

# GAN의 Residual Block 사용에 관한 구조적 비교 연구

권주원, 유지상, 권순철\*

광운대학교, \*광운대학교

02kjjw0203@gmail.com, jsyoo@kw.ac.kr, \*ksc0226@gmail.com

## A Study on the Structural Comparison of Residual Block in GANs

Kweon Ju-won, Yoo Ji-Sang, Kwon Soonchul\*

Kwangwoon Univ., \*Kwangwoon Univ.

### 요약

본 논문은 딥러닝 기반 생성 모델인 GAN에서 Residual 구조를 사용했을 때와 사용하지 않았을 때 Generator의 성능을 측정하여 비교한다. GAN에서 Residual 구조를 이용하면 약간의 네트워크의 파라미터 수가 증가하는 단점이 존재하지만 Residual Block을 사용하지 않았을 경우에 비해 FID 값이 작아지는 이점이 있다. FID는 두 분포간의 거리를 측정하는 지표로 작을수록 두 분포가 가깝다는 것을 의미하며 값이 작을수록 모델이 좋은 성능을 갖는다. 실험 결과로 Residual Block을 사용한 경우 크기는 38.3197부터 적게는 1.39까지 FID 값이 작아짐을 보였고 이는 네트워크의 성능이 증가함을 의미한다. 본 실험을 통해 작은 크기의 데이터셋에서는 GAN 구조에 Residual Module을 이용하면 성능이 높아짐을 알 수 있다.

### I. 서론

최근 딥러닝 분야의 급속한 발전은 다양한 분야에서 좋은 영향을 미치고 있습니다. 고전적인 문제들인 분류, 객체 탐지, 객체 추적 등의 문제 외에도 새로운 데이터를 생성하는 영역에 대한 관심도 비약적으로 증가하였다. 뉴럴 네트워크를 이용해 데이터셋의 분포를 학습하여 가상의 데이터를 생성하는 GAN[1]은 2014년 논문이 발표된 이후 최고의 딥러닝 아이디어로 평가 받았다.

현재에 이르러서 GAN은 구조적 측면, 학습의 안정을 가져오는 측면 등에 대해 다양한 논문들이 발표되었으며 대표적인 논문으로는 초창기 DCGAN[2], CGAN[3]부터 BigGAN[4], StyleGAN[5], WGAN[6], WGAN-GP[7] 까지 여러 발전된 연구들이 있다. 이러한 많은 연구들은 단순히 가상 데이터를 생성하는데 끝나지 않고 데이터 획득이 어렵거나 부족한 문제들을 해결하고 있다. 대표적으로는 CT, MRI, 병리 데이터 등의 의료 데이터들이 있다.

본 논문에서는 딥러닝 기반의 생성 모델인 GAN의 구조적 측면에서 Residual Block을 사용하는 것이 효과가 있는지 알아보는 연구를 진행하였다. Residual 구조를 처음으로 사용해 큰 주목을 받았던 ResNet은 gradient vanishing 문제를 해결하고 더 깊은 네트워크를 학습시킬 수 있음을 보였다. 마찬가지로 GAN에서도 Residual Block을 사용한 구조를 사용하는 모델들이 있다. 하지만 Residual Block을 사용하지 않은 연구들의 수가 많았기에 모든 Convolution Block을 Residual Block으로 변경하였을 때 효과를 확인해보기 위해서 실험을 하였다.

실험은 DCGAN의 구조를 이용해 모든 네트워크를 학습하였으며 데이터셋은 28x28의 해상도를 갖는 손글씨 숫자 데이터셋인 MNIST 데이터셋과 CIFAR-10 데이터셋, 고양이 얼굴 데이터셋 등 다양한 해상도에서 실험을 진행하였다. 모든 실험에서 성능을 평가하기 위해 사용된 평가 지표는 FID를 사용했다. 결과로 Residual Block을 사용하였을 경우 미약하게 Generator의 성능이 조금 올라가는 것을 알 수 있었다.

### II. 본론

본 논문에서는 딥러닝 기반의 생성모델인 GAN에서 Residual 구조를 사용하였을 때 어느정도의 성능향상과 이점이 있는지를 확인하기 위해서 MNIST, CIFAR-10, CATS 등 다양한 데이터셋에 대해서 성능을 측정하고 비교하였다. MNIST 데이터셋은 0~9까지의 손글씨 숫자 데이터셋으로 28x28의 크기를 갖으며 training data는 60,000장의 흑백 데이터로 이루어져 있다. CIFAR-10은 10개의 클래스를 갖는 32x32 크기의 60,000장의 RGB 데이터로 구성되어 있다. CATS 데이터셋은 고양이 얼굴 데이터셋으로 64x64의 크기를 갖으며 15,747장의 RGB 데이터로 구성되어 있다. 그림.1은 각 데이터셋의 원본 데이터, Residual Block을 사용하지 않은 생성 결과, Residual Block을 사용한 생성 결과는 각각 4개씩 보여준다. 위부터 MNIST, CIFAR-10, CATS 데이터셋이다.

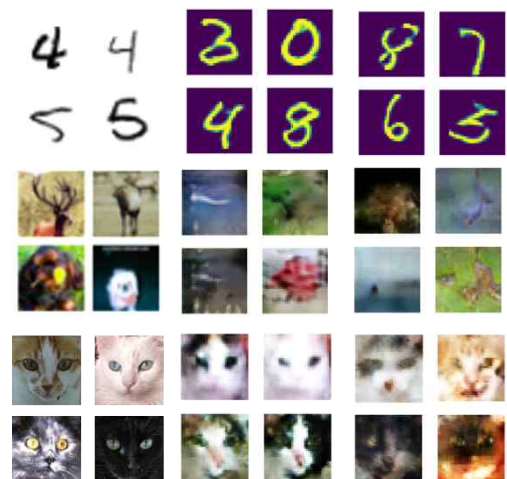


그림.1 데이터셋 예시와 생성된 결과 이미지

실험에 사용한 네트워크의 구조는 기본적으로 DCGAN의 구조를 이용하였다. DCGAN은 기존 GAN의 학습이 불안정한 것을 해소할 수 있도록

구조적인 가이드라인을 제시한 연구이다. BatchNormalization의 사용, ReLU와 LeakyReLU를 Generator와 Discriminator에 각각 activation으로 사용, Generator의 output layer를 Tanh를 사용하는 등의 조건들을 그대로 따랐다.

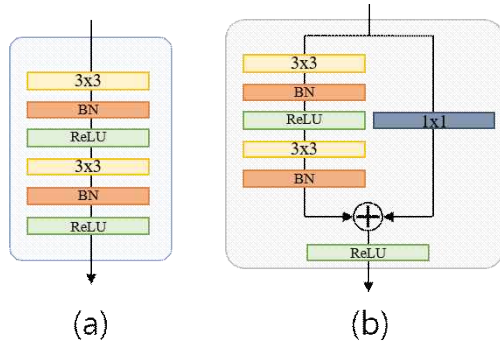


그림.2 (a) Convolution Block과 (b) Residual Block 내부의 구조

다만 차이를 둔 점은 Convolution Block을 Residual 구조를 사용하는 Block으로 대체한 것이다. 그림.2 는 우리가 사용한 Convolution Block과 Residual Block의 구조를 보여준다. Convolution Block은 3x3 Convolutional Layer-BatchNorm-ReLU를 2번 반복하며, Residual Block은 그림.2 (b)처럼 skip connection을 통해 이어지는 1x1 Convolution Layer와 두 번째 BatchNorm를 더해준 후 ReLU를 통과하는 구조를 갖는다. 모든 Convolution Block을 Residual Block으로 변경할 때 네트워크 파라미터 수가 약간 증가한다. 하지만 이 비율은 Residual Block을 사용하지 않은 모델에 비해 약 4% 정도의 증가로 큰 차이가 있지 않다.

데이터셋	데이터 크기	FID	
		w/o Residual	w/ Residual
MNIST	28x28	10.9461	<b>9.5561</b>
CIFAR-10	32x32	75.3115	<b>36.9918</b>
CATS	64x64	148.9732	<b>140.8342</b>

표.1 데이터셋 별 결과 FID 값

우리는 두 모델의 성능을 측정하기 위해서 FID (Fréchet Inception Distance)를 이용하였다. FID는 최근 많은 GAN 논문들에서 성능을 측정하기 위해서 사용 중인 성능 지표로 실제 데이터 분포와 생성된 데이터의 분포의 거리를 계산하여 품질을 평가합니다. InceptionV3 network를 사용하여 중간 layer에서 추출된 feature의 평균과 분산을 통해 두 분포간의 거리를 계산하며 낮을수록 높은 성능을 의미한다. 표.1는 데이터셋에 따른 FID 측정 결과를 보여준다. MNIST의 경우 Residual 구조를 사용하였을 때 약 9.5561로 사용하지 않은 경우보다 FID가 1.39 정도 낮아졌다. CIFAR-10은 약 38.3197 정도로 가장 크게 FID 값이 낮아졌으며 CATS 데이터셋 역시 약 8.139 정도로 FID가 작아지는 것을 알 수 있다.

본 논문에서 실험한 방법은 Intel® Core™ i9-10900k CPU와 RTX 3090 GPU에서 실험되었다.

### III. 결론

본 논문에서는 딥러닝 기반의 생성 모델 GAN의 가장 간단하고 효과적인 구조인 DCGAN을 바탕으로 모든 Convolution Block을 Residual Block으로 변경하였을 때 네트워크의 성능이 좋아지는지 실험하였다. 결과적으로 GAN의 성능을 측정하는 지표인 FID는 작은 값일수록 원데이터 분포와 비슷함을 의미하며 성능이 좋음을 의미한다. 실험결과로 3가지

데이터셋 MNIST, CIFAR-10, CATS 모두 FID 값을 작게는 1.39부터 크게는 38.3197까지 낮출 수 있었다. 하지만 Residual Block을 사용할 시 네트워크의 파라미터 수가 약 4% 증가한다. 네트워크 파라미터 수의 증가는 영향이 커질 경우 학습이 느려지거나 네트워크가 과하게 무거워질 수 있다. 하지만 약 4%의 적은 네트워크 파라미터가 증가로 전체적인 성능을 증가시킬 수 있는 방법이다.

### ACKNOWLEDGMENT

"본 연구는 과학기술정보통신부 및 정보통신기획평가원의대학 ICT 연구센터육성지원사업의 연구결과로 수행되었음"(IITP-2021-2016-0-00288)

### 참 고 문 헌

- [1] Goodfellow, Ian J., et al. "Generative adversarial networks." arXiv preprint arXiv:1406.2661 (2014).
- [2] Radford, Alec, Luke Metz, and Soumith Chintala. "Unsupervised representation learning with deep convolutional generative adversarial networks." arXiv preprint arXiv:1511.06434 (2015).
- [3] Mirza, Mehdi, and Simon Osindero. "Conditional generative adversarial nets." arXiv preprint arXiv:1411.1784 (2014).
- [4] Brock, Andrew, Jeff Donahue, and Karen Simonyan. "Large scale GAN training for high fidelity natural image synthesis." arXiv preprint arXiv:1809.11096 (2018).
- [5] Karras, Tero, Samuli Laine, and Timo Aila. "A style-based generator architecture for generative adversarial networks." Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2019.
- [6] Arjovsky, Martin, Soumith Chintala, and Léon Bottou. "Wasserstein generative adversarial networks." International conference on machine learning. PMLR, 2017.
- [7] Gulrajani, Ishaan, et al. "Improved training of wasserstein gans." arXiv preprint arXiv:1704.00028 (2017).