

데이터 증강 기법이 의료 이미지 분류 모델의 예측 성능에 미치는 영향

한예찬, 손우진, 김재윤*
순천향대학교

ychan.sch@gmail.com, swj000317@naver.com, *kimym38@sch.ac.kr

The Effect of Data Augmentation Techniques on Prediction Performance of Medical Image Classification Model

Han Yechan, Shon Woojin, Kim Jaeyun*
Soonchunhyang Univ.

요 약

컴퓨팅 기술의 발전으로 인간의 능력을 뛰어넘는 인공지능이 개발이 시도되고 있으며 이러한 인공지능은 다양한 분야에 통찰력을 제공하고 있다. 의료 분야에서도 인공지능을 접목시켜 환자의 상태를 진단하거나 미래의 질병을 예측하는 작업을 수행하여 의사의 진단에 도움을 주고 있다. 하지만 한정적인 의료 데이터를 구하는 것은 쉽지 않으며 그 양 또한 매우 적은 편이다. 이러한 문제는 예측 모델의 성능 문제로 이어질 수 있기 때문에 본 논문에서는 의료 이미지 데이터에 데이터 증강 기법을 적용하며 증강된 데이터가 모델의 예측 성능 개선에 영향을 미치는지 살펴본다.

1. 서 론

과거 의료 분야에서의 컴퓨터는 환자의 상태에 대한 정보를 기록만 하는 역할을 수행하였다. 하지만 컴퓨팅 기술이 나날이 발전함에 따라 인간의 능력을 뛰어넘는 인공지능이 개발되고 있으며 다양한 분야에서 이를 활용하여 비약적인 성장을 이뤄내고 있다. 최근 의료 분야에서도 인공지능을 접목시키려는 사례가 늘어나고 있으며 의사의 진단에 도움을 주기 위해 노력한다.[1] 많은 의료 인공지능 연구자들은 머신러닝, 딥러닝 같은 기술을 활용하여 환자의 현재 상태를 진단하거나 미래에 발생할 병을 예측한다. 하지만, 의료 데이터의 특성 상 데이터의 수가 매우 적고 한정적이기 때문에 모델 개발 시 다양한 편향이 발생하게 된다. 이러한 편향은 예측 성능 저하, 과적합 등의 인공지능 모델 개발의 고질적인 문제로 이어질 수 있다. 따라서 연구자들은 더 많은 데이터를 수집하기 위해 데이터 증강, GAN 등의 기법을 활용하며 이를 통해 생성된 데이터를 모델의 학습데이터에 추가하여 예측 성능을 끌어올리기 위해 노력하고 있다.[2] 본 논문에서는 의료 이미지 데이터에 데이터 증강 기법을 적용하여 모델의 예측 성능에 얼마나 영향을 미치는지 살펴본다.

2. 본 론

2.1. 실험 데이터

본 논문에서는 Kaggle 의 Dataset 중 Chest X-Ray Images (Pneumonia)을 사용했다.[3] 데이터는 train, val, test 로 나뉘어져 있으며 이 중 train 과 test 를 사용한다. 데이터셋의 Class 는 normal 과 pneumonia 총 2 개가 있으며 데이터에 대한 설명은 표 1 과 같다.

표 1. 실험 데이터 셋 명세

Dataset	Normal	Pneumonia	Image shape
train	1341	3875	일정하지 않음
test	234	390	일정하지 않음

2.2. 실험 방법

본 논문에서는 이미지 분류를 위한 CNN 모델을 Python 으로 직접 설계하였으며 train 데이터셋은 모델 학습에, test 데이터셋은 모델 평가에 활용되었다. 또한, 데이터 증강을 위한 기법으로 RandomZoom (RZ) 방법과 PGGAN 모델을 활용했다. PGGAN 을 통해 생성된 이미지는 그림 1 과 같다. 모델 학습을 위해 input image 의 크기는 128 x 128 로 맞춰주었으며 batch size 는 64, epochs 는 100 으로 설정했다.

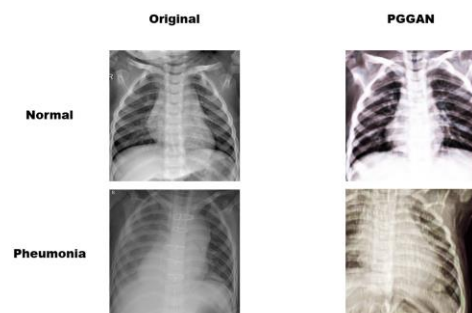


그림 1. 원본 이미지(좌)와 PGGAN 으로 생성된 이미지(우)

2.3. 실험 결과

표 2의 실험 결과 요약표를 보면 RandomZoom 방식을 활용했을 때의 정확도가 0.916으로 가장 높았으며 그 다음은 PGGAN 모델을 통해 생성된 이미지를 추가하여 학습을 진행한 모델의 정확도가 0.833으로 높았다. 요약된 표만 보았을 때는 CNN만 이용하여 학습을 하는 것보다 PGGAN으로 생성된 이미지를 학습 데이터셋에 추가하여 학습하는 것이 좋아 보이지만 그림 2, 4의 세부적인 학습 그래프를 보면 그렇지 않음을 알 수 있다.

그림 2는 데이터 증강 기법을 적용하지 않았을 때의 CNN 모델의 학습 그래프를 나타내며, 학습이 진행되면서 test 데이터셋의 예측 정확도가 0.750에서 횡보하고 있는 것을 볼 수 있다. 그림 4는 PGGAN을 이용하여 학습 데이터를 추가한 CNN 모델의 학습 그래프를 나타내며 test 데이터셋의 예측 정확도가 0.700에서 횡보하고 있는 것을 볼 수 있다. 학습이 진행되는 동안 얻은 최대 예측 정확도는 CNN만 사용했을 때가 0.875, PGGAN 이미지를 추가했을 때가 0.865로 나타났다. 이는 PGGAN으로 생성된 이미지를 학습 데이터셋에 추가했을 때, 모델의 학습이 더 불안정해진다는 것을 의미한다.

그림 3은 데이터 증강 기법 중 RandomZoom 방법을 적용했을 때의 CNN 모델의 학습 그래프를 나타낸다. RandomZoom 방법을 적용한 학습 그래프는 예측 정확도가 0.900에서 횡보하고 있는 것을 볼 수 있으며 학습 중 최대 예측 정확도는 0.932로 다른 방법에 비해 훨씬 우수한 성능을 보여주었다. 이는 랜덤으로 사진을 확대 또는 축소를 하는 데이터 증강 기법이 CT 이미지로 된 pneumonia를 식별하는데 도움을 줄 수 있다는 것을 의미한다.

표 2. 실험 방법 별 실험 결과 요약

Method	Learning data size	Accuracy	Max Accuracy
CNN	5216	0.790	0.875
CNN + RZ	10432	0.916	0.932
CNN + PGGAN	5616	0.833	0.865

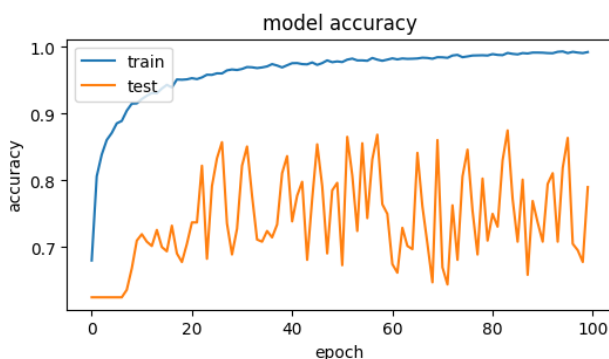


그림 2. 데이터 증강 기법 적용 전 학습 그래프

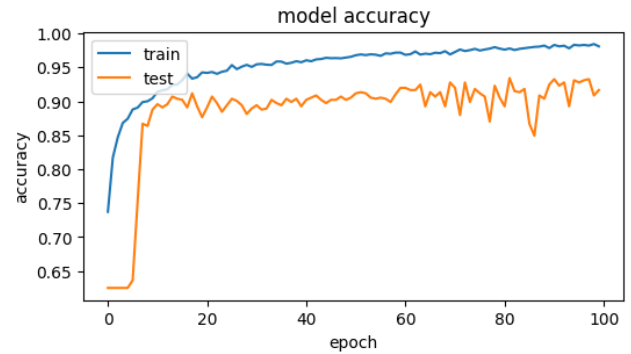


그림 3. RandomZoom 기법 적용 후 학습 그래프

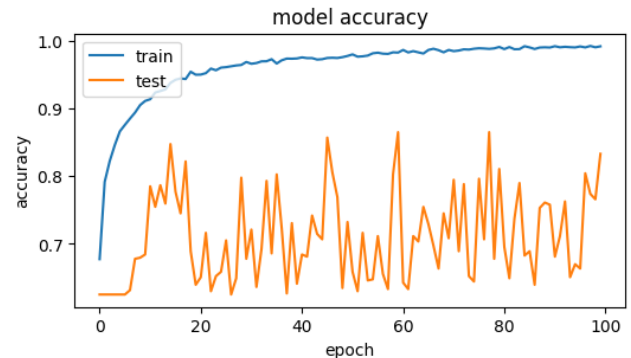


그림 4. PGGAN 기법 적용 후 학습 그래프

3. 결론

Pneumonia를 판별하는 CT image 데이터를 통해 실증분석을 진행한 결과 RandomZoom 방식의 데이터 증강 기법을 사용하는 것이 Pneumonia 분류 모델의 예측 정확도를 크게 향상시킬 수 있다는 것을 확인했다. 또한, PGGAN 모델은 주어진 데이터셋을 잘 생성하지 못해 모델의 학습을 불안정하게 만드는 것을 확인했다. 본 논문은 pneumonia를 판별하는 CT image 데이터에 대해서만 실험을 진행했다는 한계가 있다. 따라서, 논문에서 다룬 데이터 이외의 다른 의료 데이터에도 다음과 같은 데이터 증강 기법을 적용하여 의료 데이터 분석 시 데이터 증강 기법의 유용성을 검증해 볼 필요가 있다.

ACKNOWLEDGMENT

본 연구는 2023년 과학기술정보통신부 및 정보통신기획평가원의 SW 중심대학사업의 연구결과로 수행되었음 (2021-0-01399). 본 논문은 한국연구재단 4단계 두뇌한국 21사업(4단계 BK21사업)의 지원을 받아 작성되었음 (과제번호: 5199990514663).

참고 문헌

- [1] Buch, V. H., Ahmed, I., & Maruthappu, M. (2018). Artificial intelligence in medicine: current trends and future possibilities. *British Journal of General Practice*, 68(668), 143-144.
- [2] Perez, L., & Wang, J. (2017). The effectiveness of data augmentation in image classification using deep learning. *arXiv preprint arXiv:1712.04621*.
- [3] <https://www.kaggle.com/datasets/paultimothymooney/chest-xray-pneumonia>