

Tobigs 15기 Week10 GAN – 15기 이성범

DCGAN 논문 리뷰

리뷰

본 논문이 나오기 전까지는 CNN의 경우 지도학습은 크게 적용되어 왔고 많은 관심을 받았다고 한다. 따라서 본 논문은 DCGAN이라는 CNN을 활용한 GAN 아키텍처를 제안하여 CNN을 활용한 비지도 학습의 연구 개발에 참여와 관심을 높이기 위해서 작성되었다.

과거의 CNN을 사용하여 GAN의 성능을 높이려는 시도들이 있었지만 성공적이지 않았다. 본 논문은 CNN을 활용하여 GAN의 성능을 높이기 위해서 다른 접근법을 개발했다. 본 접근법은 LAPGAN의 저자에게 많은 영감을 얻었다 한다.

아래는 안정적인 DCGAN 아키텍처를 위한 논문에서 제시한 가이드 라인이다.

1. pooling layer를 판별자의 경우 strided convolution 으로 대체, 생성자의 경우 fractional-strided convolution으로 대체 -> 코드를 보면 strid를 활용하여 판별자의 경우 이미지의 크기를 줄여나가는 것을, 생성자의 경우 이미지의 크기를 증가시키는 것을 확인할 수 있음
2. 학습의 안정성을 높이고자 생성자와 판별자에 Batch Normalization을 사용 -> 모든 레이어에 사용하면 모델의 학습이 불안정하여 생성자의 출력레이어, 판별자의 입력 레이어에는 적용하는 것을 피했다함
3. 더 깊은 모델을 만들기 위해서 Fully Connected layer를 제거하였다.
4. 생성자의 모든 레이어에 ReLU를 활성화 함수로 사용하고 출력의 경우에만 Tanh을 사용함
5. 판별자의 모든 레이어에 LeakyReLU를 활성화 함수로 사용

아래는 본 논문에서 진행한 학습 방법이며 아래의 학습 방법을 통해서 학습이 안정되었다 한다.

1. 학습에 활용한 이미지 데이터의 경우 전처리가 적용되지 않았으며 Tanh 함수에 의해서 -1~1의 범위로 데이터가 스케일링이 되었다.
2. SGD 방식으로 미니배치사이즈를 128로 학습을 진행했다.
3. 모든 가중치 들은 평균은 0, 표준 편차는 0.02로 초기화 되었다.
4. LeakyReLU의 경우 slope를 0.2로 설정했다.
5. Adam optimizer을 사용했으며 learning late의 경우 0.001은 너무 커 학습이 불안정하다는 것을 발견하여 learning late를 0.0002로 두어 학습을 진행했으며 beta1을 0.9로 두었을 때 학습이 불안정하여 0.5로 두어 학습을 진행했다.

본 논문은 위에서 제시한 DCGAN 가이드라인으로 아키텍처를 구성하고 학습을 진행함으로써 다양한 Task에서 실험을 진행했다. 실험 결과 아직 모델의 불안정성은 존재하며 모델을 오래 학습시키는 것은 판별자와 생성자가 서로를 속고 속이며 제자리를 맴돌아 양쪽 모두 전역해로 수렴할 수 없게 되는 진동 상태에 빠진다고 한다. 하지만 본 논문은 DCGAN이라는 새로운 아키텍처를 제안했다는 것에 대해서 큰 의의를 가지며 GAN이 지도 학습과 생성 모델을 위한 이미지 표현도 잘 학습할 수 있다는 것을 증명했다.

참고자료

- <https://ysbsb.github.io/gan/2020/12/05/DCGAN.html>
- <https://jgrammer.tistory.com/entry/%EB%85%BC%EB%AC%B8-%EB%A6%AC%EB%B7%B0-DCGAN-2016>
- <https://hyunlee103.tistory.com/61>