Hung-yi Lee Generative Adversaria

Generator: a neural network

Gan discriminator predict

假设数据集内部元素呈现线性分布, 于是可以遍历的方式拿到所有可能的数据集, 这些数据及在discriminator中分数最高的即为预测值, 这些数据集中, 属于training的应该让dircriminator给出高分, 不属于training的dircriminator应该给出低分, 借此完成discirminatro的独立training

Gan Feature Extraction

infoGan

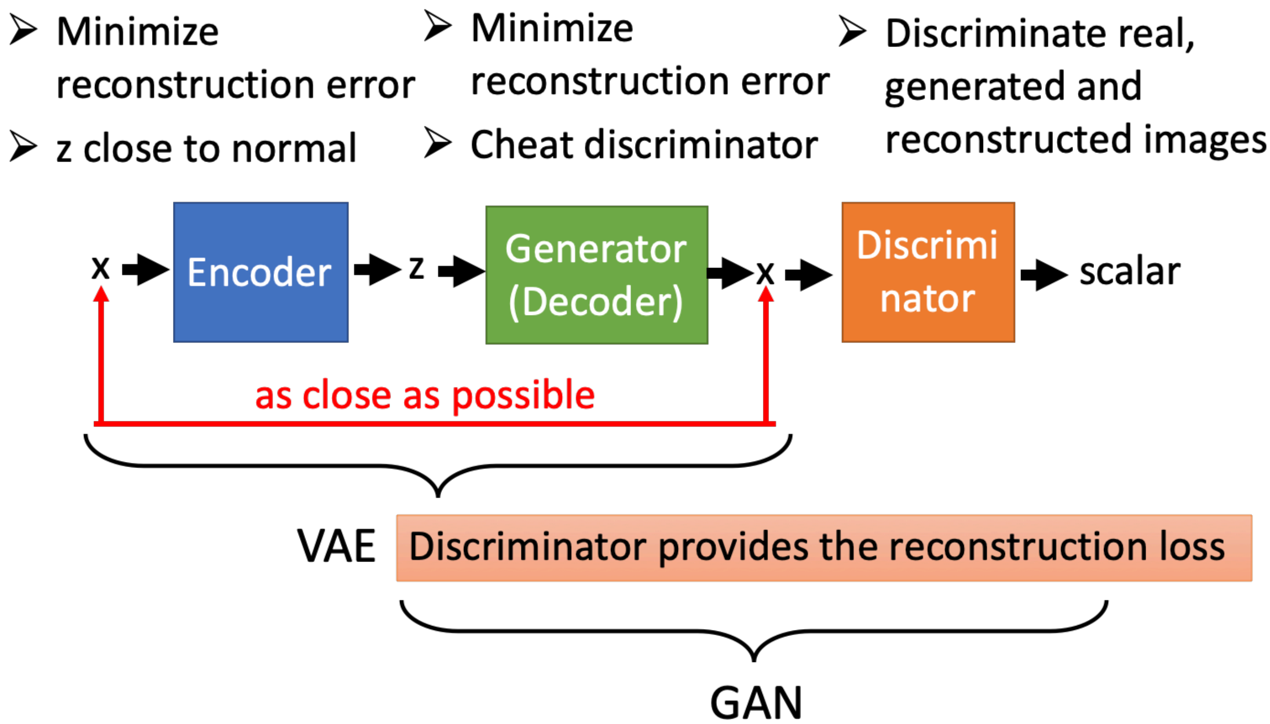
InfoGan

假设我们打算生成像MNIST那样的手写数字图像，每个手写数字可以分解成多个维度特征：代表的数字、倾斜度、粗细度等等，在标准GAN的框架下，我们无法在上述维度上具体指定Generator生成什么样的手写数字。

为了解决这一问题，文章对GAN的目标函数进行了一些小小的改进，成功让网络学习到了可解释的特征表示（即论文题目中的interpretable representation）。

infoGan理解

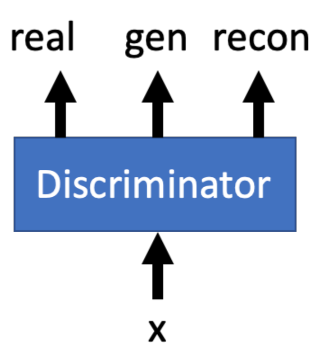
VAE-GAN



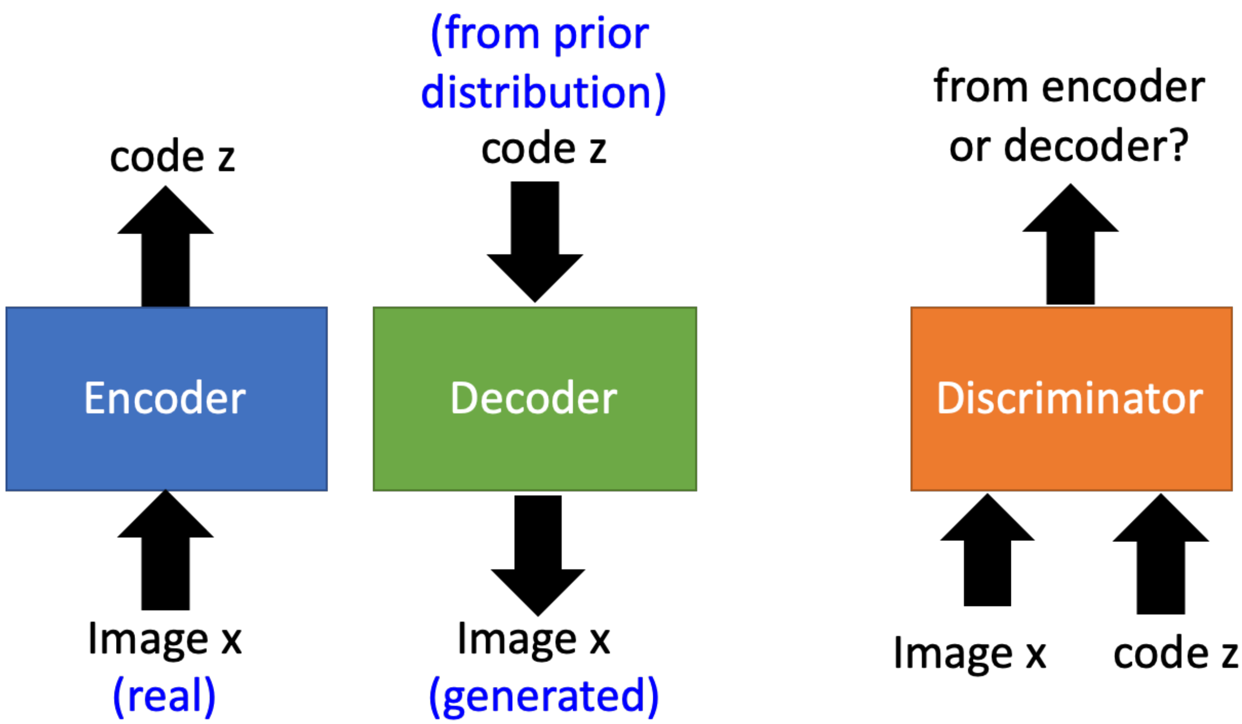
Algorithm

- Initialize En, De, Dis
- In each iteration:
  - Sample M images $x^1, x^2, \cdots, x^M$ from database
  - Generate M codes $\tilde{z}^1, \tilde{z}^2, \cdots, \tilde{z}^M$ from encoder
    - $\tilde{z}^i = En(x^i)$
  - Generate M images $\tilde{x}^1, \tilde{x}^2, \cdots, \tilde{x}^M$ from decoder
    - $\tilde{x}^i = De(\tilde{z}^i)$
  - Sample M codes $z^1, z^2, \cdots, z^M$ from prior P(z)
  - Generate M images $\hat{x}^1, \hat{x}^2, \cdots, \hat{x}^M$ from decoder
    - $\hat{x}^i = De(z^i)$
  - Update En to decrease $\|\tilde{x}^i - x^i\|$, decrease KL(P($\tilde{z}^i$|x^i)||P(z))
  - Update De to decrease $\|\tilde{x}^i - x^i\|$, increase $Dis(\tilde{x}^i)$ and $Dis(\hat{x}^i)$
  - Update Dis to increase $Dis(x^i)$, decrease $Dis(\tilde{x}^i)$ and $Dis(\hat{x}^i)$

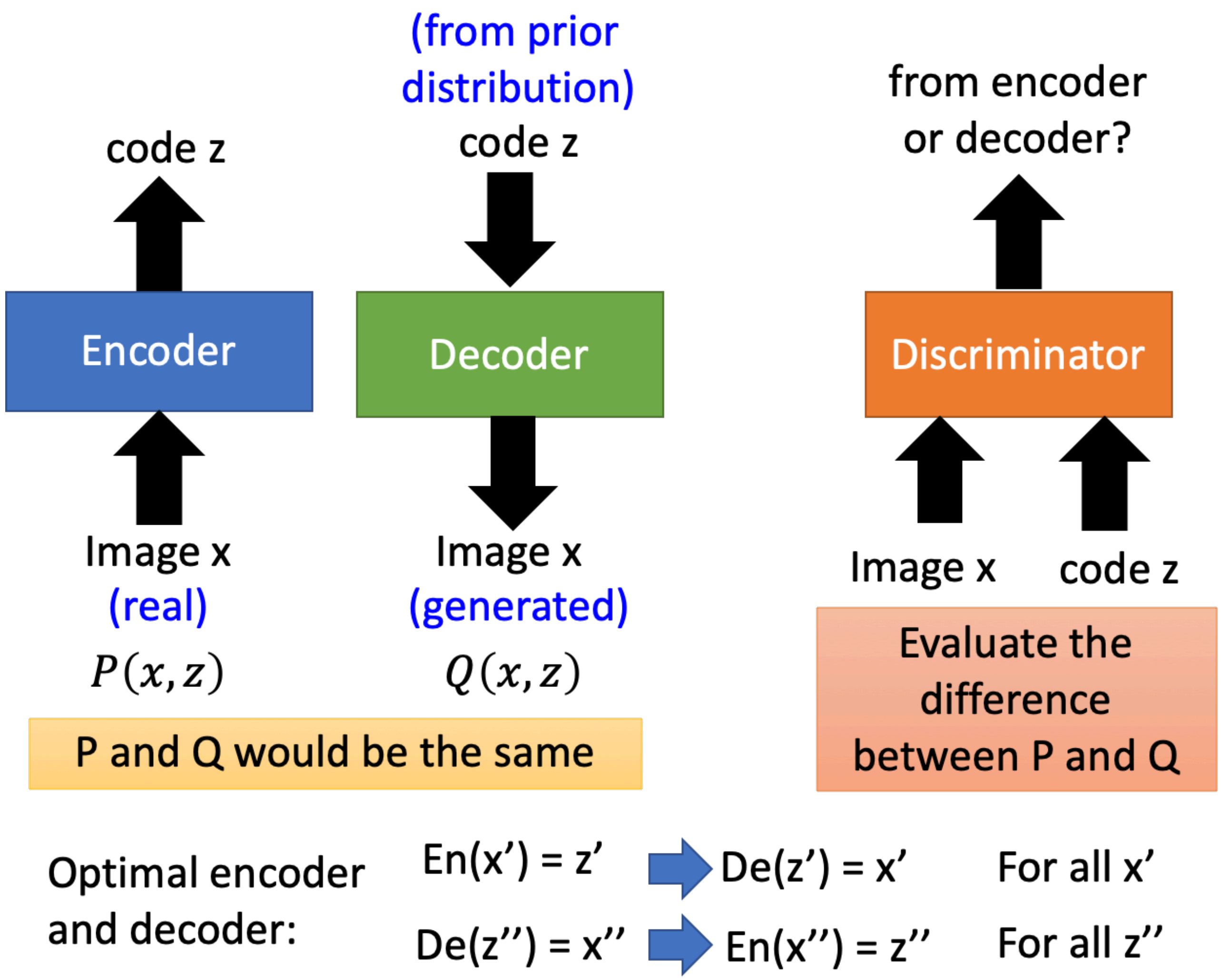
Another kind of discriminator:



BiGan



- Initialize encoder En, decoder De, discriminator Dis
- In each iteration:
  - Sample M images $x^1, x^2, \cdots, x^M$ from database
  - Generate M codes $\tilde{z}^1, \tilde{z}^2, \cdots, \tilde{z}^M$ from encoder
    - $\tilde{z}^i = En(x^i)$
  - Sample M codes $z^1, z^2, \cdots, z^M$ from prior P(z)
  - Generate M codes $\tilde{x}^1, \tilde{x}^2, \cdots, \tilde{x}^M$ from decoder
    - $\tilde{x}^i = De(z^i)$
  - Update Dis to increase $Dis(x^i, \tilde{z}^i)$, decrease $Dis(\tilde{x}^i, z^i)$
  - Update En and De to decrease $Dis(x^i, \tilde{z}^i)$, increase $Dis(\tilde{x}^i, z^i)$

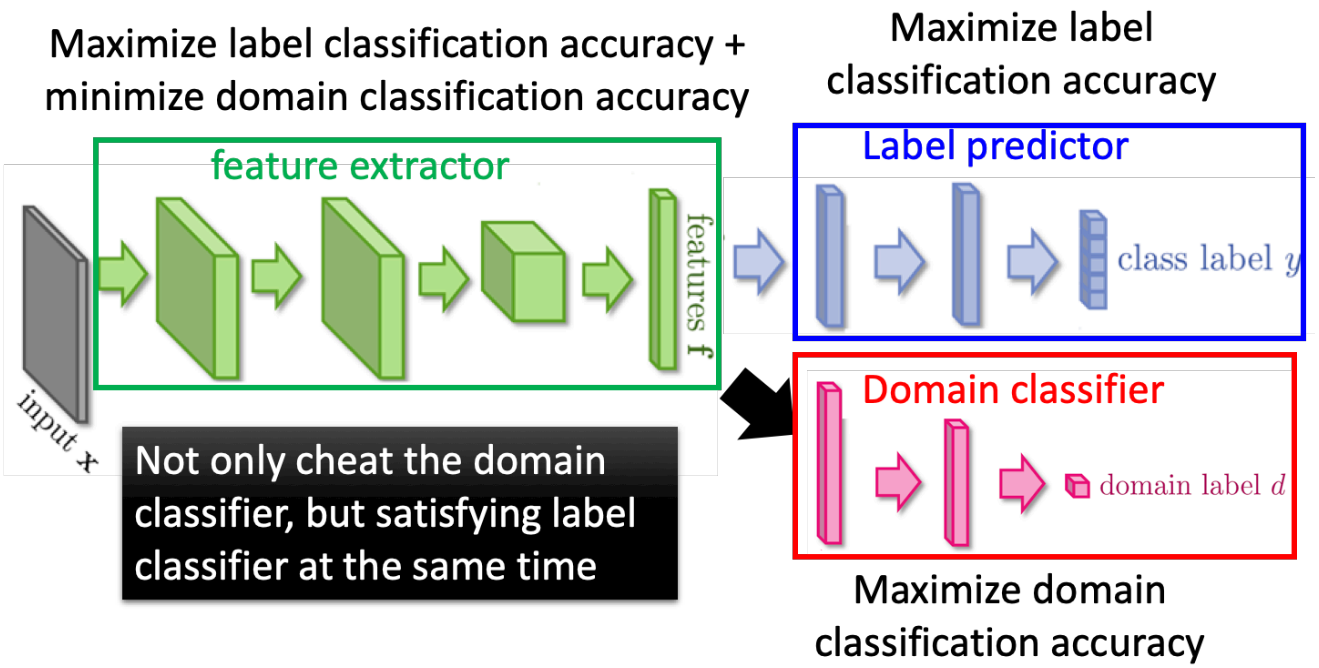


让encoder与decode越相似越好

Bigan得到的auto encoder与一半的auto-encoder特性不一样

Triple Gan

Domain-adversarial training



This is a big network, but different parts have different goals.