G A N s t r u d y

박 이 삭

Info_GAN

Interpretable Representation Learning by Information Maximizing Generative Adversarial Nets

Contents

Unit 01 | disentangled representation

Unit 02 | background

Unit 03 | information theory

Unit 04 | info_GAN

Unit 05 | result

Generate Model?

- P(x)의 분포를 모르지만 샘플은 얻을 수 있는 경우, 유사한 Q(x)를 만든다.

- 유사함의 기준: KL 다이버 젼스

\+ 만들기에 그치는 것이 아니라 얼굴데이터를 생성할 경우
눈의 색, hairstyle, 안경의 유무 등의 representation을 학습 할 수 있다면?

Motivation

기존 GAN에서 학습을 진행할 때
$\min_G \max_D V(D,G) = E_{x\sim p_{data}}[logD(x)] + E_{z\sim p_z(z)}[log(1-D(G(z))]$ z에 제약이 없기 때문에 이미지의 특징을 잡아 낼 필요가 없었다. (혹은 잡아내지 못했다.)

How can we achieve
unsupervised learning of disentangled representation?

In general, learned representation is entangled,
i.e. encoded in a data space in a complicated manner

그렇기 때문에 이미지의 특성이 꼬인(entangled) 형태로 학습이 된다.

When a representation is **disentangled**, it would be
more interpretable and easier to apply to tasks

꼬임을 풀어(disentangled) 학습을 진행 한다면 특징들을 잡아 낼 수 있을 것이다.

지도 학습보다는 비지도 학습으로

Label을 이용해 특징을 학습하기 보다 연속형(혹은 범주형)변수를 이용해 특징을 학습시키는 것

Latent code 'C' 는 데이터의 특징을 목표로 하는 변수

예) $c_1 \sim category(k=10, p=0.1)$ 이면 mnist의 0~9까지 대응시켜 CGAN의 역할도 가능

$c_2 \sim unif(-1, 1)$ 이면 회전 이나 두께에 대응시킬 수 있다.

의미없는 code를 학습하는 것을 막기위해 info_GAN은 생성분포 G 와 c 에

상호 정보량(Mutual Information)이 높도록 제약조건을 부여한다.

$\boldsymbol{I[c\,;G(z,c)]}$ 높아야 한다.

Variational Information Maximization I

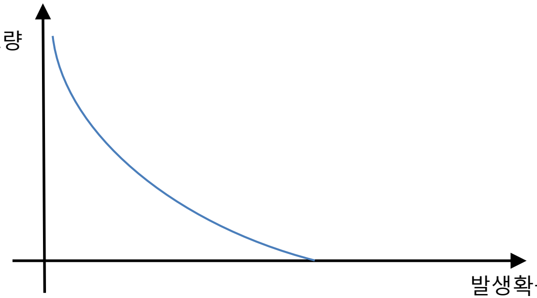
- 상호정보량 $I(\boldsymbol{c}; G(\boldsymbol{z}, \boldsymbol{c}))$을
  - 계산에 $P(\boldsymbol{c}|\boldsymbol{x})$가 필요하지만 직접 최대화가 불가능
  - 보조분포 $Q(\boldsymbol{c}|\boldsymbol{x})$를 사용하여 low limit을 구한다.
    - Variational Information Maximization을 이용 [Barber and Agakov, 2003]
      - 보조분포를 도입시 KL 정보량으로 low limit을 설정

$$
\begin{aligned}
I(c; G(z,c)) &= H(c) - H(c|G(z,c)) \\
&= \mathbb{E}_{x\sim G(z,c)}[\mathbb{E}_{c'\sim P(c|x)}[\log P(c'|x)]] + H(c) \\
&= \mathbb{E}_{x\sim G(z,c)}[\underbrace{D_{\mathrm{KL}}(P(\cdot|x) \parallel Q(\cdot|x))}_{\geq 0} + \mathbb{E}_{c'\sim P(c|x)}[\log Q(c'|x)]] + H(c) \\
&\geq \mathbb{E}_{x\sim G(z,c)}[\mathbb{E}_{c'\sim P(c|x)}[\log Q(c'|x)]] + H(c)
\end{aligned}
$$

$I[c\,;G(z,c)]$?엔트로피

정보의 양을 $h(x) = -log\,p(x)$로 정의하면

(낮은 확률로 발생하는 사건을 높은 정보량을 부여하기 위해)



$h(x)$의 기댓값 $-\sum_x p(x)log\,p(x) = H(x)$ 라고 정의하고 '엔트로피' 라고 부른다.

H(x,y)는 흔히 말하는 cross-entropy 라고 하며 $H(x,y) = -\sum_x p(x)log\,q(y)$ 로 정의한다.

$I[c\,;G(z,c)]$?

KL 다이버젼스는 두 확률분포 p q 간의 괴리로 정의하며

$KL(p||q) = H(x,y) - H(x)$로 정의하고 GAN에 빗대어 설명한다면

"어떤 확률분포 p 가 있을 때 P를 알 수 없어, P어서 얻은 sample을 이용해 근사적으로 q를 만들어 p대신 사용할 때 생기는 엔트로피 변화의 의미" 이다.

$$= -\sum_x p(x) log\, q(x) - \left(-\sum_x p(x) log\, p(x) \right)$$

$$= -\sum_x p(x) log \frac{p(x)}{q(x)}$$

$I[c\,;G(z,c)]$?

결합 확률분포(joint)도 GAN에서 빼놓을 수 없다…

$p(x,y) = \int_x p(x)p(y|x)\,dx$로 정의 된다. X를 이미 알고 있다면 결합분포를 구하는데 필요한 정보량은 $-log\,p(y|x)$가 된다.

이를 조건부 엔트로피라고 하고

$H(y|x) = -\sum_x \sum_y p(x,y) log\,p(y|x)$ 로 정의한다.

$= \sum_x p(x) \sum_y p(y|x) log\,(y|x)$ 베이즈 규칙

조건부 엔트로피는 다음과 같은 성질이 있다.

$$H(x,y) = H(y|x) + H(x)$$

(x,y)를 설명하기 위해서는 x만의 정보 H(X) 와 x가 주어졌을 때 y를 기술하는데 필요한 정보의 합

$I[c\ ; G(z, c)]$?

마지막으로 결합분포(joint) 와 주변분포(marginal)를 정보 기준으로 살펴보자.

만약, x와 y가 독립이라면 p(x,y) = p(x)p(y) 가 될 것이다.

하지만 독립이 아니라면 둘의 차이는 존재하고 이를 KL로 측정할 수 있다.

$KL(\ p(x,y)||p(x)p(y)\ )$

$$= \sum_x \sum_y p(x,y) log\, p(x)p(y) - \{-\sum_x \sum_y p(x,y) log p(x,y)\}$$

$$= \sum_x \sum_y p(x,y) log \frac{p(x,y)}{p(x)p(y)} = H(x) - H(x|y)$$

이를 상호 정보라고 한다.

$\boldsymbol{I[x\,;y] \geq 0}$ 이고 독립일때 0을 만족한다. 즉, $\boldsymbol{I[x;y]}$는 연관이 있다면 0보다 큰 값을 갖는다.

$$\min_{G}\max_{D} V_I = V(G,D) - \lambda I(c\,;G(z,c)\,)$$

Info_GAN 에서 상보 정보량은 c 가 변하면 생성함수 G도 변화시켜 만들어 내는 이미지를 변화하게끔 학습시키는 것!!

$$I[c\,;G(z,c)] = H(c) - H(c|G(z,c))$$

$$= \sum_{c}\sum_{x} p(c,x)\, log\, p(c|x), \text{단}\ x\sim G(z,c)$$

$$= \sum_{x} p(x) \sum_{c} p(c|x) log\, p(c|x)$$

$p(c|x)$가 필요하지만 p(x)는 구할 수 없다.

$$= \sum_{x} p(x) \sum_{c} p(c|x) log \frac{P(c|x))}{Q(c|x)} Q(c|x)$$

$Q(c|x)$라는 보조분포(auxiliary)를 이용해 근사시키는 방법을 이용한다.

$$= \sum_{x} \boldsymbol{p(x)} \{ \underbrace{\sum_{c} \boldsymbol{p(c|x)log\frac{p(c|x)}{Q(c|x)}}}_{\text{KL다이버 젼스} \geq 0} + \sum_{c} \boldsymbol{p(c|x)logQ(c|x)} \}$$

$$\therefore I(c\,;G(z,c)) \geq \sum_{x} \boldsymbol{p(x)} \sum_{c} \boldsymbol{p(c|x)logQ(c|x)} + \boldsymbol{H(c)}$$

$$
\begin{aligned}
\mathbb{E}_{x\sim X, y\sim Y|x}\left[f(x,y)\right] &= \int_{\mathcal{X}}\int_{\mathcal{Y}} P(x)P(y|x)f(x,y)dydx && (\because \text{ Definition of pdf})\\
&= \int_{\mathcal{X}}\int_{\mathcal{Y}} P(x,y)f(x,y)dydx && (\because \text{ Bayes rule})\\
&= \int_{\mathcal{X}}\int_{\mathcal{Y}} P(x',y)f(x',y)dydx' && (\because \text{ no difference})\\
&= \int_{\mathcal{X}}\int_{\mathcal{Y}} \left[\int_{\mathcal{X}} P(x|y)dx\right] P(x',y)f(x',y)dydx' && \left(\because \int_{\mathcal{X}} P(x|y)dx = 1\right)\\
&= \int_{\mathcal{X}}\int_{\mathcal{Y}} \left[\int_{\mathcal{X}} P(x|y)\frac{P(y)}{P(y)}dx\right] P(x',y)f(x',y)dydx' && (\because \text{ Bayes rule})\\
&= \int_{\mathcal{X}}\int_{\mathcal{Y}} \left[\int_{\mathcal{X}} P(x,y)dx\right] P(x'|y)f(x',y)dydx' && (\because \text{ Bayes rule})\\
&= \int_{\mathcal{X}}\int_{\mathcal{Y}} \left[\int_{\mathcal{X}} P(x)P(y|x)dx\right] P(x'|y)f(x',y)dydx' && (\because \text{ Bayes rule})\\
&= \int_{\mathcal{X}}\int_{\mathcal{Y}}\int_{\mathcal{X}} P(x)P(y|x)P(x'|y)f(x',y)dx'dydx && (\because \text{ Fubini theorem})\\
&= \mathbb{E}_{x\sim X, y\sim Y|x, x'\sim X|y}\left[f(x',y)\right]
\end{aligned}
$$

$$L_I(G,Q) = E_{c\sim P(c).x\sim G(z,c)}[logQ(c|x)] + H(c)$$ ← Goal eq

$$= E_{c\sim P(c),x\sim P_G(x|z,c)}[logQ(c|x)] + H(c)$$

$$= E_{c\sim P(c),x\sim P_G(x|z,c),c'\sim P(c|x)}[logQ(c'|x)] + H(c)$$ ← Lemma 5.1

$$= E_{x\sim P_G(x|z,c),c'\sim P(c|x)}[logQ(c'|x)] + H(c)$$

$$= E_{x\sim P_G(x|z,c)}\left[E_{c'\sim P(c|x)}[logQ(c'|x)]\right] + H(c)$$

$$\leq I(c; G(z,c))$$

최종적으로 다음과 같은 식을 얻을 수 있다.

$$\min_{G,Q} \max_{D} V_{InfoGAN}(D,G,Q) = V(D,G) - \lambda L_I(G,Q)$$
$$L_I(G,Q) = E_{c\sim P(c).x\sim G(z,c)}[logQ(c|x)] + H(c)$$

생성된 c는 G에 영향을 준다.
G에서 x를 생성한 다음 Q를 통해 다시 c로 돌아 온다.

그런데!! 다시 돌아온 c 가 처음과 같아야 한다는 내용!

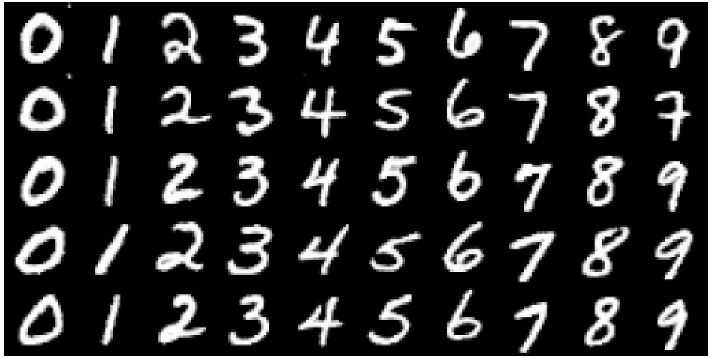
이를 Reconstruction loss 라고 한다.

최종적으로 다음과 같은 식을 얻을 수 있다.

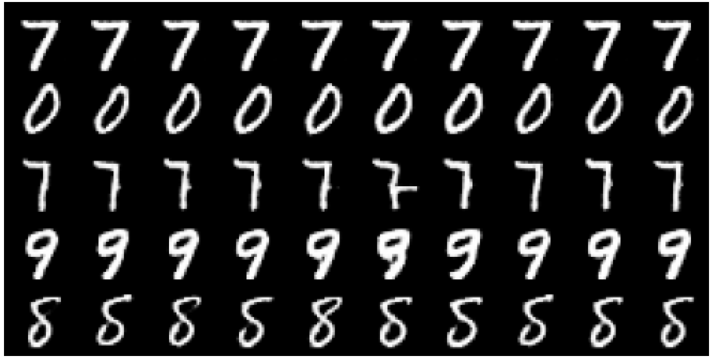
$$\min_{G,Q}\max_{D} V_{InfoGAN}(D,G,Q) = V(D,G) - \lambda L_I(G,Q)$$
$$L_I(G,Q) = E_{c\sim P(c).x\sim G(z,c)}[logQ(c|x)] + H(c)$$

Intuitions
* V(D,G)는 기존 GAN과 같다.
* G와 Q는 L_I도 최대화 해야한다.
* 즉 Q는 G(z,c)를 다시 c로 잘 바꿔야 하고,
* G 는 Q가 잘 바꿀 수 있도록 x=G(z,c)를 생성해야 한다.

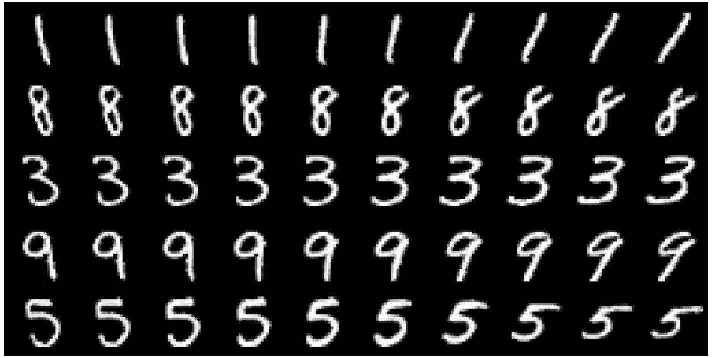
Unit 05 | result



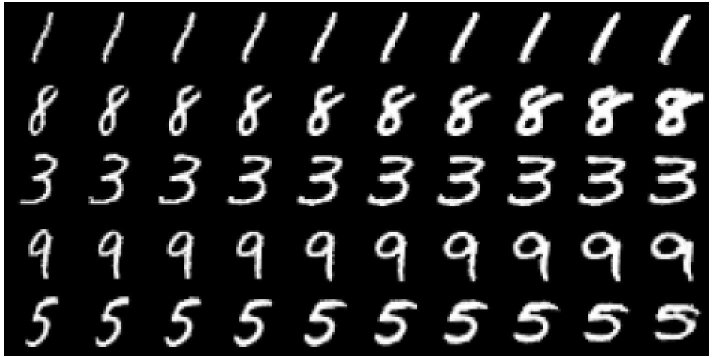
(a) Varying c_1 on InfoGAN (Digit type)



(b) Varying c_1 on regular GAN (No clear meaning)



(c) Varying c_2 from -2 to 2 on InfoGAN (Rotation)



(d) Varying c_3 from -2 to 2 on InfoGAN (Width)

Unit 05 | result

(a) Azimuth (pose) (b) Elevation

(c) Lighting (d) Wide or Narrow

(a) Rotation (b) Width

PR-022:infoGAN https://www.youtube.com/watch?v=_4jbgniqt_Q&t=905s

초짜 대학원생 http://jaejunyoo.blogspot.com/2017/03/infogan-1.html

infoGAN 코드 : https://github.com/openai/infoGAN

Lemma5.1 해석:
https://www.facebook.com/groups/TensorFlowKR/permalink/484676605206736/

상호 정보량: http://sanghyukchun.github.io/62/
http://norman3.github.io/prml/docs/chapter01/6.html