

Autoencoding beyond pixels using a learned similarity metric

齐家兴

Hebei University

July 2, 2019

Outlines

- 1 VAEGAN解决的问题
- 2 引言
 - Siamese网络
 - VAEGAN
- 3 VAEGAN的训练

Outlines

1 VAEGAN解决的问题

2 引言

- Siamese网络
- VAEGAN

3 VAEGAN的训练

- 结合GAN和VAE，用来学习一个更好的相似性度量。用特征的相似性替代了原有的数据的相似性。
- 解决了数据的相似性在图像数据上对少量的平移敏感的问题
- 使VAE生成的图像更加清晰真实

- 结合GAN和VAE，用来学习一个更好的相似性度量。用特征的相似性替代了原有的数据的相似性。
- 解决了数据的相似性在图像数据上对少量的平移敏感的问题
- 使VAE生成的图像更加清晰真实

- 结合GAN和VAE，用来学习一个更好的相似性度量。用特征的相似性替代了原有的数据的相似性。
- 解决了数据的相似性在图像数据上对少量的平移敏感的问题
- 使VAE生成的图像更加清晰真实

The main contributions of VAEGAN:

- 结合GAN和VAE，提出了VAEGAN模型.
- 在VAEGAN中，GAN的判别器用来学习数据的特征表示。
- 将Siamese网络的思想与GAN结合
- 将GAN的生成器，与VAE结合

The main contributions of VAEGAN:

- 结合GAN和VAE，提出了VAEGAN模型.
- 在VAEGAN中，GAN的判别器用来学习数据的特征的代表。
- 将Siamese网络的思想与GAN结合
- 将GAN的生成器，与VAE结合

The main contributions of VAEGAN:

- 结合GAN和VAE，提出了VAEGAN模型.
- 在VAEGAN中，GAN的判别器用来学习数据的特征的代表。
- 将Siamese网络的思想与GAN结合
- 将GAN的生成器，与VAE结合

The main contributions of VAEGAN:

- 结合GAN和VAE，提出了VAEGAN模型.
- 在VAEGAN中，GAN的判别器用来学习数据的特征的代表。
- 将Siamese网络的思想与GAN结合
- 将GAN的生成器，与VAE结合

Outlines

1 VAEGAN解决的问题

2 引言

- Siamese网络
- VAEGAN

3 VAEGAN的训练

Outlines

1 VAEGAN解决的问题

2 引言

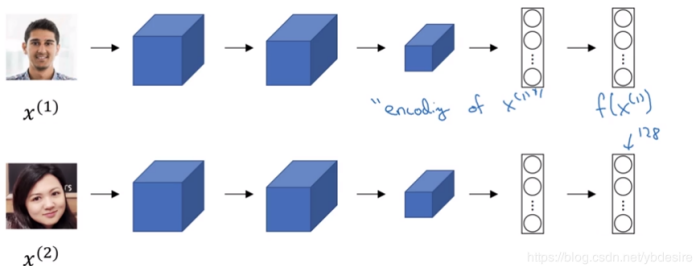
- Siamese网络
- VAEGAN

3 VAEGAN的训练

在人脸识别中，存在所谓的**one-shot**问题。举例来说，就是对公司员工进行人脸识别，每个员工只给你一张照片（训练集样本少），并且员工会离职、入职（每次变动都要重新训练模型）。有这样的问题存在，就没办法直接训练模型来解决这样的分类问题了。

为了解决**one-shot**问题，我们会训练一个模型来输出给定两张图像的相似度，所以模型学习得到的是**similarity**函数。

网络结构：



Siamese网络并不直接输出类别，而是输出一个向量。
所以通过对上下两个模型输出的向量做欧氏距离计算，就能得到
输入两幅图像的相似度。

训练方式:

Learning Objective

Diagram illustrating the learning objective of a Siamese network. It shows two pairs of face images. The first pair, labeled "Anchor" (A) and "Positive" (P), shows two photos of the same person. The second pair, labeled "Anchor" (A) and "Negative" (N), shows two photos of different people. Below the images, the learning objective is written as:

$$\text{Want: } \frac{\|f(A) - f(P)\|^2}{d(A, P)} \leq \frac{\|f(A) - f(N)\|^2}{d(A, N)}$$

对图中的一幅照片A，如果给定了同一个人的另一幅照片P，则模型的输出向量 $f(A)$ 和 $f(P)$ 应该是距离比较小的。如果给定了另一个人的照片N，则模型的输出向量 $f(A)$ 和 $f(N)$ 之间的距离就比较大。所以 $d(A, P) < d(A, N)$ 。

Loss Function:

$$\sum_i^N \left[\|f(x_i^a) - f(x_i^p)\|_2^2 - \|f(x_i^a) - f(x_i^n)\|_2^2 + \alpha \right]_+$$

其中，参数 α 用于做margin，能避免模型输出的都是零向量。

Outlines

1 VAEGAN解决的问题

2 引言

- Siamese网络
- **VAEGAN**

3 VAEGAN的训练

- **VAE** 可以实现从原始分布编码到一个正态分布，再从正态分布解码回原始分布的功能，然后解码器可以单独被提取出来作为生成器使用。
- VAE目标是重构误差损失越小越好，但是越小不见得图片就会越清晰。
- 因此VAEGAN可以理解为，利用GANs去提升VAE的图像生成质量。

- VAE 可以实现从原始分布编码到一个正态分布，再从正态分布解码回原始分布的功能，然后解码器可以单独被提取出来作为生成器使用。
- VAE目标是重构误差损失越小越好，但是越小不见得图片就会越清晰。
- 因此VAEGAN可以理解为，利用GANs去提升VAE的图像生成质量。

- **VAE** 可以实现从原始分布编码到一个正态分布，再从正态分布解码回原始分布的功能，然后解码器可以单独被提取出来作为生成器使用。
- **VAE**目标是重构误差损失越小越好，但是越小不见得图片就会越清晰。
- 因此**VAEGAN**可以理解为，利用**GANs**去提升**VAE**的图像生成质量。

- GAN的判别器的第L层输出，作为数据的特征表示，用该层的特征衡量不同输入之间的相似性。
- VAE中的解码器，和GAN中的生成器参数共享

- GAN的判别器的第L层输出，作为数据的特征表示，用该层的特征衡量不同输入之间的相似性。
- VAE中的解码器，和GAN中的生成器参数共享

VAEGAN结构:

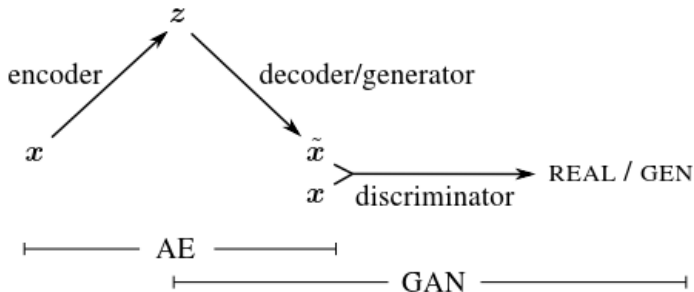
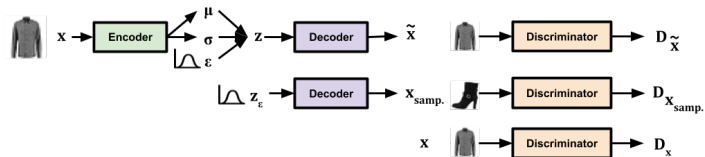


Figure 1. Overview of our network. We combine a VAE with a GAN by collapsing the decoder and the generator into one.



VAEGAN的优点:

- VAEGAN利用GANs判断生成图片的真实性，解决了VAE生成图片不清晰的问题

Outlines

1 VAEGAN解决的问题

2 引言

- Siamese网络
- VAEGAN

3 VAEGAN的训练

损失函数:

$$\mathcal{L} = \mathcal{L}_{\text{prior}} + \mathcal{L}_{\text{like}}^{\text{Dis}} + \mathcal{L}_{\text{GAN}}$$

其中:

$$\mathcal{L}_{\text{prior}} = D_{\text{KL}}(q(\mathbf{z}|\mathbf{x})||p(\mathbf{z}))$$

$$\mathcal{L}_{\text{like}}^{\text{Dis}} = -\mathbb{E}_{q(\mathbf{z}|\mathbf{x})} [\log p(\text{Dis}_I(\mathbf{x})|\mathbf{z})]$$

$$\mathcal{L}_{\text{GAN}} = \log(\text{Dis}(\mathbf{x})) + \log(1 - \text{Dis}(\text{Gen}(\mathbf{z})))$$

VAEGAN训练过程:

Algorithm 1 Training the VAE/GAN model

 $\theta_{\text{Enc}}, \theta_{\text{Dec}}, \theta_{\text{Dis}} \leftarrow$ initialize network parameters**repeat** $\mathbf{X} \leftarrow$ random mini-batch from dataset $\mathbf{Z} \leftarrow \text{Enc}(\mathbf{X})$ $\mathcal{L}_{\text{prior}} \leftarrow D_{\text{KL}}(q(\mathbf{Z}|\mathbf{X})\|p(\mathbf{Z}))$ $\tilde{\mathbf{X}} \leftarrow \text{Dec}(\mathbf{Z})$ $\mathcal{L}_{\text{llike}}^{\text{Dis}_l} \leftarrow -\mathbb{E}_{q(\mathbf{Z}|\mathbf{X})} [p(\text{Dis}_l(\mathbf{X})|\mathbf{Z})]$ $\mathbf{Z}_p \leftarrow$ samples from prior $\mathcal{N}(\mathbf{0}, \mathbf{I})$ $\mathbf{X}_p \leftarrow \text{Dec}(\mathbf{Z}_p)$ $\mathcal{L}_{\text{GAN}} \leftarrow \log(\text{Dis}(\mathbf{X})) + \log(1 - \text{Dis}(\tilde{\mathbf{X}}))$
 $\quad + \log(1 - \text{Dis}(\mathbf{X}_p))$

// Update parameters according to gradients

 $\theta_{\text{Enc}} \xleftarrow{+} -\nabla_{\theta_{\text{Enc}}} (\mathcal{L}_{\text{prior}} + \mathcal{L}_{\text{llike}}^{\text{Dis}_l})$ $\theta_{\text{Dec}} \xleftarrow{+} -\nabla_{\theta_{\text{Dec}}} (\gamma \mathcal{L}_{\text{llike}}^{\text{Dis}_l} - \mathcal{L}_{\text{GAN}})$ $\theta_{\text{Dis}} \xleftarrow{+} -\nabla_{\theta_{\text{Dis}}} \mathcal{L}_{\text{GAN}}$ **until** deadline

Thanks!