

GAN을 이용한 손글씨 이미지 생성

이경근, 송현철, 최광남*
 중앙대학교 컴퓨터공학부
 *e-mail : knchoi@cau.ac.kr

Handwriting Images Based on GAN

Kyung Keun Lee, Hyun Chul Song, Kwang Nam Choi*
 Dept of Computer Science and Engineering, Chung-Ang University

요 약

Generative Adversarial Networks(GAN)는 생성자가 만든 가짜 이미지를 최대한 실존 이미지에 가깝게 만들어지도록 디자인된 모델이다. 모델의 구조는 실제와 가까운 샘플들을 생성하려고 하는 생성자, 학습에 필요한 이미지와 생성자가 만든 이미지를 판별하는 구분자로 구성되어 있다. 모델은 생성자와 구분자가 서로 대립하는 형태로 학습된다. 본 논문의 GAN은 숫자형태의 손글씨로 된 이미지를 생성한다.

1. 서론

비지도학습 (Unsupervised Learning)의 대표적인 모델인 GAN은 실제 이미지와 매우 비슷한 이미지를 만들어내는 결과를 보여주면서 많은 관심을 가지게 되었다.

그림 1은 구글에서 발표한 Boundary Equilibrium Generative Adversarial Networks(BEGAN) 모델로 생성한 이미지이다. 모두 실존하는 인물이 아니며 가짜이미지이다. 이미지와 레이블을 사용하여 지도학습을 하는 모델과는 달리 레이블 없이 새로운 데이터를 생성해 낼 수 있다는 점에서 BEGAN은 학습한 데이터의 특징을 이해하고 분류할 수 있다는 것을 보여준다. [1]



그림 1. Google의 BEGAN모델로 생성한 이미지

2. 관련연구

기계학습(Machine Learning)에서 라벨링을 하여 모델을 학습하는 방법에 대한 연구는 지속적으로 진행되어 왔으

며, 지도학습(Supervised Learning)을 기반으로 하는 Convolution Neural Networks(CNN)이나 Recurrent Neural Networks(RNN)에서 분류문제나 회귀문제를 해결하는 것이 대표적이라 할 수 있다.[2]

지도학습(Supervised Learning)은 각각의 이미지에 대하여 Annotation을 반드시 필요로 한다. 이점은 모델을 학습하기 위해 필요한 데이터를 만드는데 어려운 점으로 작용한다.

이에 반해 GAN은 이미지에 대한 Annotation을 필요로 하지 않기 때문에 학습 데이터를 구성하는데 어려운 점을 해결 할 수 있다.

3. 설계 및 구현

GAN은 2가지 요소인 구분자와 생성자로 설계 되어있다. 구분자는 이미지를 받아 실제 존재하는 이미지는 1으로, 생성자가 생성한 가짜이미지는 0으로 판단한다.

GAN의 손실함수는 Cross Entropy를 사용한다. 수식 (1)의 첫 번째 항은 실존이미지 x 에 대한 구분자의 Loss이고 두 번째 항은 생성자가 생성한 가짜이미지 $G(z)$ 에 대한 구분자의 Loss이다. 학습에 사용한 데이터는 28*28 크기를 가지고 있으며 출력은 정확도를 나타내는 확률 값한자리로 나타낸다.

$$L_{adv} = E_x [\log D_{src}(x)] + E_z [\log (1 - D_{src}(G(z)))] \quad (1)$$

그림2의 구분자D는 실제 이미지를 받아 손실함수를 통해 오차를 구한다. 생성자 G는 임의의 노이즈 이미지를

생성하고, 생성된 노이즈 이미지는 구분자에 입력된다. 구분자는 생성자가 만든 이미지를 가짜이미지로 구분하려 하지만 생성자는 구분자가 실존이미지로 판단하도록 학습된다. 최종적으로 구분자가 가짜이미지를 판별할 수 없는 상태가 되면 생성자는 실존이미지와 비슷한 이미지를 생성할 수 있는 모델로 학습이 된 것이다.[3]

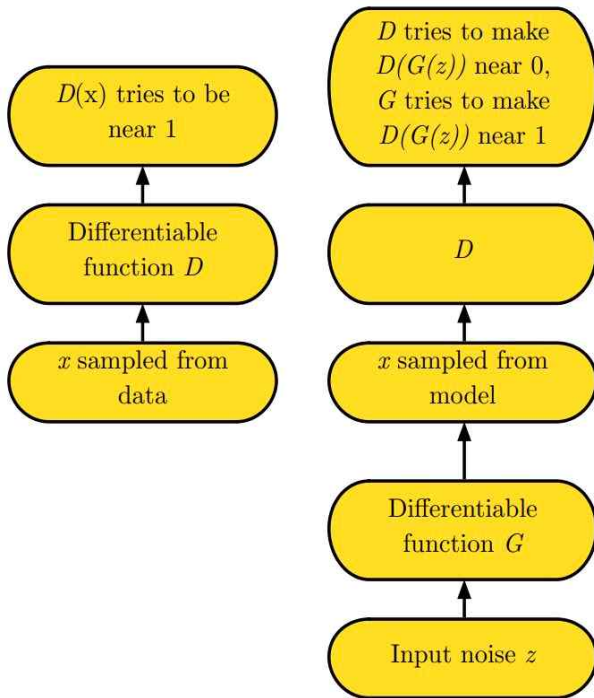


그림 2 GAN의 구조

4. 실험 결과

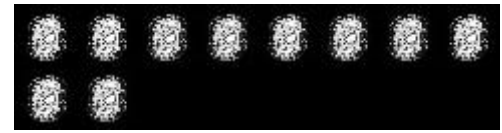
GAN의 구분자는 생성자가 만든 이미지의 진위여부를 판단하기 위해 학습된다. 학습데이터는 총 60,000장이며 Pytorch에서 제공하는 28x28사이즈의 MNIST를 사용하였다. 생성자는 구분자를 속이기 위해 실존이미지와 비슷한 이미지를 생성한다. GAN의 생성자는 초기에 임의의 노이즈 이미지를 그림3과 같이 생성한다.



그림 3 노이즈 이미지

생성자는 초기의 랜덤한 노이즈 이미지에서 실존이미지와 비슷한 이미지를 만들어 낸다. 그림 4는 생성자가 만들어 낸 이미지이고 3000번째 이미지에서 숫자와 비슷한 형태를 만들어 냈으며 20000번째 이미지에서는 선명한 글씨

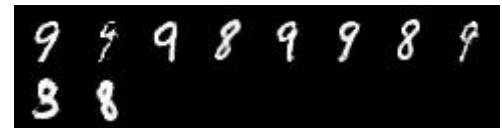
를 확인 할 수 있다.



<100회 학습>



<3,000회 학습>



<20,000회 학습>

그림 4 GAN의 학습과정

5. 결론

본 논문에서는 GAN을 이용하여 손글씨 데이터를 학습할 수 있는 모델을 구현하였다. 학습데이터에 대한 어노테이션이 필요하지 않으므로 Annotation에 많은 시간과 비용이 드는 다량의 데이터셋의 경우 유용하게 사용할 수 있다. 이러한 특징을 응용하여 음성 및 자연어 처리로 확장하여 적용 시킨다면 더 넓은 영역에서 GAN을 활용할 수 있을 것이다.

참고문헌

- [1] David Berthelot, Thomas Schumm, Luke Metz , “BEGAN: Boundary Equilibrium Generative Adversarial Networks,” *arXiv:1703.10717*, 2017.
- [2] Z. Liu, P. Luo, X. Wang, and X. Tang, “ Deep learning face attributes in the wild,” IEEE International Conference on Computer Vision (ICCV), 2015.
- [3] I. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, and Y. Bengio., “ Generative adversarial nets” *NIPS 2016*, 2016.