



Data Augmentation Techniques for Deep Learning-Based Medical Image Analyses

딥러닝 기반 의료영상 분석을 위한 데이터 증강 기법

Mingyu Kim, PhD¹ , Hyun-Jin Bae, PhD^{2*} 

¹Department of Convergence Medicine, University of Ulsan College of Medicine, Seoul, Korea

²Promedius Inc., Seoul, Korea

Medical image analyses have been widely used to differentiate normal and abnormal cases, detect lesions, segment organs, etc. Recently, owing to many breakthroughs in artificial intelligence techniques, medical image analyses based on deep learning have been actively studied. However, sufficient medical data are difficult to obtain, and data imbalance between classes hinder the improvement of deep learning performance. To resolve these issues, various studies have been performed, and data augmentation has been found to be a solution. In this review, we introduce data augmentation techniques, including image processing, such as rotation, shift, and intensity variation methods, generative adversarial network-based method, and image property mixing methods. Subsequently, we examine various deep learning studies based on data augmentation techniques. Finally, we discuss the necessity and future directions of data augmentation.

Index terms Deep Learning; Medical Imaging; Data Analysis

서론

1943년에 처음 개발된 인공신경망(Artificial Neural Network) 기술은 여러 차례 암흑기를 맞으며 발전해왔다. 당시 인공신경망 학습에 가장 큰 문제는 학습층을 깊게 만들 수 없고, 충분한 연산 능력을 확보하지 못했다는 점이었다. 2012년 컨볼루션 뉴럴넷(Convolutional Neural Network)을 여러겹으로 쌓아 만든 심층신경망(Deep Neural Network)인 알렉스넷(AlexNet)이 이미지넷 분류 대회에서 사람보다 높은 성능을 보이면서 컴퓨터 비전 분야에서 각광을 받기 시작했고, 이후 이러한 신경망 학습 기법을 딥러닝(deep learning)이라 부르게 되었다. 현재 딥러닝 기술은 의료영상의 다양한 분야에 적용되어, 의료영상을 특정 질환

Received September 2, 2020
Revised September 22, 2020
Accepted September 24, 2020

*Corresponding author
Hyun-Jin Bae, PhD
Promedius Inc.,
B-604, 10 Olympic-ro 35ga-gil,
Songpa-gu, Seoul 05510, Korea.

Tel 82-2-423-8511
Fax 82-0303-3445-8511
E-mail hjbae@promedius.ai

This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (<https://creativecommons.org/licenses/by-nc/4.0>) which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

ORCID iDs

Mingyu Kim 
<https://orcid.org/0000-0003-1236-2220>
Hyun-Jin Bae 
<https://orcid.org/0000-0001-5134-5517>

에 따라 분류하거나, 병변의 위치를 찾고, 장기를 분할하는 등 다양한 형태로 사용되고 있다(1, 2).

딥러닝은 많은 양의 데이터와 이에 대한 정답(레이블)을 동시에 학습시키는 방식으로 학습이 이루어진다. 이를 통해 딥러닝 모델은 학습한 데이터셋에는 높은 성능을 보이게 되지만, 새로 입력 받는 데이터의 속성이 학습에 쓰인 데이터의 속성과 차이를 보이는 경우 일반적으로 낮은 성능을 보인다. 이러한 현상을 과적합(overfitting)이라 부르며, 과적합을 줄이기 위한 다양한 기법이 적용되고 있다(3). 예를 들어, 정규화(regularization)나 Dropout, 모델 복잡도를 낮추는 등의 방법이 효과를 보이고 있지만(4), 가장 좋은 방법은 학습 과정에서 대량의 데이터를 학습시키는 것이다(5). 학습용 데이터는 데이터의 다양한 특징들을 대부분 포함하고 있어야 하며, 구별하고자 하는 특징들이 골고루 분포하고 있어야 한다. 하지만, 의료영상은 개인 정보 문제 등으로 인해 대량의 데이터를 수집하기가 어렵다. 또한, 정답 데이터(레이블)를 만들기 위해 의료인의 노력이 필요하기 때문에 많은 비용이 든다는 어려움이 있다(6, 7), 또한, 데이터를 수집하는 과정에서 각 질환의 유병률 등에 따라 수집된 데이터의 균형이 맞지 않을 가능성이 높다(8). 이처럼 원천 데이터를 추가로 구하기 어려운 상황에서 데이터의 다양성을 학습시켜 딥러닝의 성능을 높이고 싶을 때 사용하는 기법이 데이터 증강(data augmentation)이다. 즉, 데이터 증강 기법은 모델 학습에 사용되는 데이터의 다양한 변이를 배우게 함으로써, 학습 모델의 과적합을 방지하기 위해 유용하게 쓰인다(9, 10).

데이터 증강 기법을 크게 분류하면, 전통적 영상처리 기반의 증강 기법과 적대적생성네트워크(Generative Adversarial Network; 이하 GAN) (11) 기반의 증강 기법이 있다. 전통적 영상처리 기반의 증강 기법에는 영상의 회전, 색 변환, 블러링 등이 있는데, 이러한 영상 밝기나 공간적 정보를 변형하는 등의 데이터 증강 기법은 변형 전 원본 영상의 속성과 대개 닮아 있다. 따라서, 영상이 가진 속성들의 다양한 변이를 배우지 못해 큰 성능 향상을 거두지 못하는 경우가 많다. 이러한 한계를 극복하기 위해 GAN 기반의 데이터 증강 기법이 다양한 연구에서 적용되고 있다. 또한 최근에는 학습 영상의 일부를 잘라내거나 영상들을 섞는 방식으로 간단하면서도 학습 성능을 크게 끌어올릴 수 있는 데이터 증강 기법들도 소개되고 있다.

본 종설에서는 딥러닝 연구에서 유용하게 쓰이는 다양한 데이터 증강 기법에 대해 정리하고, 이러한 기법을 활용한 대표적인 영상의학 관련 딥러닝 연구들을 소개하고자 한다. 마지막으로 딥러닝 연구에 있어 데이터 증강의 필요성을 고찰하고 앞으로 해결해야 할 문제들을 짚어본다.

영상처리 기반 데이터 증강 기법

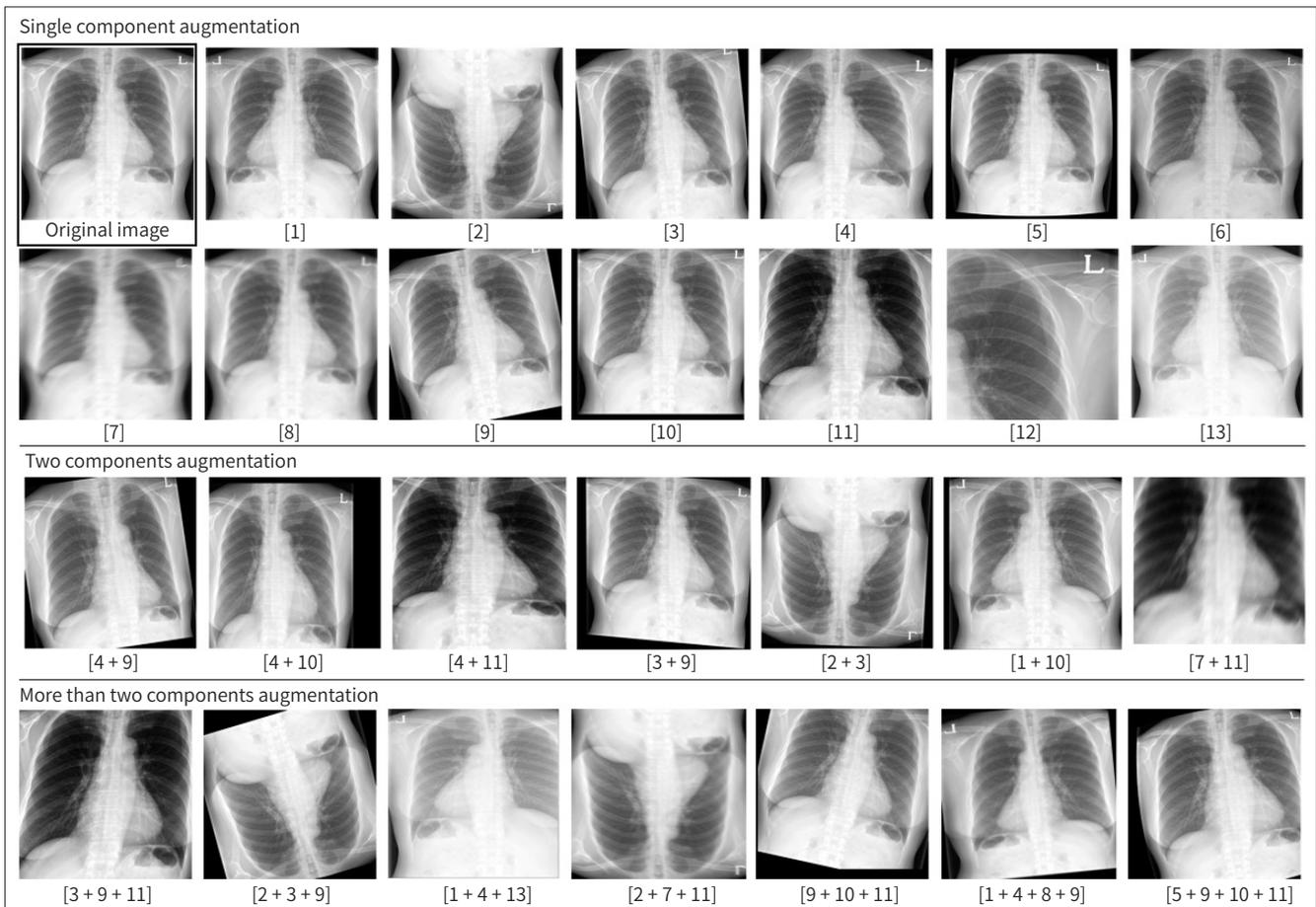
가장 기본적인 데이터 증강 기법 중 하나는 영상처리를 기반으로 하는 기법이다. 이는 입력 영상에 대해 다양한 영상처리 기법을 적용하여 원본의 특성을 살리면서도 다양한 특성을 갖도록 만드는 방법이다. 영상처리 기반 데이터 증강 기법은 기존의 영상처리 알고리즘을 활용하여 쉽게 구현할 수 있고 모델의 성능을 높이는 데 있어서도 효과적이기 때문에 꾸준히 사용되고 있다. 데이터 증강에 사용되는 영상처리 기법에는 무작위로 뒤집기, 회전하기, 자르기 등 영상의 물리적 형태를 변환하는 기법부터, 역상, 채널 섞기, Contrast Limited Adaptive Histogram Equalization 등 영상의 색상이나 밝기를 변화하는 기법들까지 다양하다. 실제 딥러닝 학습에서는 이러한

증강 기법을 단독으로 사용하는 경우보다는 여러 기법을 함께 사용하거나, 2개 이상의 기법을 동시에 적용하는 경우들이 대부분이다.

최근에는 영상처리 기반 데이터 증강을 위한 파이썬 라이브러리도 개발되어, 연구자들이 편리하게 데이터 증강 기법을 사용할 수 있도록 도와주고 있다. 그중 대표적인 파이썬 라이브러리 중 하나는 Albumentations (<https://albumentations.ai/>)다. Albumentations는 Pytorch, Keras 등 다양한 딥러닝 프레임워크에서 동작할 수 있도록 설계되었기 때문에, 기존의 코드에서 단지 몇 줄의 코드를 추가하는 방식으로 쉽게 적용할 수 있다. 또한 입력되는 영상의 레이블이 경계를 갖는 박스이거나, 키포인트, 혹은 마스크인 경우, 해당 레이블도 같은 방식으로 변환시켜 주기 때문에, 영상 분류뿐만 아니라, 영상 검출, 영상 분할 등 다양한 연구에 적용할 수 있다. 또 다른 파이썬 라이브러리 중 하나는 imgaug (<https://github.com/aleju/imgaug>)이며, 앞서 언급한 라이브러리와 거의 유사한 기능을 갖고 있으므로 사용자의 취향에 따라 골라서 사용하면 된다. Fig. 1에 앞서 소개한 라이브러리 중 하나인 Albumentations를 이용하여 증강한 흉부 X-ray 영상의 예시를 보였다.

위와 같은 영상처리 기반 데이터 증강 기법은 다양한 의료 인공지능 연구에서 성능을 높이기 위

Fig. 1. Various traditional augmentation examples of chest X-ray images using Albumentations. Numbers below images represent: [1] horizontal flip, [2] vertical flip, [3] elastic transform, [4] grid distort, [5] optical distort, [6] sharp, [7] motion blur, [8] gaussian blur, [9] rotation, [10] shift, [11] zoom, [12] crop, and [13] gamma correct.



한 기본적인 기법으로 사용되고 있다. Roth 등(12) 연구에서는 딥러닝을 통해 CT 영상 내 다섯 가지 장기를 분류하는 연구를 수행했다. 이 과정에서 랜덤이동, 랜덤회전, 그리고 비강체변형(non-rigid deformation)을 통한 데이터 증강 기법을 적용했다. 데이터 증강 기법을 적용하기 전과 후의 평균 area under the curve (이하 AUC)를 비교해보면 0.994에서 0.998로 증가했고, 오차율은 9.6%에서 5.9%로 감소함을 보였다. 이러한 결과는 데이터 증강 기법이 딥러닝 모델의 성능을 올리기 위해 필수적인 과정이라는 것을 잘 보여주고 있다.

Hao 등(13) 연구에서는 딥러닝을 통한 MRI 영상에서 전립선암 검출 연구를 수행하면서 5개의 영상처리 기반 데이터 증강 기법(회전하기, 위아래 뒤집기, 좌우 뒤집기, 자르기, 이동하기)을 각각 적용하여 각각의 기법들이 모델 성능에 어떻게 영향을 미치는지 살펴보았다. 그 결과 랜덤회전 기법을 적용한 모델의 AUC (85.04%)가 기본 모델의 AUC (80.84%)에 비해 5.2% 더 높은 성능을 보이며 가장 효과적인 증강 기법으로 파악되었다. 이러한 이유는 전립선 MRI 영상이 좌우대칭이기 때문인 것으로 보인다. 전립선의 대칭적 형태 때문에 뒤집기 등의 경우 데이터 증강 효과가 적지만, 회전을 통해 다양한 데이터 형태를 학습시킬 수 있는 것이다. 본 연구를 통해 우리는 영상의 특징에 맞게 최적의 데이터 증강 기법을 선택할 필요가 있다는 점을 알 수 있다.

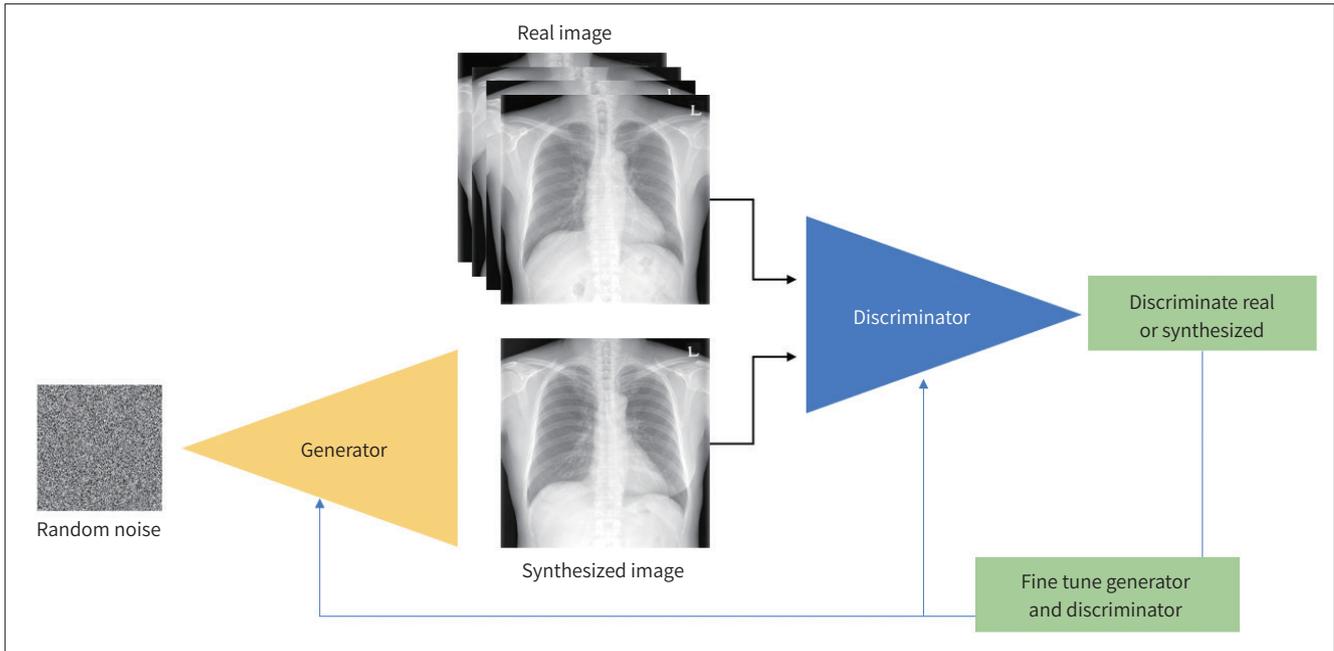
GAN 기반 데이터 증강 기법

GAN은 비지도학습(unsupervised learning) 방식의 인공지능 알고리즘 중 하나로 입력 데이터의 특징을 스스로 학습하여 새로운 데이터를 생성해내는 기법이다(11). GAN의 학습 과정에는 영상을 생성하는 생성기(generator)와 생성된 영상과 실제 영상을 판별하는 판별기(discriminator)가 사용된다(Fig. 2). 학습의 시작에 생성기는 무작위 노이즈를 초기값으로 입력받아, 생성기 네트워크를 통과시켜 영상을 생성해 낸다. 판별기는 입력 영상의 특징들을 학습하여 실제 영상과 합성 영상을 판별한 뒤, 생성기 네트워크를 업데이트한다. 이렇게 생성, 판별 및 피드백, 모델 수정 과정을 반복하여 학습이 끝나면, 학습된 생성기만 떼어내어 임의의 잠재 벡터(latent vector)에 대해 새로운 합성 영상을 만들 수 있게 된다. 이렇게 만든 합성 영상을 통해 학습 데이터 증강이 가능하다(14).

GAN의 영상 생성 기법은 의료영상 분석의 다양한 측면에 이용될 수 있다. 먼저 의료데이터는 생체 정보를 담고 있는 환자의 개인 정보로서 의료데이터를 활용하여 연구를 하는 것에 법적 윤리적 책임이 따른다. 앞서 기술한 바와 같이 GAN의 학습에는 환자의 개인 정보가 담긴 의료데이터를 사용하긴 하나, 학습이 완료된 뒤에는 학습된 생성기만 떼어내어 영상을 생성해낸다. 생성기는 네트워크를 이루고 있는 가짜 영상을 생성하기 위해 학습된 가중치만 존재하므로, 환자의 개인 정보를 역으로 유추해낼 수 있는 방법은 없다. 따라서, GAN을 이용하여 생성한 데이터를 연구에 사용함에 있어 법적 윤리적 책임을 피할 수 있다.

둘째, GAN의 데이터 증강 기법을 이용하여 인공지능 학습의 결과를 향상시킬 수 있다. 인공지능 학습의 정확도는 영상의 개수에 비례하는데, 이는 인공지능이 영상에 담겨 있는 다양한 속성을 배우기 위해서는 다양한 변이를 가진 영상이 필요하기 때문이다. GAN을 이용하면 인공지능 학습에 필요한 영상을 무한히 생성할 수 있으며, 따라서 인공지능 학습의 결과를 향상시킬 수 있다.

Fig. 2. Basic architecture of generative adversarial network with example images of chest X-ray.



셋째, 인공지능 분류와 검출 문제에서는 각 클래스별 영상 수의 비율이 학습 결과에 큰 영향을 끼친다. 클래스별 영상 수의 비율이 다를 경우, 영상수가 많은 클래스의 영상 속성을 학습하므로, 학습 모델에 편향이 생기게 된다. GAN을 활용하여 각 클래스별 영상수를 증강시켜 학습시키면, 특정 클래스에 모델이 편향되는 문제를 방지할 수 있다.

넷째, 다른 도메인에서 찍은 영상이 특정 속성에 대해 영상학적 차이가 있을 경우, 임의의 도메인 속성에서 특정 도메인 속성으로 변환을 할 수 있다. 예로, 저선량(low dose) CT 영상과 일반선량(routine dose) CT 영상이 있을 때, 두 도메인의 데이터들 간의 차이를 학습하여 저선량 CT 영상을 일반선량 CT 영상으로 변환시키거나, 역으로 일반선량 CT 영상을 저선량 CT 영상으로 변환하는 것이 가능하다(15). 이를 이용하면, 특정 기관의 영상으로 학습한 모델을 다른 기관에서 사용할 때, 다른 기관의 영상을 모델이 만들어진 특정 기관의 영상의 속성으로 변환하여 사용할 수 있다. 또한, 이를 이용하여 특정 군에서만 보이는 노이즈를 제거(denoising) 할 수도 있다(16, 17).

최근 의료영상 연구에서는 GAN을 통한 데이터 증강 기법을 적용한 연구들이 다양하게 소개되었다. Frid-Adar 등(18)은 간 CT 영상에서 낭포 영역, 전이 영역, 혈관종 영역을 분류하였다. 총 182명 환자의 데이터에서 53개의 낭포 영역, 64개의 전이 영역, 그리고 65개의 혈관종 영역의 영상을 얻었고, 영상처리 기반으로 증강한 데이터와 deep convolutional GAN (이하 DCGAN) (19)을 이용하여 증강한 데이터를 포함한 데이터를 각각 학습시켰다. 그 결과 영상처리 기반으로 증강한 데이터에 GAN으로 증강한 데이터를 합하여 학습한 것이 민감도에서 10% 정도의 증가를 보였고, 특이도에서 5% 정도의 증가를 보였다. 본 연구를 통해 적은 데이터로도 GAN을 이용하여 증강한 데이터를 이용하면 분류 성능을 높일 수 있다는 것을 알 수 있다.

Zhao 등(20)은 데이터 증강 기법을 통해 폐암분류의 정확도를 높였다. 폐암의 다양성을 고려했

을 때, 기존 연구에서 폐암분류에 사용하는 데이터의 개수가 현저히 부족하다. 본 연구에서는 Forward and Backward GAN이라는 두 단계의 GAN을 사용하였는데, Forward GAN이 다양한 영상을 만들면, Backward GAN은 영상의 노이즈를 제거하여 품질을 향상시킨다. 이렇게 생성된 영상을 이용해 악성 결절과 양성 결절 분류를 진행하였다. 연구 결과, 단순히 영상처리만으로 데이터 증강을 한 학습 결과보다, 영상처리 기반의 데이터 증강을 한 후 다시 Forward GAN을 이용하여 데이터 수를 두 배 가까이 늘린 모델에서 분류 정확도가 4% 정도 더 높게 나왔다. 또한, Backward GAN을 이용하여 데이터 정제 작업을 추가로 하면 1% 정도 성능을 더 높일 수 있음을 보였다. 본 연구는 GAN을 이용한 데이터 증강의 영향뿐 아니라, 노이즈가 제거된 영상이 딥러닝 성능을 향상시킬 수 있음을 알 수 있게 한다.

Salehinejad 등(21)은 흉부 X-ray 영상에서 정상을 포함한 5가지 질환을 분류하는 연구를 진행하였다. 실험에 사용한 실제 데이터는 정상 환자 15781명, 심비대증 환자 17098명, 흉막 삼출 환자 14510명, 폐부종 환자 5018명, 기흉 환자 4013명의 흉부 X-ray로 클래스 간 데이터 불균형이 심한 것을 알 수 있다. 데이터 불균형이 분류 성능에 끼치는 영향과 GAN으로 증강하여 불균형을 없앤 것이 분류 성능에 끼치는 영향을 연구하기 위해 다음 세 가지 연구를 진행하였다. 첫 번째로 클래스 간 불균형이 있는 데이터를 그대로 학습에 사용하였다. 두 번째는 과소표집(undersampling)을 통해 클래스 간 불균형을 균형 있게 갖춘 데이터를 학습에 사용하였다. 과소표집이란 데이터의 수가 가장 작은 클래스에 맞춰 다른 모든 클래스의 데이터를 동일한 개수로 추출하는 방법이다. 마지막으로, DCGAN을 이용하여 데이터를 증강해 클래스 간 불균형을 균형 있게 맞추어 학습을 하였다. 그 결과 각 질환별 분류 정확도의 평균이 과소표집을 통해 클래스 간 데이터 수의 균형을 갖춘 경우가 가장 낮았고(58.90%), 불균형 데이터를 그대로 학습한 경우가 두 번째(70.87%), 그리고 GAN을 이용해 증강한 경우가 월등히 높은 성능(92.10%)을 보였다. 이 결과를 통해 GAN을 이용해 데이터 수를 증강하고 클래스 간 데이터 수의 균형을 맞추는 것이 분류 성능을 크게 높일 수 있음을 알 수 있다.

GAN을 이용한 데이터 증강 기법은 다음과 같이 병변 검출에도 사용되고 있다. He 등(22)은 180명의 환자로부터 얻은 256×256 사이즈의 뇌 MR 영상을 progressive growing of GANs (이하 PGGAN) (23)으로 학습하여 생성하였다. 생성된 영상의 성능을 평가하기 위해 YOLOv3 (24) 모델을 이용해 실제 영상만 가지고 종양 검출을 한 것과 GAN으로 증대한 영상을 합쳐서 종양 검출을 한 결과를 비교하였다. 그 결과 GAN으로 증대한 영상을 합쳐서 종양 검출을 학습한 결과의 평균 정밀도가 실제 영상만 이용한 경우에 비해 10% 성능 향상을 보였다.

조영증강 MR 영상은 간암 환자의 진단에 필요한 영상이다. 하지만 gadolinium을 이용한 조영증강은 부작용 위험이 크고 비용이 많이 든다는 단점이 있다. Zhao 등(25)은 조영제의 주입 없이 조영증강과 유사한 MR 영상을 생성하였고, 높은 정확도(89.4%)로 간암을 검출하는 것에 성공하였다. 본 연구에서는 tripartite-GAN을 사용하여 합성 영상 생성과 종양 검출을 동시에 수행하였다. 또한, 생성자에서 산출된 attention map을 활용하여 종양 검출 성능이 향상될 수 있음을 보였다.

GAN을 이용하여 증강한 데이터를 활용하면 영상 분할 성능도 높아진다는 연구도 진행되고 있

다. Bowles 등(26)은 CT와 MR 영상을 이용하여 뇌척수액 분할 성능을 평가하였다. 영상처리 기반으로 증강한 데이터와 GAN으로 증강한 데이터 각각에 대한 학습결과와 두 가지 기법을 함께 적용한 데이터에 대해 학습한 결과를 비교하였다. 그 결과 두 가지 증강 기법을 동시에 사용한 경우 1~5% 사이의 Dice 점수 향상을 얻을 수 있었다. 본 연구 결과를 통해 GAN은 학습 데이터의 분포를 함수화하여, 함수의 내삽(interpolation)을 통해 영상을 생성하지만, 학습 데이터의 분포를 벗어나지는 못한다는 것을 알 수 있다. 따라서, 적절한 영상처리 기반의 데이터 증강을 이용해 학습 데이터의 다양성을 높여주고, GAN을 이용해 학습 데이터의 특징을 내삽함으로써 다양한 영상을 생성해 분할 성능을 향상시킬 수 있음을 알 수 있다.

Russ 등(27)은 20개의 조영증강 복부 CT에서 증강 기법에 따른 혈관 분할 성능을 평가하였다. XCAT 팬텀을 CycleGAN (28)을 이용해 CT로 변환하여 데이터를 증강하였다. 실제 CT 영상의 혈관 분할 성능이 XCAT 팬텀으로부터 변환된 CT 영상을 학습한 혈관 분할 성능보다 높았지만(Dice 점수 = 0.78), 두 개의 데이터를 모두 사용한 경우에 Dice 점수가 0.83으로 가장 높았다. 이를 통해 CycleGAN으로 변환한 영상이 실제 데이터와 완벽히 일치하지는 않지만, 다양한 데이터 분포를 학습함으로써 분할 성능을 높인 것을 알 수 있다.

지금까지 살펴본 단순 생성 GAN과 달리, 영상 변환에 유용한 GAN을 활용하여 영상이 부족한 클래스의 데이터 수를 늘린 연구도 있다. Gupta 등(29)은 X-ray 영상(상완골, 경골, 대퇴골)에서 병이 있는 뼈를 분류하는 연구를 진행하였다. 하지만, 병이 있는 영상 데이터의 부족으로 클래스 간 불균형이 심하였다. 이를 해결하기 위해, 영상 변환 기법을 활용하여 병이 없는 영상을 병이 있는 영상으로 변환(image translation) 하여 데이터 증강을 시도하였다. 그 결과 증강하지 않은 원본 데이터로 학습한 것보다 증강한 데이터로 학습한 것이 경골 분류의 AUC를 5% 정도 향상시켰다.

일반적으로 조영증강 CT가 종양 분류 등에 더 유용하여, 조영증강 CT를 많이 찍게 된다. 하지만, 환자의 상태에 따라 조영증강 CT를 찍지 못하는 경우가 있다. 따라서, 학습에 사용하기 위해 모든 CT 데이터에는 조영증강 CT와 조영증강이 되지 않은 CT가 섞여 있어 학습 성능의 저하를 일으키게 된다. Sandfort 등(30)은 이 문제를 해결하기 위해 CycleGAN을 이용하여 조영증강 CT를 조영증강이 되지 않은 CT로 변환하여 신장, 간, 비장 분할을 진행하였다. 각각의 분할 학습을 위해 총 네 가지 실험을 하였는데, 첫째로 데이터 수의 증강 없이 실제 데이터만을 사용하였고, 두 번째로 회전, 변형 등과 같은 영상처리 기반으로 증강한 데이터를 사용하였다. 세 번째로는 히스토그램 평활화(equalization)만을 적용한 데이터를 사용하였고, 마지막으로 CycleGAN으로 조영증강 CT를 조영증강하지 않은 CT로 변환한 영상을 사용하였다. 영상 분할은 U-Net (31)으로 학습하였으며, 학습결과 신장, 간, 비장의 분할 결과를 평균한 Dice 점수를 첫 번째부터 각각 0.916, 0.924, 0.935, 0.932로 얻었다. 이 결과만 보면 영상처리 기법에 따른 분할 성능의 차이는 크게 없음을 알 수 있다.

본 연구에서는 외부 병원에서 얻은 조영증강하지 않은 CT 영상에 대해서도 같은 결과를 얻는지 추가 실험을 진행하였다. 그 결과 각각 0.101, 0.535, 0.529, 0.747의 Dice 점수를 얻었다. Dice 점수가 낮아진 이유는 CycleGAN을 이용한 네 번째 실험을 제외하고 나머지 실험들이 조영증강 CT를 이용해 학습한 결과를 조영증강이 되지 않은 외부 데이터에 사용하였기 때문이다. 반면에 조영

증강 CT를 조영증강이 되지 않은 CT로 변환하여 학습한 네 번째 실험의 경우 외부 데이터에도 강인하게 좋은 성능을 내는 것을 확인할 수 있었다. 본 연구를 통해 GAN을 이용하여 다양한 데이터의 속성(본 연구에서는 조영증강 여부)을 단일 속성으로 표준화함으로써 성능을 높일 수 있음을 알 수 있다.

또한, GAN을 이용하여 영상에서 병변을 지우거나 만드는 방식으로 데이터를 증강한 연구도 있다. 악성 종양이 포함된 유방조영상과 정상 유방조영상의 데이터 수는 일반적으로 큰 차이가 있다. Wu 등(32)은 이러한 불균형을 해소하기 위해 conditional infilling GAN (이하 ciGAN)을 이용하여 악성이 아닌 유방조영상에 악성영역을 생성하고, 악성 종양의 유방조영상에서 악성 종양을 지운 뒤, 악성 영역과 그 외의 영역을 분류하는 성능을 평가하였다. 총 세 가지 데이터셋에 대해 실험을 진행하였는데, 데이터 증강을 하지 않은 원본 데이터, 영상처리 기법을 통해 증강한 데이터, 그리고 ciGAN과 영상처리 기법을 통해 증강한 데이터 각각에 대해 학습한 결과를 비교하였다. 그 결과 ciGAN과 영상처리 기법을 같이 사용하여 증강한 데이터로 학습한 결과에서 분류 성능이 가장 높게 나왔다(AUC = 0.896).

Onishi 등(33)은 Wasserstein GAN (이하 WGAN) (34)을 이용해 생성된 영상으로 분류기를 만들고, 실제 영상으로 분류기를 미세조정(fine tuning) 하는 연구를 진행하였다. 본 연구에서 악성 종양이 있는 CT 27건과 33건의 양성 종양이 있는 CT를 이용해 분류 성능을 평가하였다. 성능 평가는 전이학습을 하지 않고 원래 데이터로 폐결절의 악성 종양 여부를 분류하는 실험과, ImageNet (9)으로 학습한 모델을 전이학습(transfer learning) 한 실험과, 본 논문에서 제안하는 WGAN을 이용하여 생성한 영상으로 학습한 분류 모델을 전이학습한 실험을 비교하였다. 그 결과 마지막 실험의 분류 정확도가 앞선 두 실험들에 비해 15%가량 상승하였다. 이는 생성 영상을 학습한 후 전이학습을 하는 경우 분류 정확도를 높일 수 있음을 의미한다.

앞서 살펴본 바와 같이 GAN으로 데이터 증강을 하는 방법은 많이 사용되고 있지만, 실제 데이터에 비해 품질이 떨어지는 경우가 많다. 이러한 이유로 GAN을 이용하여 생성한 영상의 품질을 검증하는 연구도 활발히 진행되고 있다. Han 등(35)은 DCGAN과 WGAN을 이용해 128 × 128 사이즈의 뇌 MR 영상을 생성하여, 시각 튜링 테스트(visual Turing test)를 진행하였다. 한 명의 전문 의료인이 시각 튜링 테스트를 실시한 결과 DCGAN으로 생성한 영상의 경우 평균 65%의 정확도를 보였고, WGAN으로 생성한 영상의 경우 60%의 정확도를 보였다. 비록 한 명에 대한 결과이지만, GAN을 이용하여 생성한 영상이 실제 영상과 구별이 쉽지 않다는 것을 보여줬다.

Chuquicusma 등(36)은 DCGAN을 이용하여 CT 영상에서 폐결절이 있는 56 × 56 크기의 영상을 생성하였고, 두 명의 의료진을 통해 시각 튜링 테스트를 진행하였다. 영상에 대한 두 명의 의료진 답안 중 50%가 서로 불일치를 나타냈고, 이는 GAN을 통해 실제 영상과 구분하기 어려운 고품질의 폐결절 영상을 생성할 수 있음을 보여준다.

마지막으로 의료영상은 아니지만 학습 데이터의 수가 적을 경우, 많은 학습 데이터로 학습한 GAN의 모델을 적은 학습 데이터에 전이학습을 하여 고품질의 영상을 생성한 연구도 있어 소개한다. Zhao 등(37)은 대량의 데이터로 잘 학습된 GAN 판별기를 이용해 소량의 데이터(1000개 샘플 이하)를 가진 프로젝트에 전이학습으로 사용하였다. 본 연구 결과 소량의 데이터만으로도 GAN의

전이학습을 이용하면, 고품질의 영상을 생성할 수 있음을 알 수 있다. Fig. 3은 PGGAN을 이용하여 생성한 흉부 X-ray의 예시 그림을 보여주고 있다.

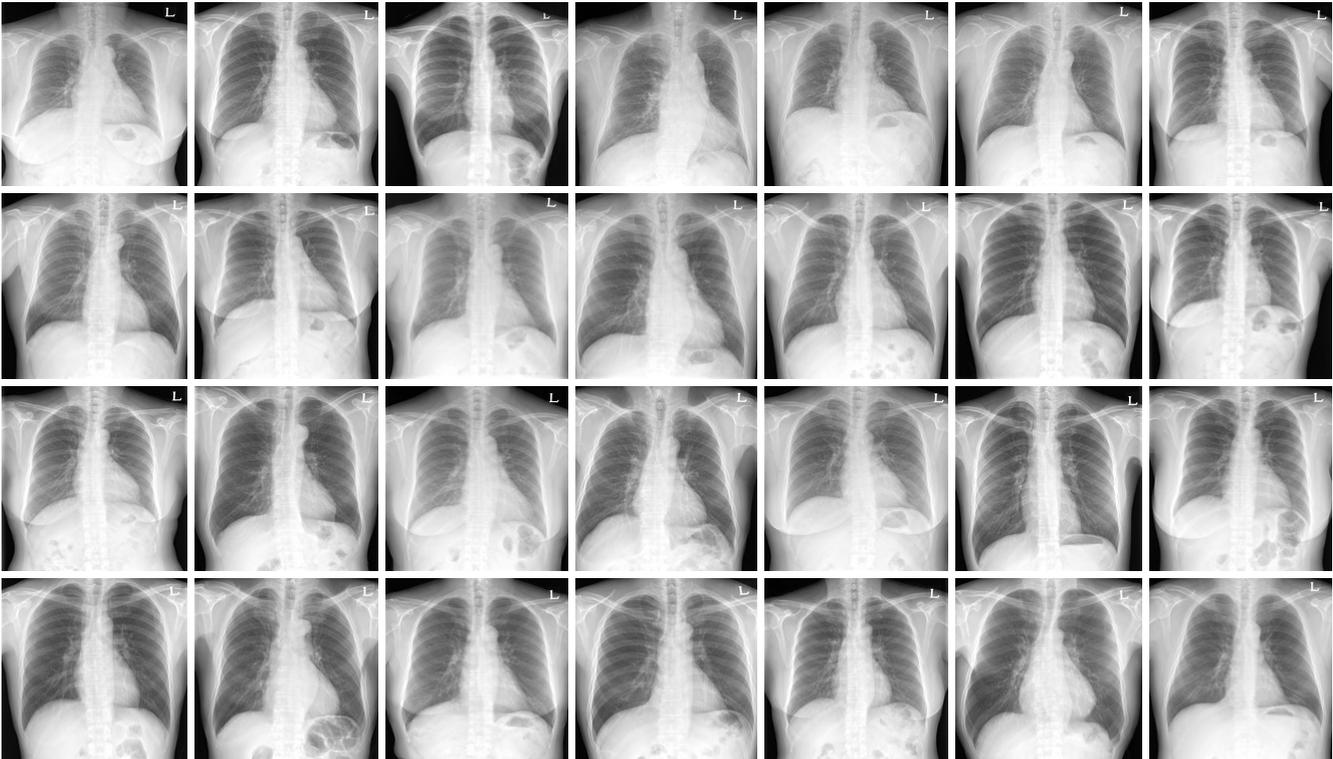
최신 데이터 증강 기법 동향

기존의 영상처리 기반 혹은 GAN 기반 데이터 증강 기법이 활발히 사용되고 있는 가운데, 새로운 데이터 증강 기법들 또한 꾸준히 소개되고 있다. 이 중에서도 주목을 받고 있는 몇 가지 대표적인 기법들을 소개하고자 한다.

먼저 Cutout은 학습 영상의 일부분을 무작위로 가리는 데이터 증강 기법이다(38). 해당 기법은 쉽게 구현할 수 있고 추가 계산이 많이 요구되지 않는다는 장점이 있지만, 데이터의 크기와 종류에 따라 최적의 마스크 크기를 찾아야 한다는 점이 단점이 있다. 예를 들어, 10개의 클래스를 갖는 Canadian Institute For Advanced Research (이하 CIFAR)-10 데이터셋(32×32)에 대해서는 16×16 크기의 마스크를 사용했을 경우 제일 높은 성능 향상을 보이며, 100개의 클래스를 갖는 CIFAR-100 데이터셋(32×32)에 대해서는 8×8 크기의 마스크를 사용한 경우 가장 높은 성능 향상을 보이고 있다.

다음으로 Mixup은 두 개의 서로 다른 클래스에 해당하는 데이터를 하나로 섞는 방식으로 데이터를 증강하는 기법이다(39). 해당 기법은 베타분포(beta distribution)에서 사용되는 하이퍼파라미터(hyperparameter)를 사용하여 두 영상 사이의 섞임 정도를 조절할 수 있고, 영상 데이터뿐만

Fig. 3. Examples of chest X-ray images generated using progressive growing of generative adversarial networks.



아니라 음성 데이터 등에도 적용할 수 있다는 점을 실험으로 보였다. Cutout과 달리 도메인에 의존하지 않는 장점이 있으며, 레이블 노이즈가 있는 경우에도 좋은 성능을 보인다는 장점이 있다. 하지만 Mixup 기법은 영상 분할에 있어서는 낮은 성능을 보인다는 단점이 있다(22).

이후에 나온 CutMix 기법은 영상의 일부를 랜덤하게 자른 뒤 해당 영역에 다른 클래스의 영상을 떼어 붙이는 방식을 통해 데이터를 증강하는 기법이다(40). Cutout에서는 영상을 잘라내는 것에 그쳤다면, CutMix는 잘라낸 부분에 다른 클래스의 영상을 붙인다는 점이 차이점이다. CutMix는 단순하면서도 다양한 딥러닝 문제에서 Cutout이나 Mixup보다 더 나은 성능을 보여준다는 장점이 있다. 다만 영상을 떼어 붙이는 과정에서 원본 영상이 갖는 의미론적 속성(semantic properties)을 잃을 수 있는 단점이 있다.

영상을 자르거나 섞거나 붙이는 증강 기법의 단점을 보완하며 제시되는 새로운 증강 기법이 AugMix 기법이다(41). 앞서 언급한 Mixup, CutMix 등의 기법들과 달리 원본 영상의 의미론적 속성을 보존하며 데이터를 증강한다는 장점이 있다. AugMix는 입력 영상에 대해 여러 번의 영상처리 기반 변환을 통한 증강 영상을 만들고, 이러한 증강 영상을 각기 다르게 여러 장 만들어 이 영상들을 픽셀 단위의 볼록 조합(convex combinations)을 통해 하나의 영상으로 합친다. 이렇게 만들어진 영상에 다시 입력 영상을 볼록 조합을 통해 합쳐주게 된다. AugMix는 앞서 언급한 기법들과 비교해 영상 분류 문제에 있어 가장 높은 성능 향상을 보여주고 있다(Fig. 4). 하지만 영상 분할이나 검출 문제에 대해서는 성능 평가가 이루어지지 않아 이 부분에 대한 추가 연구가 필요할 것으로 보인다. Fig. 5는 다양한 최신 데이터 증강 기법을 사용하여 변형한 영상의 예시를 보이고 있다.

딥러닝을 활용한 의료영상 연구에서도 앞서 소개한 최신 데이터 증강 기법 혹은 유사한 기법들이 시도되고 있어, 몇 가지 대표적인 연구들을 소개한다.

Bae 등(42)은 딥러닝을 통한 흉부 CT 영상에서 미만성간질성폐질환에 대한 분류 연구를 수행하였다. 해당 연구에서는 게임 등에서 주로 사용되는 펄린 노이즈(Perlin noise) (43, 44)를 활용한 데이터 증강 기법을 새롭게 적용하여 기존의 영상처리 기반 데이터 증강 기법과 분류 성능을 비교하였다. 펄린 노이즈는 무작위로 자연스러운 경계선을 만들 수 있으며, 이러한 경계선을 이용해

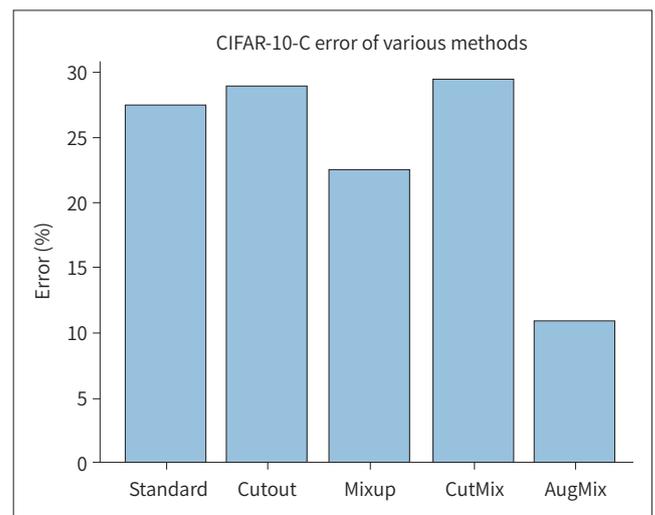
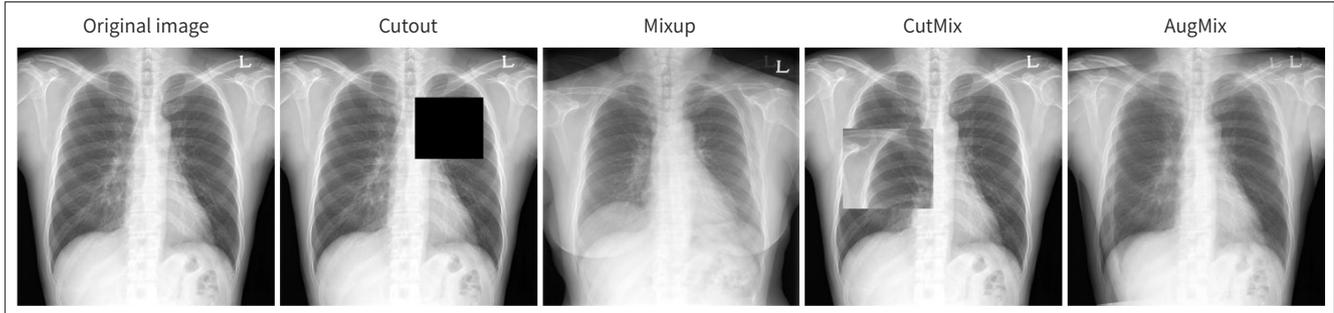


Fig. 4. Classification error rates of deep learning models trained with various data augmentation techniques on the CIFAR-10-C dataset. Adapted from Hendrycks et al. ArXiv Preprint 2019;arXiv:1912.02781 (41).
CIFAR = Canadian Institute For Advanced Research

Fig. 5. Examples of various chest X-ray augmented images using recent augmentation techniques.



학습에 쓰이는 6개 클래스 중 두 개의 클래스를 섞어 데이터 증강을 하는 방식이다. 해당 방식은 하나의 입력 영상 안에 두 개의 클래스를 무작위로 섞는다는 점에서 CutMix와 유사한 방식으로 볼 수 있다. 펄린 노이즈 기반 증강 기법을 사용했을 때 평균 분류 정확도는 89.5%로 영상처리 기반 증강 기법을 사용했을 때의 82.1%에 비해 통계적으로 유의미하게 높은 성능을 보여주는 것을 확인했다.

Noguchi 등(45)은 딥러닝을 통해 전신 CT 영상에서 뼈를 분할하는 연구를 수행하였다. 이 과정에서 기존의 영상처리 기반 기법, Mixup, 그리고 random image cropping and patching (이하 RICAP) (46)이라는 세 가지의 데이터 증강 기법을 각기, 그리고 혼합하여 분할 성능을 비교하였다. RICAP은 CutMix와 유사하지만 네 개의 각기 다른 클래스 영상을 하나로 만든다는 점에서 차이가 있다. 분할 모델의 성능을 각기 비교한 결과 Mixup을 사용한 경우보다 영상처리 기반 및 RICAP을 동시에 사용한 경우에서 가장 높은 성능 향상을 보였다. 앞서 언급한 것처럼 Mixup은 영상 분할에 있어 낮은 성능을 보인다는 점을 다시 한번 확인하였으며, 문제의 종류에 따라 적절한 데이터 증강 기법을 선택해야 한다는 점을 보여주는 연구이다.

토론

의료영상은 데이터에 대한 레이블링을 하는 과정에 많은 비용과 시간이 들기 때문에 딥러닝에 필요한 만큼 대량의 데이터를 얻기 힘들다. 또한, 클래스별 데이터 수가 불균형일 경우가 많다. 이렇게 데이터가 불충분하면 학습된 모델의 편향을 일으킨다. 따라서, 학습 성능은 우수하다 할지라도 테스트 데이터나 외부 데이터에 대해서는 낮은 성능을 보이게 된다(36). 데이터의 속성을 다양하게 늘리고, 클래스별 데이터 수의 균형을 맞추기 위해 데이터 증강을 해야 함을 본 종설에서 소개한 연구들을 통해 알 수 있다. 본 종설에서 소개한 데이터 증강 기법으로는 회전, 반전, 감마값보정 등의 영상처리 기반 기법과, GAN 학습을 통하여 새로운 영상을 생성해내는 기법, 그리고 영상의 부분을 잘라 붙이거나 다양한 영상처리 기법을 섞는 등의 최신 데이터 증강 기법들이다.

본 종설에서 소개한 연구들에 따르면 클래스별 영상 수의 불균형을 균형 있게 맞추어주고, 학습 데이터 수를 늘리면 딥러닝 학습 성능이 올라간다는 사실을 알 수 있다. 하지만, 데이터를 단순히 늘리는 방식으로 과대표집(oversampling)을 하게 되면, 모델이 학습 데이터의 특징을 쉽게 외우

게 되어, 과적합의 문제가 발생한다. 이러한 문제를 피하기 위해 영상처리 기법이나 GAN 등을 적절히 활용한 데이터 증강이 필요하고, 또한 문제의 성격과 보유하고 있는 데이터의 성격에 따라 최적의 증강 기법을 찾아야 한다.

특히, GAN 기반의 증강 기법은 영상처리 기반의 데이터 증강 기법에 비해 주의를 기울여 사용해야 할 것으로 보인다. 이는 GAN의 한계 때문인데, 첫째로 생성자와 판별자가 상호적으로 조금씩 각자의 성능을 키워야 하지만 한쪽의 능력이 크게 우세해지면 더이상 학습을 하지 못한다. 또한, 학습 데이터의 전체 분포를 배우지 못하고 쉬운 패턴만 익혀서 생성할 수 있다. 생성자는 판별자를 속여서 가짜 데이터를 진짜 데이터로 판별하게 하는데, 진짜처럼 만들 수 있는 쉬운 영상만 만들어서 판별자에게 알려주게 되고, 판별자가 생성자를 수정하게 하기 위한 피드백을 주지 못한다. 또한, 앞서 소개한 CycleGAN과 같은 영상 변환의 경우에는 실제로 존재하지 않는 영상을 만들어 내기도 한다.

이와 같은 문제를 해결하기 위해 GAN으로 생성한 영상이 실제와 얼마나 유사한지 알아보기 위한 연구가 다양하게 진행되고 있다. 예를 들어, GAN으로 생성한 영상의 품질을 평가하기 위해 Mean Absolute Error, Peak Signal-to-Noise Ratio, Structural SIMilarity, 그리고 Inception Score 등 수학적 정의를 이용한 평가 지표가 개발되거나 시각 튜링 테스트 등이 실시되기도 한다. 하지만 유사성을 정량화하는 것에는 아직도 한계가 있다. 특히, 모드 붕괴(mode collapse)와 같은 이유로 GAN이 생성한 영상의 다양성을 정량적으로 측정할 수 있는 후속 연구가 필요하다.

결론

본 종설에서 데이터 증강 기법의 동향을 소개하였다. 실제 우리가 접하는 대부분의 문제에서는 데이터의 부족이나, 클래스별 데이터의 불균형으로 인해 과대적합 및 과소적합이 나타날 수 있으며, 이를 해결하기 위한 데이터 증강 기법은 필수적이라 할 수 있겠다. 본 종설에서 소개한 연구들을 살펴보면, 영상처리 기반의 데이터 증강 기법을 주로 사용하던 연구에서 점차로 GAN을 활용한 데이터 증강 기법으로 연구 동향이 옮겨가고 있음을 확인할 수 있다. 이는 GAN을 활용하는 경우 학습 영상이 가질 수 있는 다양한 특징들을 만들어내어 모델을 더욱 일반화할 수 있기 때문이다. 또한, 흥미로운 사실은 기존의 영상처리 기반의 데이터 증강 기법과 GAN을 활용한 데이터 증강 기법을 섞게 되면, 데이터 분포의 다양성이 더욱 증대되어 모델의 성능을 높일 수 있다는 것이다. 최근에는 CutMix, AugMix 등 상대적으로 간단하게 영상을 합성하여 데이터를 증강하는 기법들이 소개되고 있으며, 기존의 데이터 증강 기법과의 비교도 시도되고 있다. 결론적으로 이러한 다양한 데이터 증강 기법 중 학습 목적과 모델 종류, 그리고 학습을 위해 보유한 데이터의 특성에 따라 적절한 선택을 해야 하며, 데이터 증강 기법에 따른 딥러닝 모델의 성능을 비교하는 정량적 연구가 필요할 것으로 보인다.

Author Contributions

Conceptualization, all authors; data curation, K.M.; formal analysis, all authors; investigation, all authors; project administration, B.H.; resources, all authors; visualization, K.M.; writing—original

draft, all authors; and writing—review & editing, all authors.

Conflicts of Interest

The authors have no potential conflicts of interest to disclose.

REFERENCES

- Kim M, Yun J, Cho Y, Shin K, Jang R, Bae HJ, et al. Deep learning in medical imaging. *Neurospine* 2019; 16:657-668
- Kim JH. Imaging informatics: a new horizon for radiology in the era of artificial intelligence, big data, and data science. *J Korean Soc Radiol* 2019;80:176-201
- Song KD, Kim M, Do S. The latest trends in the use of deep learning in radiology illustrated through the stages of deep learning algorithm development. *J Korean Soc Radiol* 2019;80:202-212
- Srivastava N, Hinton G, Krizhevsky A, Sutskever I, Salakhutdinov R. Dropout: a simple way to prevent neural networks from overfitting. *J Mach Learn Res* 2014;15:1929-1958
- Yamashita R, Nishio M, Do RKG, Togashi K. Convolutional neural networks: an overview and application in radiology. *Insights Imaging* 2018;9:611-629
- Do S, Song KD, Chung JW. Basics of deep learning: a radiologist's guide to understanding published radiology articles on deep learning. *Korean J Radiol* 2020;21:33-41
- Jia X, Luo T, Ren S, Guo K, Li F. Small sample-based disease diagnosis model acquisition in medical human-centered computing. *J Wireless Com Network* 2019;2019:212
- Han C, Rundo L, Araki R, Furukawa Y, Mauri G, Nakayama H, et al. *Infinite brain MR images: PGGAN-based data augmentation for tumor detection*. In Esposito A, Faundez-Zanuy M, Morabito F, Pasero E, eds. *Neural approaches to dynamics of signal exchanges*. Singapore: Springer 2020:291-303
- Krizhevsky A, Sutskever I, Hinton GE. Imagenet classification with deep convolutional neural networks. *Adv Neural Inf Process Syst* 2012;25:1097-1105
- Roth HR, Lu L, Liu J, Yao J, Seff A, Cherry K, et al. Improving computer-aided detection using convolutional neural networks and random view aggregation. *IEEE Trans Med Imaging* 2015;35:1170-1181
- Goodfellow I, Pouget-Abadie J, Mirza M, Xu B, Warde-Farley D, Ozair S, et al. Generative adversarial nets. *Adv Neural Inf Process Syst* 2014:2672-2680
- Roth HR, Lee CT, Shin HC, Seff A, Kim L, Yao J, et al. *Anatomy-specific classification of medical images using deep convolutional nets*. New York: 2015 IEEE 12th International Symposium on Biomedical Imaging (ISBI) 2015:101-104
- Hao R, Namdar K, Liu L, Haider MA, Khalvati F. A comprehensive study of data augmentation strategies for prostate cancer detection in diffusion-weighted MRI using convolutional neural networks. *ArXiv Preprint* 2020;arXiv:2006.01693
- Zhao Z, Zhang Z, Chen T, Singh S, Zhang H. Image augmentations for GAN training. *ArXiv Preprint* 2020;arXiv:2006.02595
- Tang C, Li J, Wang L, Li Z, Jiang L, Cai A, et al. Unpaired low-dose CT denoising network based on cycle-consistent generative adversarial network with prior image information. *Comput Math Methods Med* 2019; 2019:8639825
- Yang Q, Yan P, Zhang Y, Yu H, Shi Y, Mou X, et al. Low-dose CT image denoising using a generative adversarial network with Wasserstein distance and perceptual loss. *IEEE Trans Med Imaging* 2018;37:1348-1357
- Dong Z, Liu G, Ni G, Jerwick J, Duan L, Zhou C. Optical coherence tomography image denoising using a generative adversarial network with speckle modulation. *J Biophotonics* 2020;13:e201960135
- Frid-Adar M, Diamant I, Klang E, Amitai M, Goldberger J, Greenspan H. GAN-based synthetic medical image augmentation for increased CNN performance in liver lesion classification. *Neurocomputing* 2018;321:321-331
- Radford A, Metz L, Chintala S. Unsupervised representation learning with deep convolutional generative adversarial networks. *ArXiv Preprint* 2015;arXiv:1511.06434
- Zhao D, Zhu D, Lu J, Luo Y, Zhang G. Synthetic medical images using F&BGAN for improved lung nodules classification by multi-scale VGG16. *Symmetry* 2018;10:519
- Salehinejad H, Valaee S, Dowdell T, Colak E, Barfett J. *Generalization of deep neural networks for chest pa-*

- thology classification in X-rays using generative adversarial networks*. Calgary: 2018 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP) 2018:990-994
22. He T, Zhang Z, Zhang H, Zhang Z, Xie J, Li M. *Bag of tricks for image classification with convolutional neural networks*. Long Beach: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition 2019:558-567
 23. Karras T, Aila T, Laine S, Lehtinen J. Progressive growing of gans for improved quality, stability, and variation. *ArXiv Preprint* 2017;arXiv:1710.10196
 24. Redmon J, Farhadi A. Yolov3: an incremental improvement. *ArXiv Preprint* 2018;arXiv:1804.02767
 25. Zhao J, Li D, Kassam Z, Howey J, Chong J, Chen B, et al. Tripartite-GAN: synthesizing liver contrast-enhanced MRI to improve tumor detection. *Med Image Anal* 2020;63:101667
 26. Bowles C, Chen L, Guerrero R, Bentley P, Gunn R, Hammers A, et al. Gan augmentation: augmenting training data using generative adversarial networks. *ArXiv Preprint* 2018;arXiv:1810.10863
 27. Russ T, Goerttler S, Schnurr AK, Bauer DF, Hatamikia S, Schad LR, et al. Synthesis of CT images from digital body phantoms using CycleGAN. *Int J Comput Assist Radiol Surg* 2019;14:1741-1750
 28. Zhu JY, Park T, Isola P, Efros AA. *Unpaired image-to-image translation using cycle-consistent adversarial networks*. Venice: Proceedings of the IEEE Conference on Computer Vision 2017:2223-2232
 29. Gupta A, Venkatesh S, Chopra S, Ledig C. Generative image translation for data augmentation of bone lesion pathology. *ArXiv Preprint* 2019;arXiv:1902.02248
 30. Sandfort V, Yan K, Pickhardt PJ, Summers RM. Data augmentation using generative adversarial networks (CycleGAN) to improve generalizability in CT segmentation tasks. *Sci Rep* 2019;9:16884
 31. Ronneberger O, Fischer P, Brox T. *U-net: convolutional networks for biomedical image segmentation*. In Navab N, Hornegger J, Wells W, Frangi A, eds. *International conference on medical image computing and computer-assisted intervention*. Cham: Springer 2015:234-241
 32. Wu E, Wu K, Cox D, Lotter W. *Conditional infilling GANs for data augmentation in mammogram classification*. In Stoyanov D, Taylor Z, Kainz B, Maicas G, Beichel RR, Martel A, et al. *Image analysis for moving organ, breast, and thoracic images*. Cham: Springer 2018:98-106
 33. Onishi Y, Teramoto A, Tsujimoto M, Tsukamoto T, Saito K, Toyama H, et al. Automated pulmonary nodule classification in computed tomography images using a deep convolutional neural network trained by generative adversarial networks. *Biomed Res Int* 2019;2019:6051939
 34. Arjovsky M, Chintala S, Bottou L. Wasserstein gan. *ArXiv Preprint* 2017;arXiv:1701.07875
 35. Han C, Hayashi H, Rundo L, Araki R, Shimoda W, Muramatsu S, et al. *GAN-based synthetic brain MR image generation*. Washington D.C.: 2018 IEEE 15th International Symposium on Biomedical Imaging (ISBI 2018) 2018:734-738
 36. Chuquicuma MJ, Hussein S, Burt J, Bagci U. *How to fool radiologists with generative adversarial networks? A visual turing test for lung cancer diagnosis*. Washington D.C.: 2018 IEEE 15th international symposium on biomedical imaging (ISBI 2018) 2018:240-244
 37. Zhao M, Cong Y, Carin L. On leveraging pretrained GANs for generation with limited data. *ArXiv Preprint* 2020;arXiv:2002.11810
 38. DeVries T, Taylor GW. Improved regularization of convolutional neural networks with cutout. *ArXiv Preprint* 2017;arXiv:1708.04552
 39. Zhang H, Cisse M, Dauphin YN, Lopez-Paz D. Mixup: beyond empirical risk minimization. *ArXiv Preprint* 2017;arXiv:1710.09412
 40. Yun S, Han D, Oh SJ, Chun S, Choe J, Yoo Y. *Cutmix: regularization strategy to train strong classifiers with localizable features*. Seoul: Proceedings of the IEEE International Conference on Computer Vision 2019:6023-6032
 41. Hendrycks D, Mu N, Cubuk ED, Zoph B, Gilmer J, Lakshminarayanan B. Augmix: a simple data processing method to improve robustness and uncertainty. *ArXiv Preprint* 2019;arXiv:1912.02781
 42. Bae HJ, Kim CW, Kim N, Park B, Kim N, Seo JB, et al. A Perlin noise-based augmentation strategy for deep learning with small data samples of HRCT images. *Sci Rep* 2018;8:17687
 43. Perlin K. An image synthesizer. *ACM Siggraph Computer Graphics* 1985;19:287-296
 44. Perlin K. *Improving noise*. *ACM transactions on graphics*. San Antonio: Proceedings of ACM SIGGRAPH 2002; 21:681-682
 45. Noguchi S, Nishio M, Yakami M, Nakagomi K, Togashi K. Bone segmentation on whole-body CT using con-

volutional neural network with novel data augmentation techniques. *Comput Biol Med* 2020;121:103767

46. Takahashi R, Matsubara T, Uehara K. *RICAP: random image cropping and patching data augmentation for deep CNNs*. Beijing: Asian Conference on Machine Learning 2018:786-798

딥러닝 기반 의료영상 분석을 위한 데이터 증강 기법

김민규¹ · 배현진^{2*}

영상처리 기반으로 의료영상을 분석하는 기법은 정상 환자와 비정상 환자를 분류, 병변 검출 및 장기나 병변의 분할 등에 사용되고 있다. 최근 인공지능 기술의 비약적 발전으로 의료영상 분석 연구들이 딥러닝 기술을 활용하여 시도되고 있다. 의료영상은 학습에 필요한 데이터를 충분히 모으기 어렵고 클래스별 데이터 수의 차이 때문에, 딥러닝 모델의 성능을 올리는 데 어려움이 있다. 이러한 문제를 해결하기 위해 다양한 연구가 시도되고 있으며, 이 중 하나가 학습 데이터를 증강하는 것이다. 본 중설에서는 회전, 역상, 밝기 변화 등과 같은 영상처리 기반의 데이터 증강, 적대적생성네트워크를 활용한 데이터 증강, 그리고 기존 영상의 속성들을 섞는 등의 최신 데이터 증강 기법을 알아보고, 의료영상 연구에 적용된 사례들과 그 결과를 조사해 보고자 한다. 끝으로 데이터 증강의 필요성을 고찰하고 앞으로의 방향을 짚어본다.

¹울산대학교 의과대학 융합의학과,

²프로메디우스 주식회사