

Generative Adversarial Nets

<https://arxiv.org/abs/1406.2661> (<https://arxiv.org/abs/1406.2661>)

1. Read paper and organize it in any format

Abstract

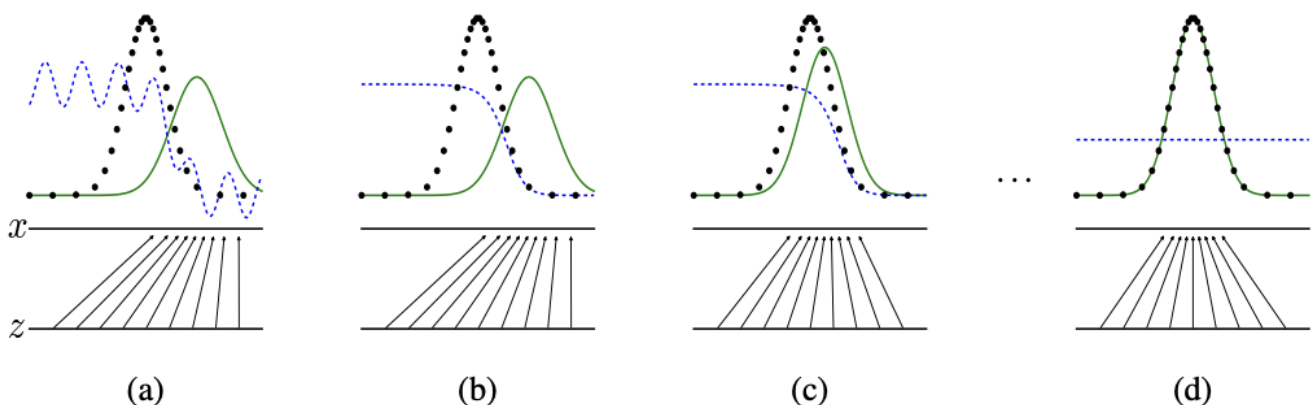
- 데이터의 분포를 따오는 생성모델 G(generator; 생성자)와 G가 아닌 훈련 데이터에서의 샘플일 확률을 추정하는 판별모델 D(Discriminator; 판별자)를 동시에 훈련합니다.
- G에 대한 훈련 절차는 D가 실수할 확률을 최대화하는 것입니다.
- G와 D가 다층 퍼셉트론에 의해 정의되는 경우, 전체 시스템은 역전파로 훈련될 수 있습니다.
- 훈련 또는 샘플 생성 중에 마르코프 체인이나 언롤드 근사 추론 네트워크가 필요하지 않습니다.

Adversarial nets

- adversarial modeling 프레임워크의 경우 모델이 모두 다층 퍼셉트론일 때 적용하기 가장 간단합니다.
- G는 z (random noise variable)로 부터 생성을 시작하고 학습이 진행될수록 실제 데이터의 분포를 따르면서 $D(G(z))$ 의 값이 1이 되도록 학습합니다.
- 반면에 D는 $D(G(z))$ 가 최대한 0에 가까운 값을 출력하도록 학습되어야 합니다.
- $\log(1 - D(G(z)))$ 를 최소화 하기 위해 G를 훈련시킵니다.

$$\min_G \max_D V(D, G) = \mathbb{E}_{\mathbf{x} \sim p_{\text{data}}(\mathbf{x})} [\log D(\mathbf{x})] + \mathbb{E}_{\mathbf{z} \sim p_z(\mathbf{z})} [\log(1 - D(G(\mathbf{z})))]$$

- 반복적이고 수치적인 접근으로 모델을 구현해야 합니다.
- 한번의 iteration마다 D를 k번, G를 1번 학습함으로써 천천히 G를 개선시킵니다.
- 학습 과정에서 inner loop에서 D를 완성하기 위해 최적화 하는 것은 불가능하며 학습 데이터셋의 한계로 오버피팅에 빠질 확률이 높기 때문에 D의 가중치를 계속 업데이트 하지 않습니다.



파란 점선 : 데이터 구분, 검은색 점선 : real data, 녹색 실선 : fake data

- (a) : 학습이 안된 상태, p_g 가 p_{data} 와 비슷함
- (b) : D가 $p_{\text{data}}(x)/(p_{\text{data}}(x) + p_g(x))$ 로 수렴하여 실제 데이터와 생성된 데이터를 구분할 수 있도록 학습
- (c) : G를 학습시킴, D의 gradient는 $G(z)$ 가 실제 데이터와 같은 데이터로 분류하게 안내
- (d) : 반복을 통해 D는 두 데이터의 차이를 구분할 수 없게 됨. $D(x) = 1/2$ 임

Theoretical Results

- 미니배치 SGD 훈련 알고리즘,

Algorithm 1 Minibatch stochastic gradient descent training of generative adversarial nets. The number of steps to apply to the discriminator, k , is a hyperparameter. We used $k = 1$, the least expensive option, in our experiments.

for number of training iterations **do**

for k steps **do**

- Sample minibatch of m noise samples $\{z^{(1)}, \dots, z^{(m)}\}$ from noise prior $p_g(z)$.
- Sample minibatch of m examples $\{x^{(1)}, \dots, x^{(m)}\}$ from data generating distribution $p_{\text{data}}(x)$.
- Update the discriminator by ascending its stochastic gradient:

$$\nabla_{\theta_d} \frac{1}{m} \sum_{i=1}^m \left[\log D(x^{(i)}) + \log (1 - D(G(z^{(i)}))) \right].$$

end for

- Sample minibatch of m noise samples $\{z^{(1)}, \dots, z^{(m)}\}$ from noise prior $p_g(z)$.
- Update the generator by descending its stochastic gradient:

$$\nabla_{\theta_g} \frac{1}{m} \sum_{i=1}^m \log (1 - D(G(z^{(i)}))).$$

end for

The gradient-based updates can use any standard gradient-based learning rule. We used momentum in our experiments.

• Experiments

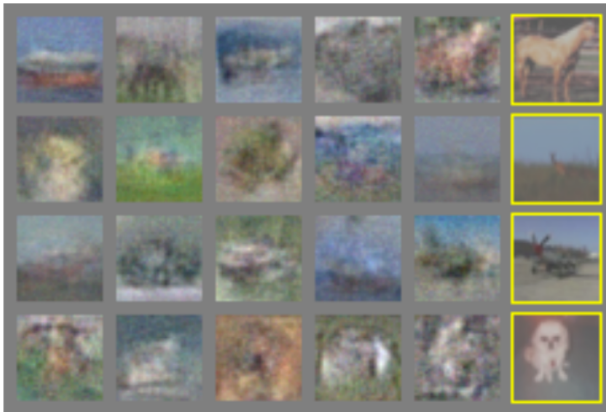
1. MNIST, TFD, CIFAR-10에 대해 훈련
2. generator net는 rectifier linear activation와 시그모이드를 혼합하여 사용합니다.
3. discriminator 훈련시 드롭아웃을 사용하고 maxout activation을 사용합니다.
4. 이론적 프레임워크에서는 generator의 중간층에 드롭아웃과 노이즈를 허용하지 않지만, 실험에서는 맨 하위 계층에 노이즈 input을 사용합니다.



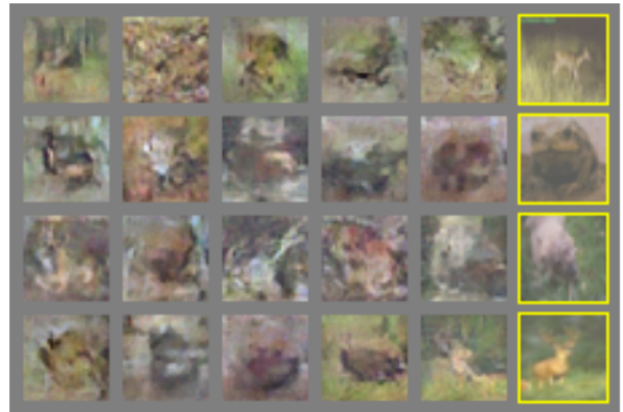
a)



b)



c)



d)

학습을 완료한 생성된 데이터

Advantages

- 역전파로 그래디언트를 구할 수 있기 때문에 markov chain이 필요 없습니다.
- 학습 과정에서 inference 를 할 필요가 없습니다.
- adversarial nets framework와 다양한 함수를 합칠 수 있습니다.

Disadvantages

- 생성된 데이터의 분포를 확실하게 표현하는 것이 없습니다.
- 학습할 때 D와 G가 잘 동기화 되어야 합니다.(synchronized well)
- helvetica scenario; 충분한 다양성을 지닐 수 없게 됩니다.

2. How to evaluate GANs?

수렴할 때까지 손실 함수로 훈련되는 다른 딥 러닝 신경망 모델과 달리 GAN 생성기 모델은 이미지를 실제 또는 생성된 것으로 분류하는 방법을 배우는 판별기라는 두 번째 모델을 사용하여 훈련됩니다. 따라서 GAN 생성기 모델을 훈련하는 데 사용되는 객관적 손실 함수가 없으며 손실만으로 훈련의 진행률과 모델의 상대적 또는 절대적 품질을 객관적으로 평가할 방법이 없습니다. 대신 생성된 합성 이미지의 품질과 다양성을 기반으로 GAN 모델의 성능을 평가하기 위해 일련의 정성 및 정량적 기술이 개발되었습니다.

1. manual evaluation

- 사람이 직접 육안으로 평가하는 가장 일반적이고 직관적인 방법 중 하나

- 생성자가 항상 판별자에게 가장 그럴듯 해 보이는 이미지를 보여주기 위해 같은 이미지를 계속 생성하는 현상인 mode-collapse 현상을 잡아낼 수 없음
- 적용 분야가 한정적

2. qualitative evaluation - rapid scene categorization

- 이미지를 보고 평가하는 시간을 매우 짧게 제한하여 이 시간동안 리얼과 페이크를 구별
- 인간 심판의 성과가 고정적이지 않다는 단점이 있음

3. qualitative evaluation - nearest neighbors

- 도메인에서 실제 이미지의 예시를 선택하고 비교를 위해 하나 이상의 가장 유사한 생성된 이미지를 찾음
- 생성된 이미지가 얼마나 현실적인지 평가하기 위한 맥락으로 쓰기 유용

4. quantitative evaluation - average log likelihood

- 생성자가 얼마나 훈련 이미지의 분포를 잘 캡처했는지 추정하는 방법
- 사소한 모델에 유리하며 시각적 충실도와 관련이 없음
- 일반적으로 gan 평가에 효과적이지 않다고 밝혀짐

5. quantitative evaluation - inception score

- 2016년 improved techniques for training gans 논문에서 처음 제안되었음
- GAN 성능평가에 널리 사용되는 지표
- 이미지의 퀄리티와 다양성을 동시에 측정함 (점수는 0부터 무한까지 매겨짐)
- 구체적으로 각 이미지가 학습된 클래스와 얼마나 많이 닮았는지, 얼마나 다양한 이미지 세트를 캡처하는지를 라벨의 확률분포로 반환하는 등의 과정

6. quantitative evaluation - frechet inception distance

- 실제 이미지와 생성된 이미지에 대해 cv 특징에 대한 통계 측면에서 두 그룹이 얼마나 유사한지 계산
- 점수가 낮을수록 성능이 좋음