

# COVID-19 심각성 예측을 위한 딥러닝 전이학습 \*

황예린<sup>01</sup> 김용일<sup>2</sup> 정교민<sup>2</sup>

<sup>1</sup> 서울대학교 협동과정 인공지능

<sup>2</sup> 서울대학교 전기정보공학부

{dpfls589, miles94, kjung}@snu.ac.kr

## COVID-19 Severity Prediction using Deep Transfer Learning

Yerin Hwang<sup>01</sup> Yongil Kim<sup>2</sup> Kyomin Jung<sup>2</sup>

<sup>1</sup> Interdisciplinary Program in Artificial Intelligence, Seoul National University

<sup>2</sup> Department of Electrical and Computer Science, Seoul National University

### 요 약

전 세계적인 팬데믹을 일으키고 있는 COVID-19 전염에 따른 공중 보건상의 추가적인 피해를 감소시키기 위하여, 개인별 신속하고 정확한 진단 및 감염 경로 파악이 요구된다. 또한 COVID-19의 심각성을 예측할 수 있다면 환자별 감염 상태에 따른 유동적인 대처가 가능하기 때문에, 최근 딥러닝 기술을 활용하여 COVID-19의 심각성을 예측하는 연구가 활발히 진행되고 있다. 본 연구에서는 전이 학습을 활용하여, COVID-19 환자의 흉부 X-ray 이미지를 통해 COVID-19의 심각성 예측하는 모델을 제안한다. 본 연구에서 제시하는 모델은 심각성을 결정하는 두가지 척도인 geographic mean과 opacity score의 예측에서 기존의 연구를 매우 상회하는 좋은 성능을 보여주어 COVID-19 심각성 예측에 효과적임을 실험적으로 증명한다.

### 1. 서 론

최근 발생한 COVID-19는 2년이 지난 현재까지도 전 세계적으로 확산세를 펼치고 있으며, 정치, 경제 및 사회 전반적으로 부정적인 영향을 미치고 있다. 이에 다양한 나라에서 COVID-19으로 인한 피해를 줄이기 위한 방역 정책을 제시하고 실천하였다. 하지만 아직까지 입증된 치료제가 없고, 변이 바이러스로 인한 돌파감염 확진자가 증가하고 있다. 2022년 4월 기준 5.1억명 이상의 COVID-19 확진자와 620만명 이상의 사망자가 확인되었다. 이렇듯 전염성과 치사율이 높은 COVID-19은 정확한 진단과 적절한 치료가 필수적이다. 따라서 현재 딥러닝을 활용하여 환자의 흉부 X-ray 사진으로 코로나를 진단하는 것에 대한 가능성 연구는 다양하게 진행되었으며, 대체로 긍정적인 결과를 보인다. 그 결과 2020년 3월부터 다수의 병원에서는 인공지능을 COVID-19 진단에 활용하고 있다.

그러나 현재까지의 연구는 단순히 코로나를 진단하는데 그쳐 그 심각성을 예측하는 데에는 한계가 있다. 공중 보건 상의 여건을 고려했을 때, COVID-19 질병의 심각성까지 예측할 수 있다면, 지금과 같이 의료 시스템이 한정적인 상황에서 환자별 감염 상태에 따른 유동적인 대처가 가능하다. 예를 들어, 심각성이 높게 예측된 환자에게 더욱 빠른 조치를 취할 것이 기대된다. 따라서 본 연구에서는 전이 학습을 이용하여 COVID-

19 심각성을 예측하는 모델을 제안한다. 기존의 여러<sup>1</sup> 질병에 관하여 좋은 성능을 보여준 사전 학습된 모델에 COVID-19 환자의 흉부 X-ray 사진을 전이 학습 방법을 통해 학습시켜 환자 상태의 심각성을 예측한다. 본 연구에서 제안된 모델은 COVID-19의 심각성을 측정하는 척도인 geographic mean 과 opacity score의 예측에서 기존의 연구보다 획기적으로 좋은 성능을 보여준다. 이에 본 연구에서 제안된 모델이 COVID-19 심각성 예측에 매우 효과적임을 실험적으로 증명한다.

### 2. COVID-19의 심각성(Severity) 분석

딥러닝을 활용하여 흉부 X-ray 사진을 통해 COVID-19 환자와 그렇지 않은 사람을 구별하는 연구가 활발히 진행 중이다. 딥러닝을 이용하여 COVID-19 환자를 찾아내는 것은 약 sensitivity 95% 이상, specificity 91% 이상의 결과가 확인되고, 이는 기존의 다른 방법보다 딥러닝을 이용하는 것이 유리하다는 것을 보여준다. 데이터 전처리 등의 어려움이 있지만, COVID-19은 빠른 진단과 치료가 필수적인 만큼 딥러닝을 이용한 빠른 판단은 매우 효과적이다.

이렇게 COVID-19의 진단의 정확도가 확보된 이후, 딥러닝을 활용하여 발병 진행도의 심각성을 예측하는

\*이 논문은 2022년도 정부(과학기술정보통신부)의 재원으로 정보통신기획평가원의 지원을 받아 수행된 연구임 [NO.2021-0-01343, 인공지능대학원지원(서울대학교)]

연구가 진행되고 있다. COVID-19 의 심각성을 나타내는 두 척도로 Opacity score 와 Geographic mean 이 존재한다. 이 두 척도를 바탕으로 COVID-19 환자의 흉부 X-ray 를 전문가가 라벨링한 데이터셋이 심각성 연구에 활용된다. 아래의 그림1에 본 연구에서 활용된 COVID-19 환자의 흉부 X-ray 예시와 두 가지 척도에 대한 라벨값을 볼 수 있다. 이 데이터셋은 2019년 12월부터 2020년 3월까지의 데이터로, 전세계 다양한 나라의 병원에서 촬영한 후 세 명의 전문가가 사진을 보고 점수를 결정한 것이며, 점수가 높을수록 더 심각한 경향성을 띤다. Geographic mean 은 COVID-19의 심각성과 밀접한 관련이 있다고 알려진 폐의 강화 정도나 Lung involvement를 확인하여 매겨진 점수로 왼쪽 폐와 오른쪽 폐의 점수 합산을 통해 0~8 사이의 스케일을 갖는다. Opacity score 는 이와 유사하게 왼쪽 폐와 오른쪽 폐의 합산으로 0~6 사이의 점수를 갖는다.

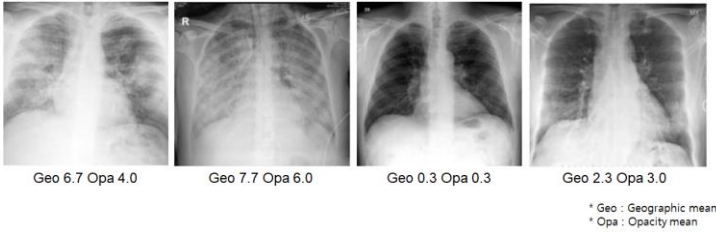


그림 1. COVID-19 환자의 흉부 X-ray 데이터셋 예시

현재 딥러닝을 활용하여 COVID-19의 심각성을 예측하는 연구도 진행 중이다. 이 논문에서는 베이스라인 모델로 Cohen et al.[1]의 모델을 활용한다. 이 연구에서는 여러 질병 데이터셋에 사전학습된 모델을 흉부 X-ray 데이터셋에 조정학습을 진행한다. 본 연구에서는 이 베이스라인 모델의 구조와 Dataset을 바탕으로 다양한 딥러닝 방법론을 활용하여 정확도를 높인 모델을 제안한다.

### 3. 제안된 모델

Feed-Forward 네트워크에서 하나의 layer를 모든 layer에 연결하여 구현된 DenseNet[2]이 흉부 X-ray 사진으로 질병을 진단하는 데 여러 CNN 구조 중에서 가장 효과적이라는 연구 결과가 존재한다[3]. 폐렴이나 기흉 등 14개의 질병을 진단하는 데 있어서 DenseNet 형태의 모델이 다른 모델들에 비해 매우 효과적이라는 결과를 확인할 수 있다.

본 논문에서는 이 DenseNet의 구조를 응용하여 121개의 Layer로 구성된 CheXNet을 100,000개 이상의 흉부 X-ray dataset인 ChestX-ray8[4]을 이용하여 모델을 사전학습시킨다. 그림 2 와 같이 pretrained 된 CheXNet 모델의 가중치를

고정시키고(Freeze) 추가적인 Linear Regression Layer를 COVID-19 환자의 흉부 X-ray dataset으로 미세 조정 학습 (fine-tuning) 하여 실험을 진행한다.

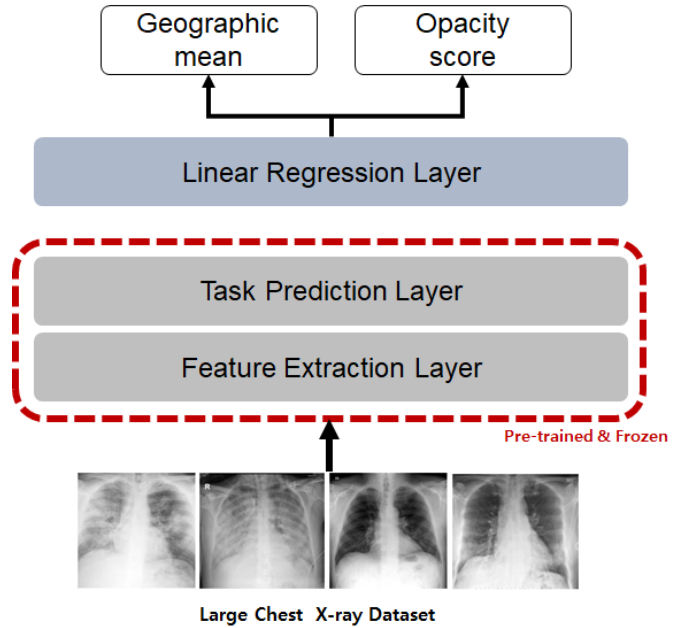


그림 2. 본 연구에서 제안된 모델

### 4. 실험 결과 및 분석

본 논문에서는 학습 데이터셋과 테스트 데이터셋을 8:2의 비율로 나누어 테스트셋에 대하여 실험한 결과를 기록하고 분석한다.

Model	Geographic mean	Opacity score
Joseph et al.	2.060	0.861
18 features	2.9067	1.1456
4 features	2.4742	1.1254
Lung opacity	<b>1.4546</b>	0.9611
+ H-flip	1.4123	0.9543
+ V-flip	1.5180	0.9329
+ H/V-flip	<b>1.4089</b>	<b>0.8437</b>

표 1. 제안된 모델의 실험 결과

#### 4.1. Feature Selection 실험 결과

제안된 모델의 task prediction layer는 Feature Extraction Layer를 거쳐 생성된 1024차원의 표현 벡터를 입력으로 받아 18개의 feature를 출력으로 나타낸다. 본 연구에서는 우선적으로 Task prediction layer가 출력하는 18개의 feature에 대하여 어떠한 조합이 DenseNet에 가장 효과적인지 실험을 진행한다.

표1에서 볼 수 있듯이, Task Prediction Layer가 출력하는 18개의 feature를 모두 사용하는 것보다 4개의 Feature를 선택하여 사용하는 것이 더 좋은

성능을 나타내고, Lung Opacity Feature 만을 사용하는 것이 가장 좋은 성능을 보여준다. Lung Opacity Feature 만을 사용했을 때 베이스라인 모델보다 Geographic mean 에 대해 MSE Loss 가 2.060 에서 1.4546으로 크게 낮아진 것을 확인할 수 있다.

#### 4.2 Data augmentation 실험 결과

다음으로 모델의 성능을 더욱 높이기 위하여 Data Augmentation 기법을 활용한다. 여러 Augmentation 기법 중 Random vertical flip과 Random horizontal flip을 사용할 때 정확도가 증가한 것을 확인할 수 있다. 표 2의 아래에서 Lung Opacity Feature 만을 선택하여 학습에 활용한 모델에 추가적으로 Random horizontal flip 기법을 적용할 때(+H-flip) MSE Loss 가 geographic mean 과 opacity score 에 대하여 모두 감소하여 성능이 좋아진 것을 확인할 수 있다. 마찬가지로, Random vertical flip(+V-flip) 을 활용할 때, geographic mean 과 opacity score 의 MSE loss가 역시 모두 감소한 것을 확인할 수 있다. 최종적으로 두 Data Augmentation 기법을 모두 활용할 때(+H/V-flip) geographic mean 과 opacity score 의 MSE Loss 가 각각 1.4089 와 0.8437 로 기존 베이스라인 모델에 비하여 월등히 좋은 성능을 보인 것을 확인할 수 있다.

결과적으로 본 연구에서 제안된 모델이 COVID-19 의 심각성을 측정하는 척도인 geographic mean 과 opacity score 의 예측에서 기존의 연구보다 획기적으로 좋은 성능을 보여주어 COVID-19 심각성 예측에 효과적임을 실험적으로 증명한다.

### 5. 결론

COVID-19의 심각성을 예측하는 연구는 한정적인 의료 시스템에서 효율적인 처치에 큰 도움을 줄 수 있는 의미있는 연구이다. 본 연구에서는 COVID-19 환자의 흉부 X-ray 이미지로 COVID-19의 심각성을 결정하는 두가지 척도인 geographic mean과 opacity score 를 예측하는 모델을 구현하였다. 여러가지 실험을 통해 다양한 딥러닝 방법론을 활용하여 연구 결과에 비해 획기적으로 좋은 성능을 보이는 모델을 제안하였다.

### 6. 참고문헌

- [1] Cohen, Joseph Paul, et al. "Predicting covid-19 pneumonia severity on chest x-ray with deep learning." *Cureus* 12.7 (2020).
- [2] Huang, Gao, et al. "Densely connected convolutional networks." *Proceedings of the IEEE conference on*

*computer vision and pattern recognition*. 2017.

[3] Rajpurkar, Pranav, et al. "Chexnet: Radiologist-level pneumonia detection on chest x-rays with deep learning." *arXiv preprint arXiv:1711.05225* (2017).

[4] Wang, Xiaosong, et al. "Chestx-ray8: Hospital-scale chest x-ray database and benchmarks on weakly-supervised classification and localization of common thorax diseases." *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2017.