 분류한다. 또한, 위장관 교차점 추정은 분류기를 통해 나온 결과를 지수 이동평균을 이용해서 일정 시간 이상 한 위장관 랜드마크라고 인식할 경우 인식할 수 있는 위장관 교차점 추정기를 설계한다. 최종적으로, 위장관 랜드마크 분류기의 성능은 무작위로 환자 데이터를 섞은 후 분류기를 학습했을 경우에는 95%의 정확도를 보였고 8-fold 교차 검증 방식을 이용해서 학습했을 경우에는 평균 정확도가 71%를 보였다. 위장관 교차점 추정기는 유문판의 경우 87.5%, 회맹판의 경우 100%의 정확도를 보였다. 기존의 연구들에 비해서 회맹판 인식이 증가하였으며, 100프레임 이내로 위장관 교차점 의심 구역을 설정해 초당 10프레임의 속도로 판독을 진행한다면 10초안에 위장관 교차점을 찾아낼 수 있다.

2. 관련 연구

기존 캡슐내시경 관련 연구는 병변 진단 보조, 위장관 랜드마크 분류 그리고 위장관 교차점 추정 등의 연구가 수행되고 있다. 본 장에서는 병변 추출과 위장관 교차점 추정에 대한 기존의 연구를 소개한다.

2.1 캡슐내시경 영상 병변 추출 기법

[4]의 연구에서는 폴립과 종양에 초점을 둔 연구이다. 종양을 검출할 때 RGB 이미지를 HSV 이미지로 변화시킨 후 로그 가버 필터를 이용한 후 모폴로지(Morphology) 연산을 이용해서 의심 구역을 탐지하고 해당 구역의 이미지를 다시 RGB 값을 사용하여 SVM(Support Vector Machine)을 사용해서 종양을 검출한다. 폴립 검출 또한 로그 가버(Gabor) 필터를 이용하고 SUSAN 엣지(Edge) 검출을 이용해서 폴립의 후보군을 검출했다. 폴립의 경우에는 민감도 100%, 특이도 67.5%였으며 종양의 경우에는 민감도 75%, 특이도 73.3%의 결과를 보여주었다. 하지만 학습에 사용한 영상 표본들이 약 50장으로 많지 않았으며, 폴립의 경우 SUSAN 엣지 검출을 사용하는 데 실시간으로 진단할 수 없다는 단점이 있다. [5]의 연구는 종양을 인식하는 연구이며 LBP(Local Binary Pattern)을 추출

한 뒤 다양한 SVM을 이용해서 색공간마다 결과를 비교하였다. [6]의 연구에서는 웨이블릿(Wavelet)과 Emphasis를 이용해서 폴립의 특징을 추출했으며, 약 96%의 정확도를 보였다.

2.2 캡슐내시경 위장관 랜드마크 분류 및 교차점 추정 기법

[10]의 연구에서는 두 개의 CNN 모델과 은닉마르코프 모델을 이용해서 위장관 랜드마크를 분류하였다. 노이즈 영상을 먼저 CNN을 통해서 분류하고, 노이즈 영상을 제외한 영상을 다시 다른 CNN모델을 통해서 위장관 랜드마크를 분류하였다. 14명의 환자 데이터를 이용했으며, 전체적으로 랜드마크 분류기의 성능은 약 88% 정도의 정확도를 보였으며, 노이즈 영상도 77% 제거된 결과를 보여주었다. [12] 특허에서는 이미지 센서 뿐만 아니라 pH, 속도, 가속도 센서 등을 사용해서 위장관 교차점을 추정한다. [11]의 연구에서는 MPEG-7을 이용하여 두 가지 특징을 추출하고(Scarable Color, Homogeneous Texture), SVM(Support Vector Machine)과 베이지안 분류기(Bayesian Classifier)를 이용해서 위장관 랜드마크를 분류하고, 위장관 교차점을 추정하였다. 식도에서 위장, 위장에서 소장, 소장에서 대장의 경우에는 중간값 오차가 각 2, 211프레임으로 낮았지만, 소장에서 대장의 경우에는 중간값 오차가 1070프레임으로 높은 수치를 보여주었다. [13, 14]의 논문에서는 캡슐내시경 영상을 작은 이미지로 나누어서 각 구역마다 특징을 추출한 뒤 SVM을 통하여 위장관 교차점을 추정하였다. [10]의 논문과 비슷하게 식도에서 위장, 위장에서 소장의 경우에는 중간값 오차가 각 10, 121 프레임으로 낮았지만, 소장에서 대장의 경우에는 평균 424프레임을 보여주었고 평균 값은 2,035프레임의 결과를 보여주었다. 결과를 보면 소장에서 대장의 교차점의 오차는 상당히 컸으며, 교차점 모두 중간값 오차에 비해 평균 오차 값이 높은 값으로, 적절한 위치로 추정하지 못한 환자가 많은 것으로 평가된다.

3. 위장관 랜드마크 분류 기반 교차점 추정 기법

3.1 캡슐내시경 이미지 특징

랜드마크별 이미지의 특징을 살펴보기 위해서 히스토그램을 이용해 분석하였다. Table 1에서는 환자별 R, G, B의 평균값을 보여준다, 위장관 랜드마크는 사람마다 다른 색상을 가지고 있는 것을 알 수 있다. 5번 환자는 다른 환자에 비해서 색공간의 값이 매우 차이 나는 것을 확인할 수 있으며 이는 학습에 영향을 미칠 수 있다고 판단된다. 또한, 대장 사진의 경우에는 환자별로 값이 상이한데, 이는 많은 배설물들이 존재하기 때문이다.

Fig. 2는 각각 위, 소장, 대장의 영상이다. 각 영상을 살펴보면 위의 경우에는 다른 두 위장관에 비해 큰 주름을 가지고 있고 탄력적인 것처럼 보인다. 소장의 경우에는 영양소를 흡수하기 위해서 융털을 가지고 있다. 이러한 특징으로 인하여 주름의 곁의 부분은 오돌토돌한 특징을 보인다. 대장의 경우

Table 1. R, G, B Mean Value Each Gastrointestinal Landmark and Patient

Label	patient No	R_mean	G_mean	B_mean
Stomach	1	123.78	79.08	66.53
	2	119.52	80.90	68.39
	3	127.71	90.45	67.77
	4	127.71	83.01	50.94
	5	129.46	94.58	81.53
	6	109.21	75.24	51.49
	7	114.21	77.34	60.30
	8	127.20	82.34	51.10
Small Intestine	1	123.03	84.25	43.33
	2	118.91	86.23	37.14
	3	115.75	76.94	31.05
	4	124.21	80.64	32.99
	5	135.75	97.20	78.54
	6	127.41	90.09	41.76
	7	130.23	91.22	45.36
	8	129.95	85.19	43.69
Large Intestine	1	94.93	80.94	37.91
	2	116.31	90.52	43.53
	3	118.35	81.02	57.44
	4	120.08	100.00	47.95
	5	146.13	99.11	61.82
	6	127.89	103.32	48.33
	7	118.93	91.99	50.24
	8	89.32	72.74	56.51



Fig. 2. Representative Gastrointestinal Landmarks Image (Stomach, Small Intestine, Large Intestine)

에는 주름이 탄력이 없어 보이며, 소장처럼 융털을 가지고 있지 않고 대부분의 영상에 이물질이 끼어있는 경우가 많다.

3.2 캡슐내시경 학습 노이즈 영상

캡슐내시경은 음식물이 소화되는 경로를 통과하며 촬영하기 때문에 음식 잔여물 등의 이유로 인해 점막 내부가 제대로 촬영되지 않은 노이즈 영상이 존재할 수 있다. 노란 액체로 가득 찬 영상도 있으며, 배설물만 보이는 영상, 캡슐내시경의 통신 불량으로 인하여 생기는 영상, 거품으로 가득 낀 영상 등 많은 영상이 존재하는데, 이는 위장관의 병변을 파악하는데 있어서 유의미한 정보를 주지 않는다. 노이즈 영상들을 전

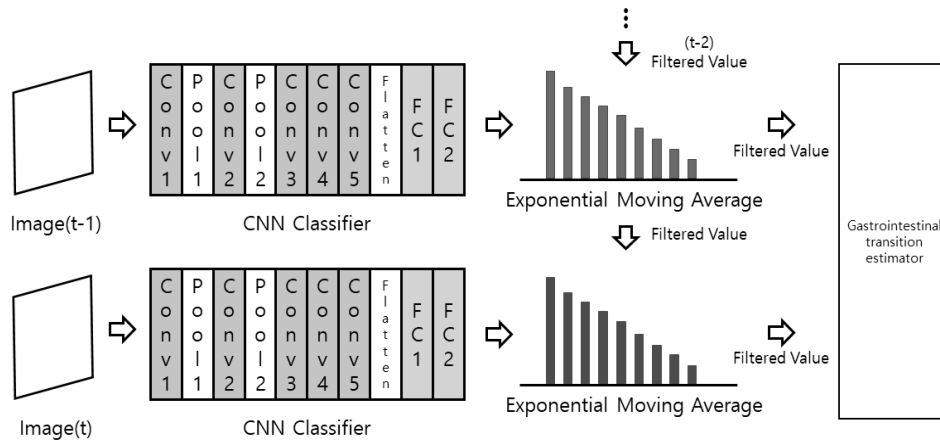


Fig. 3. Sequence of Estimating Gastrointestinal Tract Junction

처리한다면, 판독 시간을 줄일 수 있을 뿐만 아니라 의미 없는 영상을 학습하지 않기 때문에 학습에 있어서 좋은 결과를 낼 수 있다. [5]의 연구에서는 유사도와 엔트로피, 명암도를 이용하여 장시간의 판독을 요구하는 캡슐내시경 영상 중 중복된 영상을 제거하였다. [6]의 연구도 판독 시간을 줄이기 위하여 노이즈영상과 장기 사진을 분류한 후에 은닉마르코프 모델을 이용해서 영상을 분류하여 판독에 있어 쉽게 수행할 수 있게 했다. [7, 8]의 연구에서는 SVM과 K-평균 군집화를 이용하여 판독에 있어서 의미 없는 노이즈 영상을 제거하여 판독시간을 단축시키려 한다.

본 논문에서는 위장관 랜드마크 분류기의 성능을 향상시키기 위해서 노이즈 영상들을 제거하였다. 8명 환자의 데이터를 확인하여 과도하게 거품이 끼거나 잘 보이지 않는 영상들을 제거하여 데이터 셋을 만들었다. 4.2절에서는 노이즈가 학습에 어느 정도의 영향을 끼치는지 노이즈가 있는 데이터 셋과 없는 데이터 셋의 학습을 통해서 비교한다.

3.3 위장관 랜드마크 분류기 설계

본 논문에서는 위장관 교차점 추정을 위해, Fig. 3과 같은 모델을 제안한다. Fig. 3의 전반부(CNN 분류기)는 위장관 교차점 판독에 앞서 먼저 위장관 랜드마크를 분류한다. 5개의 합성곱층과 2개의 풀링 층으로 구성하였으며 모델의 상세한 사항은 Table 2에 명시되어 있다. 본 논문에서는 이와 같은 모델 구조를 기반으로 학습을 수행하였으며, 추가적으로 FC1층에서 많은 변수로 인해 분류기가 훈련 데이터에 편향될 수 있기 때문에 이를 해결하기 위해 FC1 층에 모델을 학습시킬 때는 25%의 dropout을 적용했다. 실험에서는 동일하게 설계된 모델에서 노이즈 영상에 유무에 따른 데이터 셋의 결과를 비교하며, 나온 결과 값을 이용하여 위장관 교차점을 추정한다.

3.4 위장관 교차점을 탐색을 위한 필터 설계

3.3절에서 설계한 모델을 훈련 시킨 뒤 나오는 결과값을 이용하여 위장관 교차점을 추정하는 필터를 설계한다. 먼저

Table 2. CNN Model's Detail Specification

Layer	Layer Comment	Output
Input	Input Image	128 x 128 x 32
Conv1	5 x 5 x 32 (SAME)	128 x 128 x 32
Pool1	2 x 2 max pooling	64 x 64 x 32
Conv2	5 x 5 x 64 (SAME)	64 x 64 x 64
Pool2	2 x 2 max pooling	32 x 32 x 64
Conv3	3 x 3 x 128 (SAME)	32 x 32 x 128
Conv4	3 x 3 x 128 (SAME)	32 x 32 x 128
Conv5	3 x 3 x 128 (SAME)	32 x 32 x 128
Flatten	flatten conv5 output	131072
FC1	131072 Input node	131072
FC2	256 Hidden node	256
FC_output	3 Output node	3

CNN 분류기의 결과 값을 정형화하는 과정이 필요한데, 결과값은 Equation (1)에 따라 결정할 수 있도록 하였다.

$$GL(t) = \begin{cases} 0, & \text{if image is stomach image} \\ 1, & \text{if image is small intestine image} \\ 2, & \text{if image is large intestine image} \end{cases} \quad (1)$$

GL(t)는 위장관 랜드마크(Gastrointestinal Landmark)를 의미하며, 소프트맥스를 통해서 0~2의 정숫값이 나오게 되며 각각 위장, 소장, 대장의 이미지를 의미한다. 각 환자의 시계열 이미지를 분류기를 통해서 나온 이상적인 결과는 Fig. 4와 같다. 그림과 같이 각 교차점을 기준으로 위장관 랜드마크가 변화하는 경우가 이상적인 결과이다.

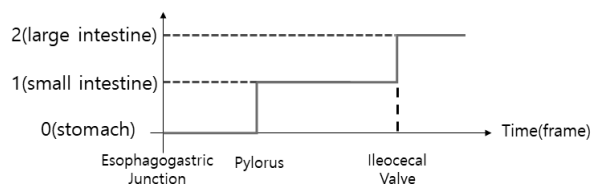


Fig. 4. Ideal Result from CNN Classifier on a Patient

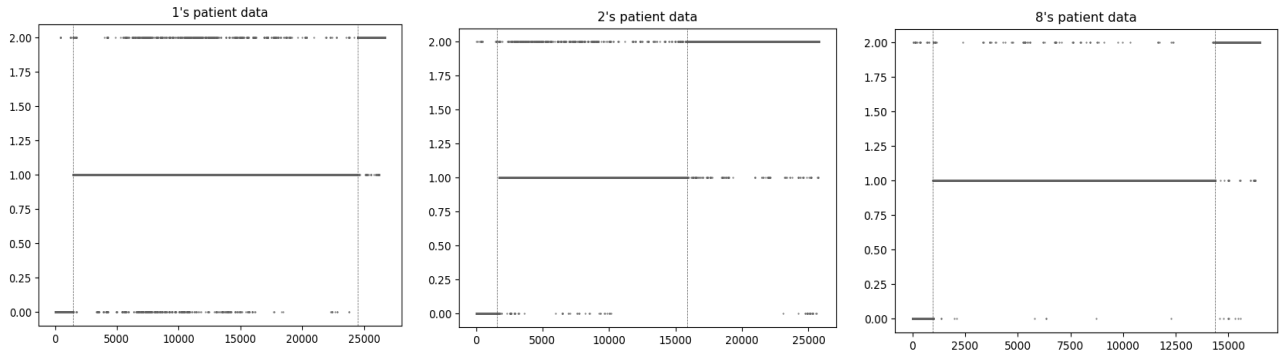


Fig. 5. Results from Several Patients Through CNN Classifier

하지만 분류기의 성능에 따라 잘못 판단하는 경우가 나오게 되며, 이를 고려하기 위해서 저역통과필터중 하나인 지수이동평균(Exponential Moving Average) 필터를 사용하였다. 사용한 필터의 식은 Equation (2)와 같으며 가중치 값을 조정하여 $GL(t)$ 의 값과 과거의 값들의 가중치를 조정할 수 있다.

$$F(t) = \begin{cases} F(t-1) * (1 - Weight) + GL(t) * Weight & \text{if } t > 0 \\ GL(t) & \text{if } t == 0 \end{cases} \quad (2)$$

위장관 교차점을 인식하는 과정은 다음과 같다. 먼저, $F(t)$ 의 값이 일정 기준 값(Threshold)을 처음 넘은 시점을 각 환자별로 저장해 놓는다. 각 환자별로 설정한 기준 값을 처음 넘었을 경우 해당 시점을 위장관 교차점 의심 구역으로 저장해 놓는다. 위장에서 소장으로 넘어가는 기준 값을 T 라고 하면 소장에서 대장으로 넘어가는 기준 값은 $(T+1)$ 이다. 이에 대한 설명은 Fig. 6에 설명되어 있다.

Algorithm 1: Estimate Gastrointestinal Transition Algorithm

Input: Filtered value $F(t)$
Output: Set of transition points (Pylorus, Ileocecal valve)

```

Initialize P_flag, IV_flag = False
While t is valid do
    if  $F(t) > \text{Threshold}$  and P_flag == False then
        Pylorus = t
        S_flag = True
    else if  $F(t) > (1 + \text{Threshold})$  and IV_flag == False then
        Ileocecal valve = t
        break
    t = t + 1
end
    
```

Fig. 6. Algorithm for Deciding the Suspected Transition Point

4. 실험 결과

4.1 실험 환경

총 8명 환자의 데이터로 노이즈를 포함하면 위장 20,175

장, 소장 114,046장, 대장 48,884장의 영상을 이용했다. 데이터를 무작위로 샘플링하는 방식과 7명의 데이터를 학습에, 1명의 데이터를 검증에 활용하는 교차 검증 방식을 수행하였으며, 가장 적은 데이터를 기준으로 1:1:1 비율로 맞춰 학습을 진행했다.

4.2 위장관 랜드마크 분류기 검증 결과

4.2절에서는 설계된 CNN 기반의 랜드마크 분류기의 성능을 검증한다. 검증 방식은 크게 두 가지로, 랜덤샘플링 기반의 학습 모델 검증과 환자 별 교차 검증이 있다. 먼저, Table 3은 랜덤샘플링 기반의 학습 모델 검증 결과이다. 두 결과를 비교하면 노이즈의 유무에 따라 모델의 정확도가 약 2% 증가한 것을 볼 수 있다. 노이즈 영상이 삭제되면서 학습 성능이 올라간 것을 확인할 수 있다. 하지만 이 경우 학습 모델이 데이터 셋에 과적합 되어 성능이 우수하게 나올 수 있기 때문에 다른 검증 방식을 진행하였다.

Table 3. CNN Classifier Result with Random Sampling

	With Noise	Without Noise
Accuracy	0.93	0.95

Table 4, 5는 교차검증(CV, Cross Validation)을 진행했을 때의 결과이며 각각 학습 노이즈 영상을 포함하고 학습한 결과와 제거한 후 학습한 결과이다. 1, 3, 5 환자의 경우에는 정확도가 매우 떨어지고, 나머지 환자의 경우에는 준수한 정확도가 나왔다. 이 결과를 보아 환자마다 데이터의 차이가 심하다는 것을 알 수 있고, 해당 환자의 특징은 다른 환자에게서 나타나지 않았다는 의미로 받아들여진다. 또한 노이즈를 제거한 후에는 1, 2, 4번의 환자의 경우에는 정확도가 증가하였고, 7번 환자의 경우에서만 정확도가 하락하였다. 랜덤 샘플링 모델의 정확도와 환자별 8-fold 교차 검증의 큰 정확도 차이를 개선하기 위해서는 각 환자와 비슷한 데이터의 환자가 존재해야 하는데 이를 해결하기 위해서는 많은 환자, 증상이 비슷한 환자의 데이터를 수집해서 학습을 시킨다면 차이를 개선할 수 있을 것으로 판단된다.

Table 4. 8-fold CV Result with Noise Data

Patient No	valuation	Stomach	Intestine	Colon	Accuracy
1	F1-score	0.15	0.47	0.36	0.36
2	F1-score	0.80	0.75	0.65	0.72
3	F1-score	0.56	0.77	0.31	0.62
4	F1-score	0.65	0.93	0.88	0.90
5	F1-score	0.00	0.10	0.46	0.30
6	F1-score	0.47	0.97	0.75	0.91
7	F1-score	0.82	0.94	0.24	0.90
8	F1-score	0.80	0.52	0.55	0.68

Table 5. 8-fold CV Result without Noise Data

Patient No	valuation	Stomach	Intestine	Colon	Accuracy
1	F1-score	0.22	0.64	0.42	0.53
2	F1-score	0.52	0.86	0.71	0.79
3	F1-score	0.42	0.85	0.19	0.62
4	F1-score	0.57	0.93	0.94	0.93
5	F1-score	0.09	0.59	0.32	0.31
6	F1-score	0.66	0.97	0.76	0.92
7	F1-score	0.59	0.93	0.34	0.88
8	F1-score	0.66	0.53	0.83	0.68

4.3 위장관 교차점 추정

위장관 교차점 추정에 사용되는 모델은 8-fold 모델은 결과가 상이하기 때문에, 노이즈를 제외하고 랜덤샘플링을 통해서 학습한 모델을 사용하였다. Fig. 5는 Equation(1)의 결과를 통해서 얻은 일부 환자의 데이터이다. X축은 프레임, 빨간 수직선은 실제 위장관 교차점을 의미한다. Table 6은 가중치 값은 0.01, 기준값은 0.85로 설정한 후의 결과이다. 오차의 값은 3.4절에서 제안한 방식을 이용해서 얻은 값과 실제 위장관 교차점의 프레임 값의 차이이다. Table 6을 보면 2번 5번 환자가 유문판의 오차가 큰 것을 알 수 있다. 8명의 환자 모두의 오차 값의 평균은 유문판의 경우 61이며, 회맹막판의 경우에는 47이다. 유문판의 평균 오차가 큰 이유는 2, 5번 환자의 오차가 크기 때문이다.

Table 6. Comparison between the Actual and Predicted Value

Patient No	Predict Pylorus	Predict ic_valve	Real Pylorus	Real Ic_valve	Pylorus Error	Ic_valve Error
1	1495	24498	1456	24459	39	39
2	1724	15893	1595	15847	129	46
3	3263	22086	3219	22048	44	39
4	876	16860	828	16815	48	45
5	126	1084	21	1023	105	61
6	1022	14413	981	14377	41	36
7	922	17639	873	17306	49	63
8	11238	13645	11202	13596	36	49
Unit is Frame				Mean	61	47

위장관 교차점 의심 구역을 지정하기 위해서, 환자마다 저장한 위장관 교차점 의심 구역의 값과 실제 위장관 교차점의 차이 값을 얻고, 모든 환자에 대해 차이 값의 평균을 얻는다. 모델을 평가할 때에는 Equation (3)을 이용한다.

$$(Predict - Error_{mean}) - 50 \leq Suspected\ area \leq (Predict - Error_{mean}) + 50 \quad (3)$$

Predict는 환자마다 저장한 값이고, Suspected area는 최종적인 위장관 교차점 의심 구역이다. 심 구역 안에 실제 위장관 교차점이 들어있다면 성공적으로 예측한 것으로 한다. 이렇게 얻은 오차의 평균값을 Equation (3)을 통해서 의심 구역을 확인해 보면 5번 환자의 경우에는 의심 구역 안에 들어오는 것을 확인할 수 있었다. Fig. 7은 환자들의 데이터를 위장관 교차점 의심 구역을 표시한 것이다. 각 환자의 첫 번째 그래프는 지수이동평균을 거친 값이며, 두 번째와 세 번째 그래프는 위장관 교차점 구역의 값을 확대하여 보여준다. 초록색 구역은 교차점 의심 구역, 빨간색 수직선은 실제 위장관 교차점 구역을 표시했다. 2번 환자는 유문판의 경우에는 실제 위장관 교차점이 의심 구역과 오차 8프레임을 보였다. 유문판의 경우에는 8명중 7명의 환자는 모두 의심 구역 안에 실제 교차점이 있었으며 회맹판은 8명의 환자 모두 의심 구역 안에 실제 교차점이 있다는 것을 확인할 수 있었다.

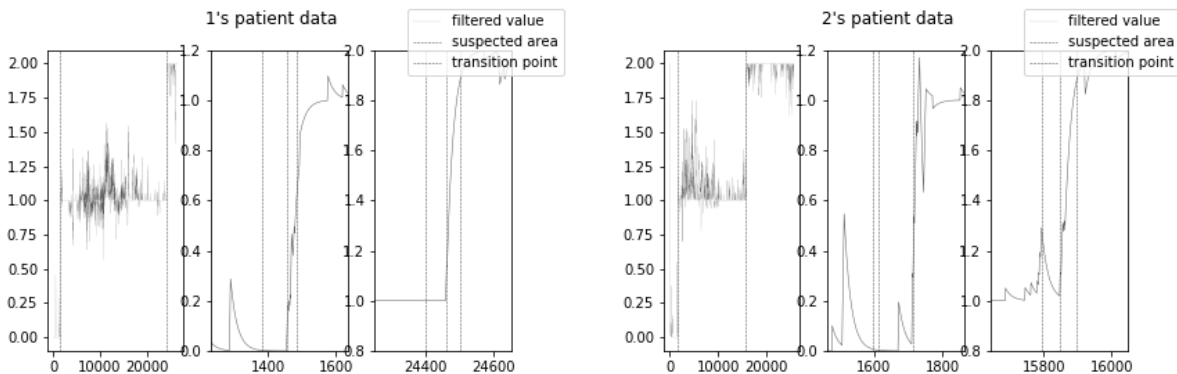


Fig. 7. Result from Gastrointestinal Tract Junction Estimator

5. 평가

위장관 랜드마크 분류기의 성능의 경우에는 노이즈의 유무에 있어 없는 경우의 정확도가 더욱 높았으며, 8명의 환자 데이터를 무작위로 학습했을 때와 8-fold 방식을 이용해서 모델을 평가했을 때 매우 상이한 결과를 볼 수 있었다. 원인을 생각해보면 3.1절에서 분석했듯이 환자마다 데이터의 분포가 상이한데 이를 간과하고 8명의 데이터로 학습을 진행하여 낮은 정확도를 가진 것으로 추측된다. 이는 표본을 늘리면 해결될 수 있다고 판단된다.

또한, 위장관 교차점을 추정할 때 쓰인 모델은 환자의 데이터를 무작위로 섞은 후 노이즈를 제거한 분류기 모델을 사용하였다. 유문판의 경우에는 8명의 환자중 7명의 환자가 의심 구역 안에 들어왔으며, 회맹판의 경우에는 8명의 환자 모두 의심 구역 안에 들어온 것을 확인할 수 있었다. Table 7에서는 기존 연구[11, 13, 14]와 에러값을 비교한 표이다. 회맹판의 중간값 에러보다 확연히 더 적은 오차를 가지며, 평균 오차 또한 더 적은 값을 가진다. 하지만 이는 8명의 환자의 데이터를 사용했다는 점에서 해당 모델이 과적합일 가능성이 있을 수 있기 때문에 이를 개선하기 위해서 더 많은 환자의 데이터를 활용해 모델의 신빙성을 높일 수 있다고 판단된다.

Table 7. Comparison of Performance in Each Method

Label	[11]	[13]	[14]	Proposed Method
Median Error				
Pylorus	211	69	91	44
Ic-valve	1070	-	285	45
Mean Error				
Pylorus	-	381	498	61
Ic-valve	-	-	2607	47

6. 결론

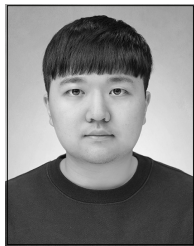
본 연구에서 8명의 데이터를 이용해서 캡슐내시경 영상에서 CNN 랜드마크 분류기를 이용한 위장관 교차점 추정 방법을 제안하였다. 위장관 랜드마크 분류기의 경우 무작위로 샘플링한 후 학습했을 경우에는 95%, 8-fold 방식을 사용해서 학습했을 때에는 평균 71%를 보였다. 위장관 교차점 의심 구역의 프레임은 100프레임으로 판독관이 초당 10프레임의 속도로 영상을 판독한다고 가정하면 판독관은 약 10초안에 위장관 교차점을 찾을 수 있다고 여겨진다. 이번 실험을 근거로 더 많은 환자의 데이터를 이용하여 랜드마크 분류기 및 위장대부분의 환자에게 적용될 수 있는 모델을 학습시키고, 이를 이용해서 일반성과 범용성을 지닌 위장관 위치 추적 모델을 개발 및 고도화할 예정이다.

References

- [1] Geet Litjens, Thijs Kooi, Babak Ehteshami Bejnordi, Arnaud Arindra Adiyoso Setio, Francesco Ciompi, Mohsen Ghafoorian, Jeroen A.W.M van der Laak, Bram van Ginneken, and Clara I. Sánchez, "A survey on deep learning in medical image analysis," *Medical Image Analysis*, Vol.42, pp.60-88, 2017.
- [2] Kim Jeong Rye, Woo Hyun Shim, Hee Mang Yoon, Sang Hyup Hong, Jin Seong Lee, Young Ah Cho, and Sangki Kim, "Computerized bone age estimation using deep learning based program: evaluation of the accuracy and efficiency," *American Journal of Roentgenology*, Vol.209, No.6, pp.1374-1380, 2017.
- [3] Eui Jin Hwang, Sunggyun Park, Jung Im Kim, So Young Choi, Jong Hyuk Lee, Jin Mo Goo, Jaehong Aum, Jae-Joon Yim, Julien G.Cohen, Gilbert R.Ferretti, and Chang Min Park, "Development and Validation of a Deep Learning-Based Automatic Detection Algorithm for Active Pulmonary Tuberculosis on Chest Radiographs," *Clinical Infectious Diseases*, Vol.69, No.5, pp.739-747, doi: 10.1093/cid/ciy967, 2018.11.
- [4] Karargyris, Alexandros, and Nikolaos Bourbakis, "Detection of small bowel polyps and ulcers in wireless capsule endoscopy videos," *IEEE Transactions on BioMedical Engineering*, Vol.58, No.10, pp.2777-2786, 2011.
- [5] Baopu Li and Max Q.-H Meng, "Tumor Recognition in Wireless Capsule Endoscopy Images Using Textural Features and SVM-Based Feature Selection," *IEEE Transactions on Information Technology in Biomedicine*, Vol.16, No.3, pp.323-329, 2012.
- [6] Meryem Souaidi, Said Charfi, Abdelkader Ait Abdelouahad, and Mohamed El Ansari, "New Features for wireless capsule endoscopy polyp detection," *Intelligent Systems and Computer Visions(ISCV)*, 2018 International Conference on IEEE, 2018.
- [7] H. G Lee, H. K Choi, D. H. Lee, and S. C Lee, "Intelligent Diagnosis Assistant System of Capsule Endoscopy Video Through Analysis of Video Frames," *Korea Intelligent Information System Society (KIIS)*, Vol.15, No.2, pp.33-48, 2009.
- [8] D.Y Yoo, Y.S Park, and J.W Lee, "SVM-based Classification of Over-residue Images for Filtering Learning-obstruction Images of Capsule Endoscopy," *Korea Computer Congress (KCC)*, pp.1865-1867, 2018.
- [9] D. Y. Yoo, Y. S. Park, B. J. Lee, and J. W. Lee, "Classification of Noise Interfering with Learning for Medical Image Data-driven Software Development," *Korea Conference on*

Software Engineering(KCSE), pp.317-322, 2019.

- [10] H. Chen, X. Wu, G. Tao, and Q. Peng, "Automatic content understanding with cascaded spatial-temporal deep framework for capsule endoscopy videos," *Neurocomputing*, 229, pp.77-87, 2017.
- [11] J. P. Silva Cuncha, M. Coimbra, P. Campos, and J. M. Soares, "Automated Topographic Segmentation and Transit Time Estimation in Endoscopic Capsule Exams," *IEEE Transaction on Medical Imaging*, Vol.27, No.1, pp.19-27, Jan. 2008.
- [12] Horn, Eli, Hagai Krupnik, and Ofra Zinaty, "System and method to detect a transition in an image stream," U.S. Patent No.7, 684,599, filed Sep.27 2005, and issued Mar 23, 2010.
- [13] Michal Mackiewicz, Jeff Barends, Mark Fisher, and Duncan Bell, "Colour and texture based gastrointestinal tissue discrimination," *IEEE International Conference on Acoustics Speech and Signal Processing Proceedings*. Vol.2, doi: 10.1109/ICASSP.2006.1660413, 2006.7.
- [14] Michal Mackiewicz, Jeff Berens, and Mark Fisher "Wireless capsule endoscopy color video segmentation," *IEEE Transactions on Medical Imaging*, Vol.27, No.12, pp.1769-1781, 2008.



장 현 응

https://orcid.org/0000-0002-2619-745X
 e-mail : jhw3169@gmail.com
 2020년 아주대학교 전자공학과(학사)
 2020년~현 재 아주대학교 전자공학과 석사과정
 관심분야 : Medical Image, Big-Data Analysis, Embedded Software, Machine Learning



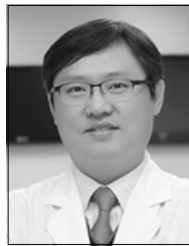
임 창 남

https://orcid.org/0000-0002-4525-5176
 e-mail : chn0714@naver.com
 2019년 아주대학교 전자공학과(학사)
 2019년~현 재 아주대학교 전자공학과 석사과정
 관심분야 : Medical Images, Machine Learning, Deep Learning, Big-Data Analysis, Embedded Software



박 예 슬

https://orcid.org/0000-0003-2584-7489
 e-mail : yeseuly777@gmail.com
 2015년 아주대학교 전자공학과(학사)
 2017년 아주대학교 전자공학과(석사)
 2017년~현 재 아주대학교 전자공학과 박사과정
 관심분야 : Bio-medical Data Modeling, Data Analysis, Medical Images, Ontology, Machine Learning, Deep Learning, Industrial Robot, Embedded Software



이 광 재

https://orcid.org/0000-0002-8534-0850
 e-mail : kjl@ajou.ac.kr
 1988년 연세대학교 의과대학(학사)
 2000년 연세대학교 의학과(박사)
 1992년 내과 전문의 취득
 2001년~2003년 벨기에 Catholic University of Leuven 소화기중개연구센터 연구원
 1995년~현 재 아주대학교 의과대학 소화기내과 강사, 조교수, 부교수, 정교수
 2015년~현 재 아주대학교 의과대학 소화기내과 주임교수겸 임상과장
 2016년~2020년 아주대병원 교육인재개발부원장
 관심분야 : Gastrointestinal Tract, Gastrointestinal Functional Diseases, Gastrointestinal Endoscope



이 정 원

https://orcid.org/0000-0001-8922-063X
 e-mail : jungwony@ajou.ac.kr
 1993년 이화여자대학교 전자계산학과(학사)
 1995년 이화여자대학교 전자계산학과(석사)
 1995년~1997년 LG종합기술원 주임연구원
 2003년 이화여자대학교 컴퓨터학과(박사)
 2003년~2006년 이화여자대학교 컴퓨터학과 BK교수, 전임강사(대우)
 2006년~현 재 아주대학교 전자공학과 교수
 관심분야 : Embedded Software, Automotive Software, Bio·Medical Data Modeling