

Application of Deep Learning for Detection of COVID-19 in CXR, Cough and CT images

CUAI 3기 Deep Learning Track F팀

박소영(수학), 신선우(컴퓨터예술), 안재우(전자전기), 이재욱(전자전기), 정서연(응용통계), 조영호(소프트웨어)

[요약]

Deep Learning을 통해 COVID-19의 자가진단을 가능하게 하고, 음성 환자를 빠르게 screening하여 진단 과정에서 발생하는 비용을 줄일 수 있는 방안을 제시했다. 기침소리 Spectrum과 CXR, CT image를 데이터셋으로 하여 다양한 인공지능망 모델로 학습시킨 후 테스트 데이터에 적용한 결과 95%에 가까운 높은 정확도를 얻을 수 있었다.

1. 서 론

COVID-19는 높은 치사율과 전파율로 현 인류에게 가장 큰 위협이 되고 있다. COVID-19의 전파를 막기 위해선 치료제, 백신의 개발뿐만 아니라 감염 여부에 대한 빠른 진단이 필수적이다. 우리 조는 Deep Learning을 통해 의료진의 빠른 진단을 돕고자 하였다.

1.1 RT-PCR

코로나19는 SARS-COV-2 바이러스를 통해 전염된다. SARS-COV-2의 RNA를 검출하여 바이러스의 감염 여부를 진단하는 방법이 현재 전세계적으로 가장 많이 사용되는 RT-PCR이다. [1, 2]

RT-PCR 검사의 가장 큰 단점은 시간 소모가 길다는 것이다. 검체 채취부터 유전자검사까지 6시간정도 소요되지만 검체 이송 및 대기시간 등을 고려하면 검사 후 1~2일이 지난 후에 결과를 확인할 수 있다. 또한 RT-PCR 검사의 경우 채취된 검체의 양이 충분하지 않을 경우 거짓 음성이 나올 수 있고, 검진자와 피검진자가 대면으로 검체를 채취해야 하므로 채취하는 과정에서 전염의 위험이 존재한다. 이에 더하여 진단 KIT 수의 부족으로 비용, 장소의 제약이 있다.

1.2 CXR, Cough Spectrum을 통한 COVID-19 진단

RT-PCR 검사의 단점을 보완하여 신속, 정확하고 전염 위험을 제거하면서 비용 문제를 해결할 수 있는 검진 모델이 필요하다. 이에 따라 국내외 연구진 사이에서 인공지능을 이용한 새로운 COVID-19 검진 모델이 활발히 연구되고 있다. COVID-19는 호흡기 질환을 유발하기 때문에 폐를 관찰할 수 있는 X-ray 영상, 그리고 기침 소리의 Spectrum을 데이터셋으로 하여 검진하는 방법을 연구 중이다.[2]

2. 본 론

2.1 이론 및 선행 연구

1) Chest X-Ray (CXR)

코로나바이러스는 바이러스성 폐렴의 일종으로 X-ray상에서 폐렴과 유사하게 촬영된다. 따라서 특정 부위에 눈에 띄는 병변이 생기기 보다 경화로 인해 전반적으로 뿌옇게 변하는 환자가 대부분인데, Figure 1과 같이 검게 찍혀야 할 폐가 전체적으로 뿌옇게 촬영된다.[3]

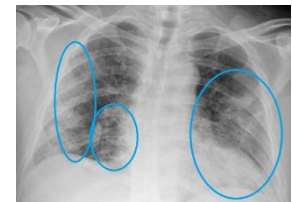


Figure 1. Covid-19 CXR

Ramsey M. Wehbe는 폐를 전면부에서 찍은 X-ray 영상을 데이터셋으로 DeepCOVID-XR이라는 모델을 통해 83%의 정확도로 COVID-19의 감염 여부를 진단해냈다. 위 연구에서는 X-ray영상의 폐 부분을 masking하여 crop하고 down-sampling을 통해 부족한 데이터셋을 보충했다. 해당 연구진은 DenseNet121, ResNet50, InceptionV3, Xception, EfficientNetB2등의 여러가지 모델을 weighted average ensemble하여 DeepCOVID-XR 모델을 설계했다.[4] KAIST의 예종철 교수 연구팀은 DenseNet을 통해 X-ray 영상에서 폐의 이미지를 검출한 후 폐 부분 중 임의의 patch들을 추출하여 데이터셋으로 사용했다. 추출된 patch들은 ResNet18을 통해 학습되었고, 그 결과 폐의 X-ray 영상을 통해 88.9%의 정확도로 COVID-19의 감염 여부를 진단할 수 있었다.[5]

2) Cough Spectrum

기침 소리는 기관의 점액이나 기류의 흐름에 따라 다르게 발생하며 사람의 목소리처럼 기침 소리는 각각 개별화된다. 일반적인 기침 소리는 3개의 phase로 나뉜다. Figure 2, 3와 같이 COVID-19에 의한 기침소리는 healthy cough와 달리 phase2, 3의 energy

distribution이 높게 나타나므로 크기와 스펙트럼이 크고 고주파 성분이 많은 것을 알 수 있다.[6][7]

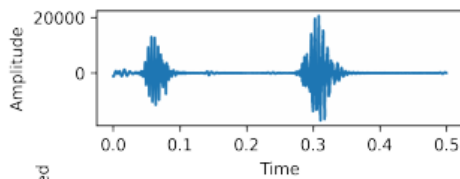


Figure 2. healthy cough

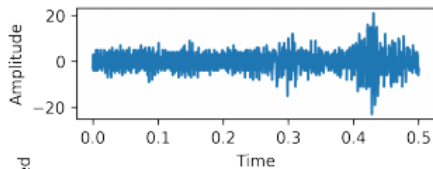


Figure 3. covid-19 cough

MIT Auto-ID 연구팀에서는 기침소리를 통해 COVID-19를 진단하는 모델을 만들었다. 특히, 무증상환자의 경우 100%의 정확도로 감염여부를 진단할 수 있다. 기침의 경우 COVID-19 감염자에게서 가장 흔하게 나타나는 증상이기 때문에 다른 증상이 없는 확진자를 식별하는데 유용하다. 해당 연구팀은 MFCC를 이용해 기침소리를 이미지화 하였으며, 이후 ResNet50을 통한 모델을 설계하여 학습시켰다.[8]

위와 같은 선행연구를 통해 X-ray 영상과 기침 소리 Spectrum에서 COVID-19 감염자의 특이성이 나타난다는 것을 알 수 있다. 이를 이론적 배경으로 하여 본론 2.2에서는 우리가 학습에 사용한 X-ray영상과 기침소리의 dataset과 전처리 과정을 설명하고 2.3에서는 데이터를 통해 COVID-19 감염 여부를 진단하는 방법론을, 2.4에서는 학습에 사용한 딥러닝 모델에 대해서 설명한다. 마지막으로 2.5에서 우리가 제안하는 모델의 결과와 정확도를 평가하였다.

3) Chest CT

CT상 코로나 환자의 가장 뚜렷한 특징은 Figure 4와 같이 폐 결절성 간유리 음영(GGO)의 양상을 보이는 것이다. GGO는 국소적 결절성 폐 침윤으로 기관지, 혈관의 경계가 잘 그려지는 결절로 정의한다. 감염이나 폐렴에서 보이는 광범위한 다발성 간유리 음영과 대조된다.[9]

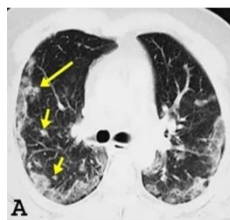


Figure 4. GGO양상의 CT

CT image를 통한 DL관련 연구의 예시로 Arnab Kumar Mishra는 CT영상을 데이터로 다양한 deep CNN기반 여러 개별 모델들을 decision fusion 기법을

사용해 COVID19 예측을 시도하였고, 평균 AUROC와 F1-score에서 각각 0.883과 0.867의 결과를 얻어 내며 CT가 COVID19 진단에 쓰일 수 있는 유효한 데이터임을 보여주었다.[10]

2.2 Data Collection 및 Pre-processing

1) CXR

Chest X-Ray dataset은 병원과 의료진 등의 public source에서 수집되어 GitHub Repository에 commit된 COVID-19 Image와[11], kaggle의 Normal data를 각각 사용하였다.[12]

좌우반전(fliping), 회전(rotation), 움직임(shifting), 절삭(crop) 등의 data augmentation을 적용하여 4044개의 COVID-19 양성 (이하 COVID) cases와 5500개의 COVID-19 음성 (이하 Non-COVID) cases로 구성하였다.

Train data와 Test data의 split비율은 0.8과 0.2로 하였으며, AP(Anterior-Posterior) 와 PA(Posterior-Anterior), 그리고 side-view(lateral)중 COVID-19 진단에 가장 적합한 PA image data를 선정하였다.

Table 1. CXR의 augmented datasets

	COVID	Non-COVID
Datasets	4044	5500
train sets	3235	4400
test sets	808	1100

2) Cough Spectrum

Cough Spectrum은 kaggle의 Cough sample dataset을 사용하였고[13], 330개의 Non-COVID(negative) cases와 1604개의 COVI(positive) cases 분포를 보인다. 이렇듯 약 5배의 데이터 불균형 문제를 해결하기 위해 단순히 negative image 개수를 5배 해주었다.

Table 2. Cough Spectrum의 augmented datasets

	COVID	Non-COVID
Datasets	1604	1650
train sets	1283	1320
test sets	3320	330

데이터 전처리는 다음과 같이 진행하였다. 기존의 RGB Images를 gray Images로 변환하고, 밝기 200을 기준으로 이진화를 수행하였다. 일정 면적 이상의 contour을 찾은 후 사각형으로 mapping하여 꼭지점의 좌표를 계산한다. 400 by 400의 크기로 perspective transform 하고, 네 변의 픽셀의 평균값을 각각 계산하여 어두운 부분이 밑에 위치하도록 회전시켰다.

3) Chest CT

Chest CT dataset은 병원에서 수집되어 GitHub Repository에 commit된 COVID-19 CT Image를 이용한다. 해당 이미지는 medRxiv, bioRxiv, NEJM, JAMA,

Lancet 등의 COVID19 관련 논문에서 수집하였다.[14]

Figure 5와 같이 원본 CT에서 GGO등 이상치가 존재하는 부위를 OpenCV모듈을 이용해 Masking 해주었다.[15]

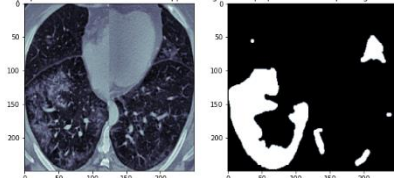


Figure 5. Chest CT와 Masking된 image

2.3 Audio Signal의 MEL Spectrogram 이론적 배경

Audio Signal을 이미지 분석에 사용하기 위해 Mel Spectrogram으로 변환하는 다음과 같은 과정을 거친다. 신호를 20ms에서 40ms 사이의 단위로 나누고, 나눠진 프레임 별로 FFT(Fast Fourier Transform)를 적용한다. FFT는 시간영역에서 시간과 진폭의 축의 오디오 신호를, 주파수 영역에서 주파수와 주파수의 세기의 축으로 변환한다. 이렇게 변환된 Spectrum에 Mel Filter를 적용해 저주파 부분을 확장한다. 해당 과정을 거친 Mel Spectrogram을 분석에 사용한다.

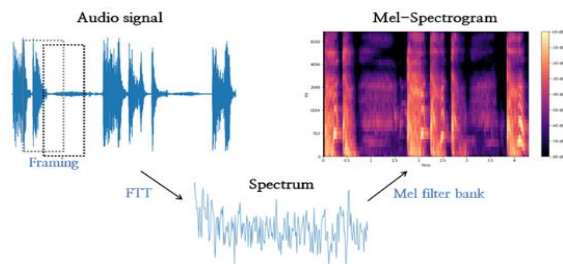


Figure 6. Audio Signal의 Mel Spectrogram

Figure 7과 같이 코로나 양성 유증상, 코로나 양성 무증상, 코로나 음성 세가지 조건의 기침 Audio signal이다.[16] 사람의 귀는 무증상자와 코로나에 걸리지 않은 사람의 기침소리를 구별할 수 없지만, 무증상자는 다시 기침하기 전 숨을 내쉬는 특징이 있다. 유증상자는 보다 짧고 목을 굽는 기침 소리를 낸다.

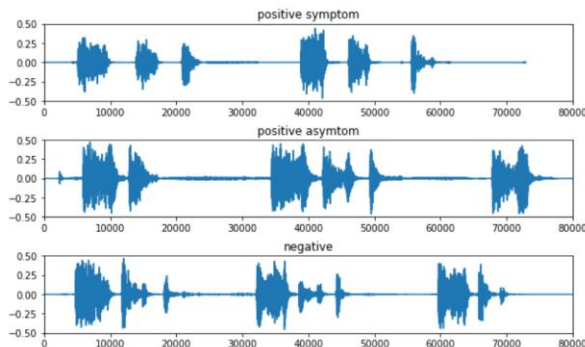


Figure 7. 세 조건의 기침 소리 waveform

Audio signal을 분석하기 위해 python의 librosa 패

키지를 활용하였다. Audio file을 array로 변환하여 신호를 frame으로 나누고, frame 별로 FFT를 적용하는 STFT(Short Time Fourier Transform)를 적용해 Mel Spectrogram으로 반환한다. Stride 10ms 마다 frame 길이를 25ms로 구성하였으며, frame이 서로 겹치도록 구성하고 연속성을 가지도록 하여 비교적 작은 frame 단위의 범위를 선택하여 시간 해상도를 높였다.

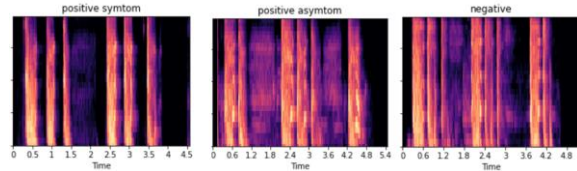


Figure 8. 세 조건의 기침 소리 Mel Spectrogram

2.4 Model과 Network Architecture

1) Grad-CAM

학습된 CXR image에서 COVID-19의 병변을 효과적으로 확인하기 위한 color visualization을 고민하였고, Gradient-weighted Class Activation Mapping (Grad-CAM) 모델을 이용하였다. 기존의 CNN 구조에서는 Fully Connected Layer(FC)를 통과시키기 때문에 Flatten 과정을 거치게 되고 위치 정보들이 소실된다. 이러한 CAM 방식은 Flatten 과정 대신, 각 채널 별로 Global Average Pooling(GAP)을 적용한다.[17] 우리는 MNIST 숫자 원본 image에 CAM을 합성하여 출력해 보았고, 각각 CXR image와 Cough Spectrum image에 이를 적용해보았다.

아래 Figure 9는 COVID-19에 양성 반응을 보인 환자의 CXR image와 Grad-CAM을 적용한 image이다.

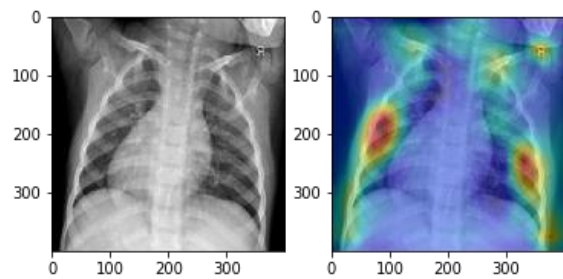


Figure 9. Grad-CAM을 적용한 CXR Image

아래 Figure 10은 COVID-19에 양성 반응을 보인 환자의 Cough Spectrum image와 Grad-CAM을 적용한 image이다.

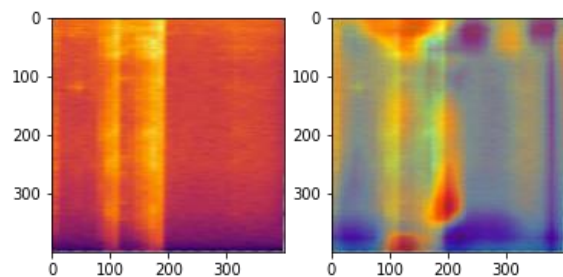


Figure 10. Grad-CAM을 적용한 Cough Spectrum

2) Chest CT의 U-net network 적용

Unet은 end-to-end방식의 FCN기반 모델이기 때문에 입력 데이터의 전처리가 필요하지 않고, 위의 GAP과 마찬가지로 FC가 없어 이미지의 위치를 유지한 채 영상을 출력한다. 따라서 이미지 내 infection part를 파악해 의료 영상처리에 매우 효과적이다.[18]

위 Figure 5 과 같이 원본 CT dataset은 x값(feature)으로, 원본 image에서 infection part masking한 pre-processed dataset은 y값(label)으로 input된다. 3-fold 교차검증으로 train하였고 result는 IOU-score, F-score, Precision, Recall로 평가하였다. 해당 Unet은 평균 F-score, IOU-score, precision, Recall이 각각 약 0.95, 0.9, 0.95, 0.95이고, COVID-19진단에 있어서 효과적인 보조모델임을 알 수 있다.[15]

2.5 Experimental Results와 Evaluation

Cough Spectrum을 학습시키는 데에는 Table 3과 같이 ResNet50, ResNet152, VGG16, VGG19 등의 신경망 모델을 사용하였다.

Table 3. Network별 accuracy와 parameter 개수

Networks	AUC	accuracy	parameters
ResNet50	0.94	0.87	25,636,712
ResNet152	0.96	0.87	60,419,944
VGG16	0.99	0.99	138,357,544
VGG19	0.99	0.98	143,667,240
InceptionV3	0.99	0.99	23,851,784
InceptionRestNetV2	0.99	0.99	55,873,736
Xception	0.99	0.98	22,910,480

ResNet50과 ResNet152의 경우 87%의 정확도, 약 0.95의 AUC 성능을 내어 COVID-19 감염자를 효과적으로 진단해낼 수 있음을 알 수 있었다. VGG16과 VGG19 모델을 사용했을 때에는 98% 이상의 정확도와 AUC 0.99의 높은 성능을 확인하였다.

Table 4. Cough Spectrum data의 Confusion matrix

Actual Class	Predicted Class	
	COVID	Non-COVID
	COVID	Non-COVID
COVID	295	25
Non-COVID	5	325

Table 4과 같이 Cough Spectrum을 학습시킨 결과 TP, FN, FP, TN이 각각 295, 25, 5, 325로, 전체적으로 TP(True Positive rate)와 TN(True Negative rate)이 수치가 높고 특히 FP(False Positive rate) 수치가 낮아 바이러스 진단 모델로 사용하기 적절하다고 판단된다. 하지만 비교적 FN(False Negative rate) 수치 높은 것으로 미루어 보아, 양성 반응을 보인 환자들을 음성이라 판단하는 문제점이 존재한다.

Table 5. Cough Spectrum의 VGG16-GAP Evaluation

# of epochs	precision	recall	AUC	accuracy
10	0.9597	0.9502	0.9890	95%
30	0.9927	0.9964	0.9966	99.4%

Table 5는 Epoch 수에 따른 Cough Spectrum의 VGG-16 evaluation결과이다. 앞서 Grad-CAM을 VGG-16 Network에 수행하였는데, Table 3보다 정확도는 소폭 하락하였음을 확인할 수 있다.

CXR image의 경우 학습에 VGG16을 이용했다. 아래 Table 6과 같이 99.4%의 정확도를 보였고, 0.9964의 재현율을 얻을 수 있었다. 거짓 음성을 피해야하는 CXR 진단 모델에서 높은 재현율이라는 유의미한 결과를 얻을 수 있었다.

Table 6. CXR의 VGG16-GAP Evaluation

# of epochs	precision	recall	AUC	accuracy
10	0.8754	0.8754	0.9529	87.5%
20	0.9833	0.9219	0.9889	95.4%

Epoch의 수를 늘려 정확도를 99.4%까지 올릴 수 있었고, 아래 Figure 11과 같은 epoch당 loss의 양상을 보임을 확인할 수 있다.

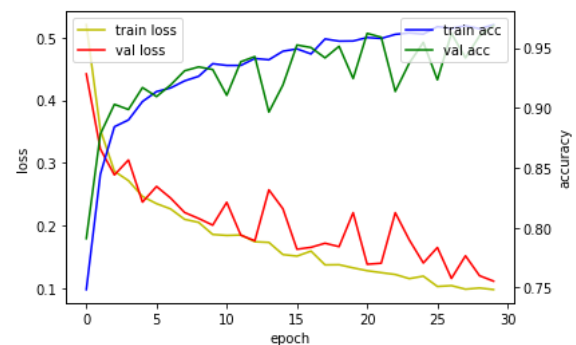


Figure 11. Audio Signal의 Mel Spectrogram

Table 7과 같이 CXR image data를 학습시킨 결과 TP, FN, FP, TN이 각각 547, 2, 4, 400으로, 전체적으로 TP와 TN이 높고 FP와 FN 수치가 낮아 바이러스 진단 모델로 사용하기 매우 적절하다고 판단된다.

Table 7. CXR의 Confusion matrix

Actual Class	Predicted Class	
	COVID	Non-COVID
	COVID	Non-COVID
COVID	547	2
Non-COVID	4	400

3. 결 론

COVID-19 감염자, 비감염자의 흉부 X-ray(CXR) 이미지, 기침소리 스펙트럼 이미지와 여러 딥러닝 모델을 통해 더욱 신속하고 정확한 COVID-19 검진 모델을 만들 수 있었다.

기침소리와 CXR을 통해 COVID-19를 진단하는 것은 바이러스 확산 속도를 늦출 수 있는 결정적인 역할을 할 수 있다고 전망한다. 첫째, 기침소리를 통한 진단은 자가진단을 가능케 한다. 증상이 있지만 RT-PCR 검사에 거부감을 느끼거나, 기침 이외에 증상이 없는 감염자의 경우 기침소리를 통한 자가진단으로 본인의 감염 가능성 여부를 판단할 수 있다. 이는 숨어 있는 감염자의 RT-PCR 검사를 유도할 수 있는 효과가 있다. 둘째, CXR을 통한 신속한 검진으로 음성환자를 빠르게 screening하여 진단KIT의 비용문제를 해결할 수 있을 것이다.

COVID-19 감염자가 빠르게 진단받을 수 있도록 하여 지역사회 전파를 차단하고, 진단에 필요한 인력과 비용을 효율적으로 관리하여 치료제와 백신 개발에 더욱더 힘을 수 있도록 하는데 이번 연구가 유용하게 사용되길 바란다.

참고 문헌

- [1] 부산광역시 홈페이지, Available: <https://www.busan.go.kr/health/corona>
- [2] 질병관리청, 코로나바이러스감염증-19, Available: <http://ncov.mohw.go.kr/#link>
- [3] Wu Jing, Feng Cailian, et al. "Novel coronavirus pneumonia (COVID-19) CT distribution and sign features", Chinese Journal of Tuberculosis and Respiratory Diseases, 2020,43(00)
- [4] Ramsey M. Wehbe, Jiayue Sheng, et al. "DeepCOVID-XR: An Artificial Intelligence Algorithm to Detect COVID-19 on Chest Radiographs Trained and Tested on a Large US Clinical Dataset", RSNA, Radiology, 24 Nov 2020
- [5] Yujin Oh, Sangjoon Park, Jong Chul Ye, "Deep Learning COVID-19 Features on CXR Using Limited Training Data Sets", IEEE, IEEE Transactions on Medical Imaging, 08 May 2020
- [6] Hanieh Chatzarrin, B.Eng, "Feature Extraction for the Differentiation of Dry and Wet Cough

- Sounds", Medical Measurements and Applications Proceedings (MeMeA), 2011 IEEE International Workshop on 2011 May, 2011년
- [7] Ankit Pal, Malaikannan Sankarasubbu, "Pay Attention to the cough: Early Diagnosis of COVID-19 using Interpretable Symptoms Embeddings with Cough Sound Signal Processing", arXiv:2010.02417v2 [cs.LG] 12 Oct 2020
- [8] Jordi Laguarta, Ferran Hueto, Brian Subirana, "COVID-19 Artificial Intelligence Diagnosis Using Only Cough Recordings", IEEE, IEEE Open Journal of Engineering in Medicine and Biology, 29 September 2020
- [9] 한림대학교 의과대학 호흡기내과 장승훈 <https://www.ekjm.org/upload/42844891.pdf>
- [10] Mishra, Arnab Kumar, et al. "Identifying COVID19 from chest CT images: a deep convolutional neural networks based approach." Journal of Healthcare Engineering 2020 (2020).
- [11] <https://github.com/ieee8023/covid-chestxray-dataset>
- [12] <https://www.kaggle.com/paultimothymooney/chest-xray-pneumonia>
- <http://dmqa.korea.ac.kr/activity/seminar/305>
- [13] https://www.kaggle.com/anasealmasry/ai-covid19-from-cough-samples?select=negative_Covid-19
- [14] <https://github.com/UCSD-AI4H/COVID-CT>
- [15] https://colab.research.google.com/github/deadskull7/One-Stop-for-COVID-19-Infection-and-Lung-Segmentation-plus-Classification/blob/master/Notebooks/task1_crossval_3folds_unet.ipynb#scrollTo=ljsQWU-d-ic0
- [16] Artificial intelligence model detects asymptomatic Covid-19 infections through cellphone-recorded coughs. Available: <https://news.mit.edu/2020/covid-19-cough-cellphone-detection-1029>
- [17] Bolei Zhou, Aditya Khosla, Aude Oliva, Antonio Torralba, "Learning Deep Features for Discriminative Localization", 14 Dec 2015
- [18] <https://medium.com/@msmapark2/u-net-%EB%85%BC%EB%AC%B8-%EB%A6%AC%E%B7%B0-u-net-convolutional-networks-for-biomedical-image-segmentation-456d6901b28a>