

GAN 기반 고해상도 의료 영상 생성을 위한 연구

고재영*, 조백환**, 정명진**

*성균관대학교 삼성융합의과학원 디지털헬스학과

**삼성서울병원 스마트헬스케어연구소 AI 연구센터

kojae8311@g.skku.edu, baekhwan.cho@samsung.com, mj1.chung@samsung.com

GAN-based research for high-resolution medical image generation

Jae-Yeong Ko*, Baek-Hwan Cho**, Myung-Jin Chung**

*Dept. of Digital Health, SAIHST, Sungkyunkwan University

**Medical AI Research Center, Smart Healthcare Research Institute, Samsung Medical Center

요 약

의료 데이터를 이용하여 인공지능 기계학습 연구를 수행할 때 자주 마주하는 문제는 데이터 불균형, 데이터 부족 등이며 특히 정제된 충분한 데이터를 구하기 힘들다는 것이 큰 문제이다. 본 연구에서는 이를 해결하기 위해 GAN(Generative Adversarial Network) 기반 고해상도 의료 영상을 생성하는 프레임워크를 개발하고자 한다. 각 해상도 마다 Scale 의 Gradient 를 동시에 학습하여 빠르게 고해상도 이미지를 생성해낼 수 있도록 했다. 고해상도 이미지를 생성하는 Neural Network 를 고안하였으며, PGGAN, Style-GAN 과의 성능 비교를 통해 제안된 모델이 양질의 고해상도 의료영상 이미지를 더 빠르게 생성할 수 있음을 확인하였다. 이를 통해 인공지능 기계학습 연구에 있어서 의료 영상의 데이터 부족, 데이터 불균형 문제를 해결할 수 있는 Data augmentation 이나, Anomaly detection 등의 연구에 적용할 수 있다.

1. 서론

인공지능 기계학습 연구에서는 품질 좋은 데이터량에 따라 연구의 성과가 달라질 수 있기 때문에 데이터량은 중요하다. 특히 의료 분야는 privacy issue, annotation 에 필요한 노동력 등 다양한 문제로 인해 정제된 양질의 데이터를 대량으로 획득하기가 어렵다.

흉부 방사선 영상 (Chest X-ray) 촬영은 병원에서 가장 대중적으로 이루어지고 있는 검사 중 하나로 의료 영상 데이터중에는 상대적으로 데이터의 개수가 많지만 각종 질환이나 병변 위치의 annotation 을 위해서는 여전히 많은 자원을 필요로 한다. 뿐만 아니라 건강 검진 등을 위해서 Chest X-ray 촬영을 많이 하기 때문에 정상(normal) 데이터가 비정상(abnormal) 데이터에 비해 데이터량이 많아 데이터 불균형 현상이 발생한다. 이러한 문제들을 해결하고자 생성 모델(Generative model)들을 이용한 연구들이 진행되고있다.

Generative Model 중 GAN 을 이용한 연구가 많은 분야에서 활발하게 이루어 지고 있다. 그 중에서 Convolutional layer 를 처음 GAN 에 적용시킨 DCGAN[5]은 가장 기본이 되는 GAN 모델이다. DCGAN 은 저해상도 영상을 빠르게 생성해주는 장점이 있지만 고해상도 영상의 경우 네트워크 구조에서

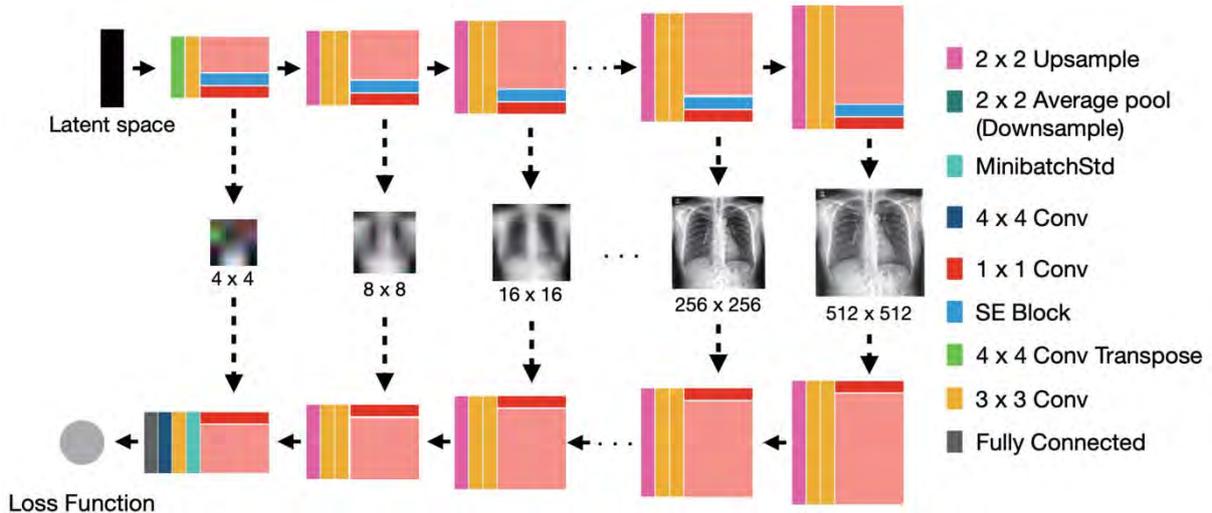
기인하는 GPU 메모리의 한계로 인해 batch 크기에 대한 제약이 있어 학습 성능이 좋지 않다.

고해상도 영상을 생성하는 대표적인 모델인 PGGAN[4]은 Discriminator 와 Generator 를 점진적으로 layer 를 쌓아 영상의 해상도를 증가시키면서 학습을 진행하는 특징을 갖고 있다. 하지만 이와 같은 방법으로 학습을 진행하는 것은 학습에 매우 많은 시간이 걸린다는 단점이 있으며, 충분한 학습 데이터량이 필요하다. 그러나 의료 데이터 특성상 다른 분야의 데이터에 비해 데이터량이 상대적으로 부족하므로 데이터량이 적은 상태로 학습을 진행할 시 학습이 불안정하게 진행이 될 우려가 크다. 따라서 우리는 본 연구에서 의료 영상 데이터의 근본적 문제인 데이터 불균형 문제와 데이터량 부족 문제를 해결하기 위해 GAN 기반 고해상도 의료 영상 이미지를 생성하는 연구를 진행했다.

2. 방법

2-1. 데이터

삼성서울병원에서 촬영한 흉부 방사선 영상은 40,301 명의 환자에서 80,675 개의 데이터로 구성된 DICOM 형식의 익명화된 흉부 방사선 데이터 셋이다. 사용된 데이터는 측면 방사선 이미지(Lateral Chest radiograph)



(그림 1) Self Attention PGGAN

와 전면 방사선 영상(Posteroanterior Chest radiograph)으로 구성 되어 있다. 본 연구에서는 전면 방사선 사진만 필요하므로 전면 방사선 사진인 PA Chest radiograph 로만 연구를 진행했다. 최종적으로 본 연구에서는 40,940 개의 데이터로 연구를 진행했다.

2-2. 전처리

방사선 전문가들은 촬영된 방사선 영상을 판독할 때 Histogram 을 조절해서 판독하기 적절한 상태로 변경하여 판독을 진행한다. 본 연구에 사용된 영상도 동일한 효과를 적용하기 위해 Contrast Limited Adaptive Histogram Equalization 을 이용하여 방사선 영상의 전처리를 진행했다. 또한 원본 방사선 영상 크기의 5% 만큼 Random Crop 을 적용한 후 512x512 크기로 resize 를 해주었다.. 또한 PGGAN 의 구조를 그대로 사용하기 때문에 Feature vector 를 RGB Color 로 만들어 주는 layer 를 적용하기 위해 Gray Scale 영상을 duplication 을 통해 3 채널 영상인 RGB 로 변경하였으며, 메모리를 효율적으로 사용하기 위해 DICOM 이미지를 Pickle 형태로 변환하였다..

2-3. 네트워크

본 연구에서는 SAGAN[9]에서 Self Attention 을 적용한 것과 마찬가지로, PGGAN 모델 구조를 응용해서 각 해상도를 담당하는 중간 계층에 SE Block[2]을 적용시켰다. 그 후 1x1 convolution layer 를 통해 각 해상도 이미지의 Feature 를 추출해서 최종적으로 (그림 1)와 같이 추출된 Feature 들이 Discriminator 의 입력으로 들어가는 구조로 학습을 진행했다.

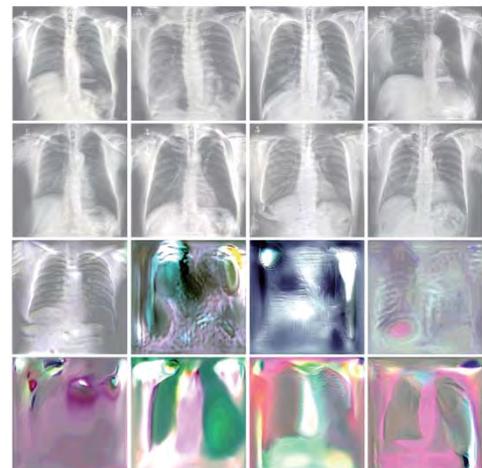
본 연구에서는 4 가지 손실 함수(Minmax Loss, WGAN-GP Loss[1], Hinge Loss, RA-Hinge Loss)를 적용하였고, 전면 흉부 방사선 영상(PA Chest radiograph) 데이터셋에는 RA-Hinge Loss[3] 가 더 빠르고 안정적으로 학습이 진행이 되는 것을 확인하였다.



(그림 2) PGGAN, Style-GAN 의 CelebA-HQ 학습 결과

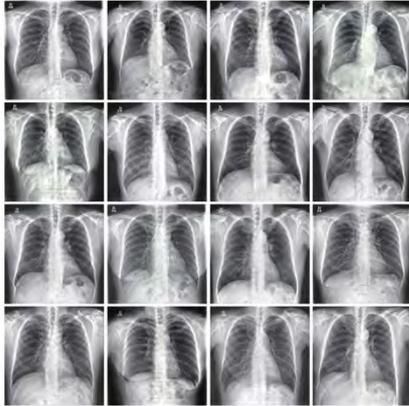
3. 연구 결과

PGGAN[4]과 Style-GAN 모델을 사용해서 CelebA-HQ 데이터로 학습을 진행한 결과는 (그림 2)와 같이 보다 안정적으로 학습이 진행되었지만 본 연구에서 사용한 PA Chest radiograph 으로 학습을 진한 결과는 (그림 3)와 같이 학습이 불안정하게 학습이 진행되어 학습 영상과 전혀 다른 형태의 영상을 생성하였다.



(그림 3) PGGAN 모델에 Chest x-ray 학습 결과 불안정하게 학습 된 GAN 의 512x512 생성 이미지.

본 연구에는 제안한 모델을 사용한 결과는 (그림 4)와 같이 안정적으로 학습이 진행된 것을 확인할 수 있었다. GAN 모델의 성능을 정량적으로 확인하기 위해 <표 1>과 같이 Epoch 마다 FID(Fréchet Inception distance)[8] Score 를 측정하였다. 기존 GAN Model 들에 비해 제안한 Self Attention PGGAN 이 FID Score 가 가장 낮았으며 특히 상대적으로 다른 모델들에 비해 Epoch 이 낮아도 안정적이고 빠르게 데이터 생성을 하고 있다는 것을 확인할 수 있었다.



(그림 4) Self Attention PGGAN 과 Relativistic Hinge GAN Loss 를 이용한 Chest X-ray 생성 결과.

<표 1> Chest x-ray 를 이용한 GAN Model 의 Epoch 에 따른 Fréchet Inception distance (FID)

Methods	FID				
	Epoch	10	50	100	150
DCGAN[5]		9045.36	8045.15	8195.39	7853.76
PGGAN[4]		305.43	105.94	208.36	207.46
Style-GAN[7]		206.53	90.63	109.34	305.28
Self Attention PGGAN		35.36	28.57	20.34	18.09

4. 결론

본 논문에서는 PA Chest radiograph 데이터를 이용하여 고해상도 이미지를 만들도록 연구를 진행했다. 제안한 모델은 PGGAN 구조에 Self Attention Module 을 적용하여, 기존의 모델에 비해 안정적으로 (그림 5)와 같이 고해상도 의료 영상을 생성하는 것을 확인하였다.



(그림 5) 생성된 512 x 512 흉부 방사선 사진.

본 연구에서 사용된 모델은 Data Augmentation, Anomaly Detection 등 의료 영상을 이용한 다양한 인공지능 기계학습 연구에 적용 가능하다.

"본 연구는 과학기술정보통신부 및 정보통신기획평가원의 대학 ICT 연구센터지원사업의 연구결과로 수행되었음" (IITP-2020-2018-0-01798)

참고문헌

[1] Gulrajani, I., et al. Improved training of wasserstein gans. Advances in neural information processing systems. (2017).
 [2] Hu, J., et al. Squeeze-and-excitation networks. Proceedings of the IEEE conference on computer vision and pattern recognition. (2018).
 [3] Jolicoeur-Martineau, A. "The relativistic discriminator: a key element missing from standard GAN." (2018).
 [4] Karras, T., et al. "Progressive growing of gans for improved quality, stability, and variation." (2017).
 [5] Radford, A., et al. "Unsupervised representation learning with deep convolutional generative adversarial networks." (2015).
 [6] Zhang, H., et al. "Self-attention generative adversarial networks." (2018).
 [7] Karras, T., et al. A style-based generator architecture for generative adversarial networks. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. (2019).
 [8] Heusel, M., et al. Gans trained by a two time-scale update rule converge to a local nash equilibrium. Advances in neural information processing systems. (2017).