

딥러닝 기반 폐색전증 진단 보조 소프트웨어 구현

노태성*, 김세정*, 신진우*, 김준형*, 김규겸*, 유종현*, 정길환*, 김대원*, 전홍영*
*원광대학교 의료융합연구센터

sexyroh@icloud.com, sejeong98105@naver.com, wkuwoo@naver.com,
hkim980323@naver.com, kgkim@wku.ac.kr, jhryu@wku.ac.kr, jeongkh1@wku.ac.kr,
kimdw@wku.ac.kr, zip80@wku.ac.kr, ,

Deep learning-based assistance software implementation for pulmonary embolism diagnosis

Roh Tae seong*, Kim Sea Jung*, Shin Jin Woo*, Kim Jun Hyung*, Kou Gyeom Kim**, Jong Hyun Ryu**, Kil Hwan Jeong**, Dae Won Kim**, Hong Young Jun**

*Dept. of Computer Software Engineering, Wonkwang University

**Medical Convergence Research Center, Wonkwang University

요약

폐색전증(Pulmonary Embolism : PE)은 사망률이 높은 질환중 하나이다. 초기에 정확한 진단 및 치료를 시행하는 경우 사망과 재발의 경우가 흔치 않으나 지연 진단이 일어나는 경우 사망률은 높기 때문에 보다 정확하고 빠른 검사법이 필요하다. 본 연구에서는 폐색전증 환자의 흉부 CT 데이터를 가지고 딥러닝 알고리즘을 생성하여 소프트웨어를 구현하고자 한다. 개발된 딥러닝 기반의 폐색전증 진단 보조 소프트웨어는 CT 촬영 후 즉각적인 질환 예측으로 판독 시간의 단축과 효율성을 제공 할 것으로 기대된다.

1. 소개

폐색전증(Pulmonary Embolism : PE)은 높은 사망률을 가진 폐질환으로 심장을 통해 색전이 폐로가는 폐동맥을 막은 것을 의미한다[1]. 폐색전증의 증상으로는 호흡곤란과 흉통, 저혈압 등이 있고 무증상 환자와 비특이적인 증상인 환자도 발생한다[2]. 이로 인해 지연 진단이 발생되며 폐색전증은 초기에 치료를 시행하는 경우 재발률이 낮지만 진단에 지연이 생기면 사망률은 20~30%에 이른다[3]. 대부분 사망의 원인은 치료실패보다 진단의 지연으로 보고되며 지연이 되는 주요 원인은 질환의 증상과 징후들이 비특이적이고 무증상환자 때문이다[3]. 이로인한 지연 진단은 치명적 결과를 초래하기 때문에 보다 정확하고 빠른 검사법이 필요하다. 진단에는 CT 조영술이 많이 시행되고 있으며 조기 진단에 유용하게 사용되고 있다.

최근 4차 산업혁명의 변화 속에서 인공지능은 활발

히 개발되고 있다. 인공지능은 인지, 학습, 추론, 판단 등 인간 사고의 과정을 알고리즘 설계로 구현하는 SW 기술이다. 특히 의료산업에서의 인공지능은 질환진단 보조소프트웨어로 유용한 역할을 수행하고 있으며, 그 중 Convolution Neural Network(CNN) 알고리즘을 이용한 의료 영상 분석 및 진단 알고리즘의 개발이 활발히 이루어지고 있다.

따라서 본 연구에서는 폐색전증의 조기 진단에 활용 할 수 있는 딥러닝 기반 진단보조 소프트웨어를 개발하고자 한다.

2. 소프트웨어 개발

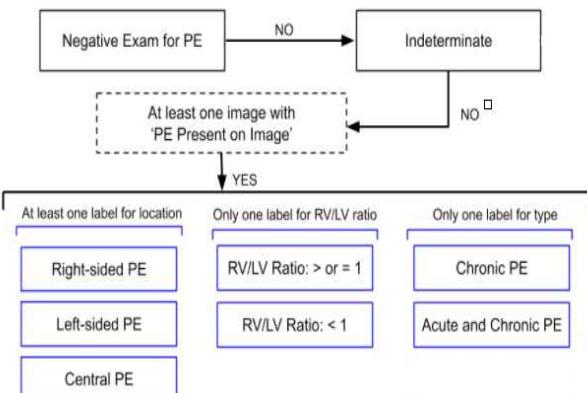
2.1 방법

본 연구에서 사용된 데이터는 Radiological Society of North America 학회에서 주최한 Pulmonary Embolism Detection Challenge 대회에서 제공된 폐

색전증 데이터¹⁾를 활용하였다. 데이터는 흉부 CT Dicom 파일 및 데이터 라벨링 된 csv 파일로 구성되어 있다. Dicom 파일의 데이터 필드로는 환자와 연구의 UID, 양성인 PE의 위치, 급성 폐색전증을 판단할 수 있는 rv/lv ratio, PE의 형태가 아닌 것을 판단하는 부분으로 구성되어 있다.

학습의 첫번째로 라벨링 데이터셋과 Dicom 파일을 Load 후 Train 데이터를 확인한다. 분류된 데이터는 정규화 과정 후 CNN 모델 선택 및 생성 후 데이터 필드와 Output layer를 추가 하였다. 또한 최적화된 activation 함수를 모델에 적용하고, Train 데이터의 학습을 진행하였다. 학습 후 Validation 데이터를 통해 생성된 모델 평가를 하였고 결과로 loss 값과 accuracy 값을 산출하였다.

학습 시 프로토콜은 그림 1과 같다.



(그림1) 데이터 필드의 학습 프로토콜

2.2 개발환경

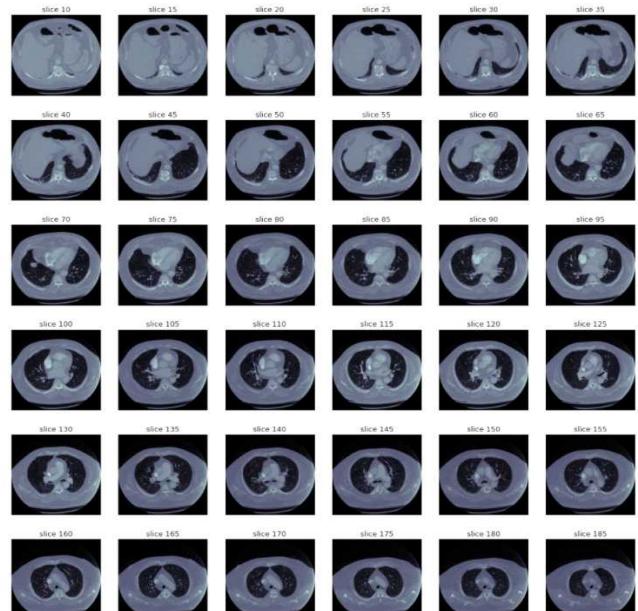
본 연구에서의 개발 환경은 Jupyter, Kaggle notebook, Tensorflow 2.3를 사용하였다. 딥러닝 모델 학습은 온라인 환경인 kaggle notebook에서 진행하였고, 모델의 알고리즘 개발은 Jupyter notebook에서 진행하였다. 딥러닝 알고리즘에 필요한 Library로 keras와 cv, numpy, pandas를 사용하였다.

2.3 학습 데이터 및 분류

폐색전증 데이터는 5개의 1)국제 연구센터에서 제

공하고 80명의 전문의의 임상 주석으로 라벨링된 데이터를 사용하였다. 그림 2는 CT 영상 데이터는 총 912G byte의 데이터로 이루어 있으며 Train 데이터와 Test 데이터로 분류한 것이다. CT 영상 데이터의 데이터 field는 그림 3와 같다. 그림4는 5개의 국제센터에서 전문의들의 임상 주석으로 라벨링이 되어진 Train.csv field를 보여준다. Train 데이터의 경우 모델 학습에 있어서 Validation split은 0.2의 비율로 진행하였다.

CT 영상데이터의 데이터 field를 구성 중 StudyInstanceUID, SerieIntanceUID, SOPInstanceUID는 데이터의 고유 ID를 의미한다. pe_present_on_image는 이미지의 폐색전증 (Pulmonary embolism, PE) 형태의 존재의 유무를 negative_exam_for_pe는 전문의의 연구에서의 PE 형태의 존재여부를 나타낸다. rv_lv_ratio는 우심실과 좌심실의 비율의 정보를 나타내는데, $rv_lv_ratio \geq 1$ 는 axis image에서는 급성(acute) PE에 대해 양성으로 간주된다. 또한 PE의 위치(좌, 우, 중앙)에 대한 필드는 (그림[2]-10,11,13,14,15)에 존재한다. 따라서 소프트웨어가 진단하는 데이터 필드는 그림[2]의 노란색 부분이 된다.



(그림 2) DICOM File의 흉부 CT image

1)

(<https://www.rsna.org/education/ai-resources-and-training/ai-image-challenge/rsna-pe-detection-challenge-2020>)

Data Field	Meaning
1 StudyInstanceUID	연구 데이터의 고유 ID
2 SeriesInstanceUID	연구 내 시리즈의 고유 ID
3 SOPInstanceUID	데이터 이미지의 고유 ID
4 pe_present_on_im age	이미지의 PE 형태의 존재 여부
5 negative_exam_for_pe	데이터필드 4번이 아닌 분류되지 않은 부류 PE에서는 음성
6 qa_motion	전문의의 연구에서 등장 문제를 지적한 여부
7 qa_contrast	전문의의 연구에서 반대의 문제를 지적한 여부
8 rv_lv_ratio_gte_1	RV/LV(우심실/좌심실) ≥ 1 의 여부
9 rv_lv_ratio_lt_1	RV/LV(우심실/좌심실) < 1 의 여부
10 leftsided_pe	좌심실의 PE 형태의 존재 여부
11 chronic_pe	慢성 PE 형태의 존재 여부
12 true_filling_defect_n ot_pe	PE 형태가 아닌 결함의 여부
13 rightsided_pe	우심실의 PE 형태의 존재 여부
14 acute_and_chronic_pe	연구에서 존재하는 PE의 급성,慢성 여부
15 central_pe	정중앙에 PE 형태의 존재 여부
16 indeterminate	연구 상 PE의 음성은 외나치면 OA 문제로 레이블 세트를 만들지 못하는지 여부

(그림 3) DICOM 데이터 field

	0	1	2	3	4
StudyInstanceUID	6897fa9de148	6897fa9de148	6897fa9de148	6897fa9de148	6897fa9de148
SeriesInstanceUID	2bf0b7fd2eab	2bf0b7fd2eab	2bf0b7fd2eab	2bf0b7fd2eab	2bf0b7fd2eab
SOPInstanceUID	c0f3cb03ed06	f57fd3883b6	41220ffda3433	13b0685dab14f	be0b7524ff4
pe_present_on_image	0	0	0	0	0
negative_exam_for_pe	0	0	0	0	0
qa_motion	0	0	0	0	0
qa_contrast	0	0	0	0	0
flow_artifact	0	0	0	0	0
rv_lv_ratio_gte_1	0	0	0	0	0
rv_lv_ratio_lt_1	1	1	1	1	1
leftsided_pe	1	1	1	1	1
chronic_pe	0	0	0	0	0
true_filling_defect_not_pe	0	0	0	0	0
rightsided_pe	1	1	1	1	1
acute_and_chronic_pe	0	0	0	0	0
central_pe	0	0	0	0	0
indeterminate	0	0	0	0	0

(그림 4) Train.csv

2.4 CNN 모델 구축 및 layer 형성

본 연구에서 사용된 컴퓨터 사양에 적합한 CNN 모델을 결정하기 위해 모델의 크기와 Parameter를 고려하였다. Image-net.org에서 보고된 결과(표 1)를 참고하여 정확도는 높고 모델의 사이즈는 작고 Parameter 수가 적은 Xception 모델을 선정하였다. Xception 모델의 weights는 “Imagenet”을 사용하였고 Activation(활성화) 함수로는 logistic, 회귀 분석에 쓰이는 “Sigmoid” 함수와 Connected layer에는 “Relu”를 사용하였다. 또한 Dense 함수를 이용하여 hidden layer의 노드수를 줄여 채널을 감소시켰다. 또한 모델의 손실함수(loss function)는 ‘Binary Cross Entropy’를 사용하였고 Optimizer로는

“Adam”을 사용하였다. 또한 모델 학습에 있어서 batch_size는 1000으로 하였고 Epoch 값은 3으로 하였다.

그림 5는 사용 모델과 추가 된 Layer를 summary 한 것이다.

	size	Parameter	Depth
VGG16	528MB	138,357,544	23
Inception V3	92MB	23,851,784	159
ResNet50	98MB	25,636,712	-
Xception	88MB	22,910,480	126
InceptionResNetV2	215MB	55,873,736	572
ResnetXt50	96MB	25,097,128	-

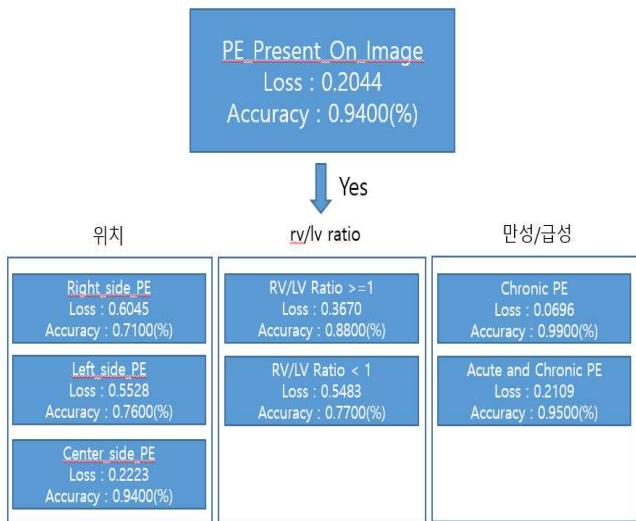
(표 1) ImageNet 데이터를 통한 CNN 모델 비교

Layer(type)	Output Shape	Param#	Connected to
input_1(InputLayer)	(None, 512, 512, 3)	0	
xception (Functional)	(None, None, None, 2)	20861480	input_1[0][0]
global_average_pooling2d	(None, 2048)	0	xception[0][0]
dropout(Dropout)	(None, 2048)	0	global_average_pooling2d[0]
dence (Dense)	(None, 1024)	2098176	dropout[0][0]
dence_1 (Dense)	(None, 256)	262400	dence[0][0]
dence_2 (Dense)	(None, 64)	16448	dence_1[0][0]
acute_and_chronic_pe (Dense)	(None, 1)	65	dence_2[0][0]
central_pe (Dense)	(None, 1)	65	dence_2[0][0]
chronic_pe (Dense)	(None, 1)	65	dence_2[0][0]
indeterminate (Dense)	(None, 1)	65	dence_2[0][0]
leftsided_pe (Dense)	(None, 1)	65	dence_2[0][0]
pe_present_on_image (Dense)	(None, 1)	65	dence_2[0][0]
rightsided_pe (Dense)	(None, 1)	65	dence_2[0][0]
rv_lv_ratio_gte_1 (Dense)	(None, 1)	65	dence_2[0][0]
rv_lv_ratio_lt_1 (Dense)	(None, 1)	65	dence_2[0][0]
Total params : 23, 239,089			
Trainable params : 2, 377,609			
Non-trainable params : 20, 861,480			

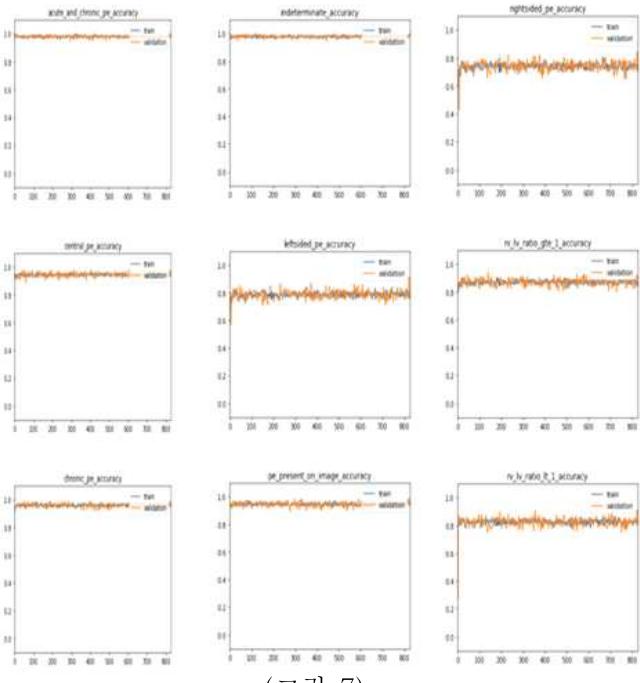
(그림 5) CNN 모델 Layer

3. 학습 결과

그림 6, 그림 7은 Train 데이터로 학습 된 모델의 Validation 데이터로 평가 된 것이다. 총 loss 값은 2.8805이며 가장 큰 loss가 있었던 것은 0.6045의 rightsided_pe_loss이다. 가장 큰 acc값은 0.99의 chronic_pe_accuracy였다. acute과 chronic PE의 진단은 높게 나왔지만, 좌, 우, 중간 위치의 PE의 진단은 낮게 측정됐다.



(그림 6) 학습된 모델의 validation 데이터 평가 값



(그림 7)

학습된 모델의 validation 데이터 평가 값의 시각화

4. 고찰

본 연구는 폐색전증의 조기 진단에 활용 할 수 있는 딥러닝 기반의 진단보조 소프트웨어를 개발하였다. 개발된 소프트웨어는 홍부 CT 영상데이터를 활용한 폐색전증의 발병유무, 만성/급성(acute/chronic)의 분류, 발생위치 분류의 정확성을 평가하였다.

CT영상에서 폐색전증의 발생여부(0.94%)의 진단 및 우심실과 좌심실의 직경 비율(rv/lv ration)에 따른 급성 폐색전증(0.95%)과 만성 폐색전증(0.99%)의

분류는 높은 정확성 나타내었다. 반면, 폐색전의 발병 위치의 분류 정확도는 다소 미흡하게 평가되었다. 향후, 부위별 정확성 향상을 위한 CNN 모델의 개선과 데이터 전처리 과정을 진행할 예정이다. 또한 모델 양상을 통해 전반적인 진단 정확도를 향상 시킬 계획이다.

5. 결론

본 연구에서 개발된 딥러닝 기반의 폐색전증 진단 보조 소프트웨어는 CT 촬영 후 즉각적인 질환 예측으로 판독 시간의 단축과 효율성을 제공 할 것으로 기대된다.

6. 사사

본 연구는 과학기술정보통신부 의료기관 창업 캠퍼스 연계 원천기술개발사업의 지원에 의하여 이루어진 것임. (No. NRF-2016M3A9E9941569

참고문헌

- Dalen JE, Alpert JS. Natural history of pulmonary embolism. Prog Cardiovasc Dis 1975;17:259–70.
- Fedullo PF, Tapson VF. The evaluation of suspected pulmonary embolism. N Engl J Med 2003;349:1247–56.
- Carson JL, Kelly MA, Duff A, Weg JG, Fulkerson WJ, Palevsky HI, et al. The clinical course of pulmonary embolism. N Engl J Med 1992;326:1240–5