
DCGAN

GAN Paper

UNSUPERVISED REPRESENTATION LEARNING
WITH DEEP CONVOLUTIONAL
GENERATIVE ADVERSARIAL NETWORKS (2016)

Alec Radford & Luke Metz
indico Research

논문 리뷰(Abstract)

강력한 비지도학습 방식의 모델 Deep Convolutional Generative Adversarial Networks (DCGANs)을 제시한다.

• DCGAN (Deep Convolutional Generative Adversarial Networks)

- 비지도학습 방식 모델
- 계층적 구조 학습
 - 다양한 데이터셋 환경에서,
 - Object level ~ Scene level까지 representation의 계층적 구조 학습
- Learned representation 품질 확인
 - 학습 feature들을 여러 task에 사용

* representation : 원시 데이터(raw data)를 특성 벡터(feature vector)에 매핑하는 행위

논문 리뷰(Introduction)

대규모의 unlabeled 데이터셋으로부터 재사용 가능한 feature representation을 학습하는 것은 매우 활발한 분야이다. 이 연구에서 그중 하나의 방법으로 GAN을 이용해 좋은 image representation을 얻는 방법을 제시한다.

• Contribution

- 안정적 훈련
 - Convolutional GAN의 구조 소개
 - 한계 제시 & 평가
- 성능 확인
 - 본 논문의 비지도학습기반 표현학습 모델을 다른 지도방식 task에 적용
 - 학습된 discriminator 사용하여 image classification task 수행
- GAN의 학습되는 필터 시각화
- generator의 vector arithmetic 특성 증명

논문 리뷰(Approach and Model Architecture)

광범위한 데이터셋에서 안정적으로 훈련될 수 있고 고해상도의 심층 생성모델이 구축될 수 있게 하는 구조를 제시한다.

• Background

- 일반적인 CNN 아키텍처 사용으로 GAN을 확장하고자 하였으나 어려움
- 광범위한 모델 탐색으로 가능한 아키텍처 식별
 - 다양한 데이터 세트에 걸쳐 안정적인 훈련을 제공
 - 더 높은 해상도와 더 깊은 생성 모델을 훈련할 수 있는 아키텍처 제품군을 식별

논문 리뷰(Approach and Model Architecture)

광범위한 데이터셋에서 안정적으로 훈련될 수 있고 고해상도의 심층 생성모델이 구축될 수 있게 하는 구조를 제시한다.

• CNN 아키텍처에 적용한 변경 사항

Architecture guidelines for stable Deep Convolutional GANs

- Replace any pooling layers with strided convolutions (discriminator) and fractional-strided convolutions (generator).
- Use batchnorm in both the generator and the discriminator.
- Remove fully connected hidden layers for deeper architectures.
- Use ReLU activation in generator for all layers except for the output, which uses Tanh.
- Use LeakyReLU activation in the discriminator for all layers.

* [fractionally-strided convolution](#) : Transposed Convolution과 동일한 업샘플링을 위한 Convolution 연산

논문 리뷰(Details of Adversarial Training)

Convolutional GAN을 훈련시키기 위한 상세 파라미터 값을 소개한다.

• Training Detail Parameter

- 이미지 픽셀값 스케일링
 - -1 ~ 1 범위로 스케일링 수행
 - Generator의 출력층에 사용된 tanh Activation value 범위와 일치
 - 모든 이미지 데이터셋에 대해 위의 scaling을 제외한 별도 augmentation 적용 안함
- Adam optimizer
- learning rate : 0.0002
- 가중치 초기값 $\sim N(0, 0.02)$
- LeakyReLU Value : 0.2
- 배치사이즈 : 128

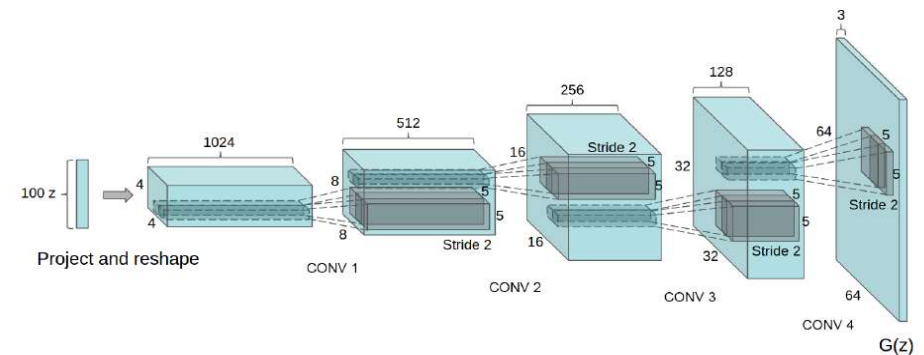


Figure 1: DCGAN generator used for LSUN scene modeling. A 100 dimensional uniform distribution Z is projected to a small spatial extent convolutional representation with many feature maps. A series of four fractionally-strided convolutions (in some recent papers, these are wrongly called deconvolutions) then convert this high level representation into a 64×64 pixel image. Notably, no fully connected or pooling layers are used.

TIP (생성기가 훈련 이미지를 기억하지 않도록 하는 방법)

생성기가 훈련 이미지를 기억해버릴 가능성을 낮추기 위해 서로 매우 비슷한 이미지들을 제거한다.
(3백만장 중 27만 5천장 제거)

• 유사 중복 이미지 제거 방법

- de-noising autoencoder 사용

- 원본 이미지들을 오토인코더의 인코더에 입력하여 coding 벡터로 인코딩
- coding 값들이 서로 비슷한 관측들을 설정한 비율에 맞는 선까지 제거

논문 리뷰(visualizing the discriminator features)

guided backpropagation 방법을 사용하여 discriminator의 학습된 필터들이 activate되는 부분을 시각화한다.

• Feature 시각화

- 목적: 비지도 방식인 DCGAN 또한 강력한 계층적인 특징을 학습할 수 있다는 것을 증명하기 위함
- Guided backpropagation 방법
 - discriminator의 학습된 필터에서 Activate되는 부분을 시각화
 - 학습 전 랜덤하게 초기화 되어있는 필터(좌측)에 비해 학습된 필터들은 특정한 가구, 창문 등에 잘 매칭되도록 되어있는 모습을 확인할 수 있음.

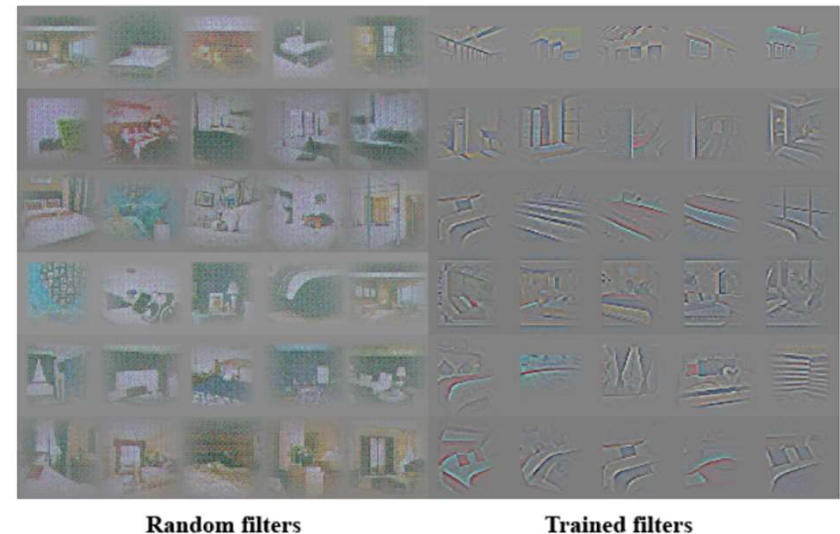


Figure 5: On the right, guided backpropagation visualizations of maximal axis-aligned responses for the first 6 learned convolutional features from the last convolution layer in the discriminator. Notice a significant minority of features respond to beds - the central object in the LSUN bedrooms dataset. On the left is a random filter baseline. Comparing to the previous responses there is little to no discrimination and random structure.

논문 리뷰(vector arithmetic on face samples)

vector들 간의 간단한 산수를 통해 latent representation space에서 풍부한 선형구조를 밝힌다.

- latent vector인 z 에 대한 산수 수행
 - Z : generator의 input
- 결과
 - 단일 벡터들에 대해 연산을 수행할 경우 불안정
 - 3개의 샘플에 대해 평균을 취하여 연산을 수행할 경우 좀 더 안정된 결과 도출
 - 결과는 아래와 같다.

