

# 编码的应用--VAE

报告人:陈扬

slides制作:陈扬

代码:陈扬

github:[https://github.com/OUCMachineLearning/OUCLML/tree/master/One Day One GAN/day15](https://github.com/OUCMachineLearning/OUCLML/tree/master/One%20Day%20One%20GAN/day15)

# 什么是编码？

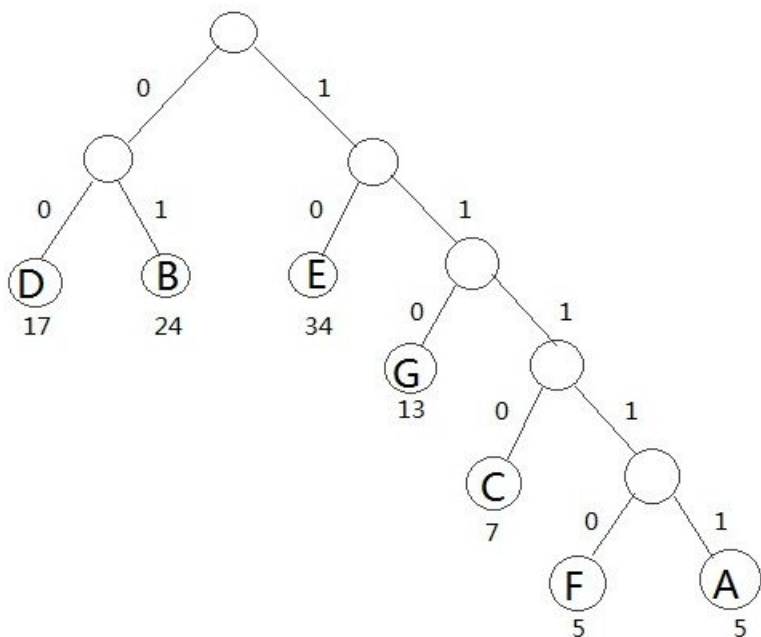
编码是信息从一种形式或格式转换为另一种形式的过程。解码，是编码的逆过程。

对于特定的上下文，编码有一些更具体的意义。

- 编码（Encoding）在认知上是解释传入的刺激的一种基本知觉的过程。技术上来说，这是一个复杂的、多阶段的转换过程，从较为客观的感觉输入（例如光、声）到主观上有意义的体验。

# 本质:数据降维

在之前,老师给我们讲解了霍夫曼编码的应用,使用**变长编码表**对源符号（如文件中的一个字母）进行编码，其中**变长编码表**是通过一种评估来源符号出现机率的方法得到的，出现机率高的字母使用较短的编码，反之出现机率低的则使用较长的编码，这便使编码之后的字符串的平均长度、**期望值**降低，从而达到**无损压缩**数据的目的。



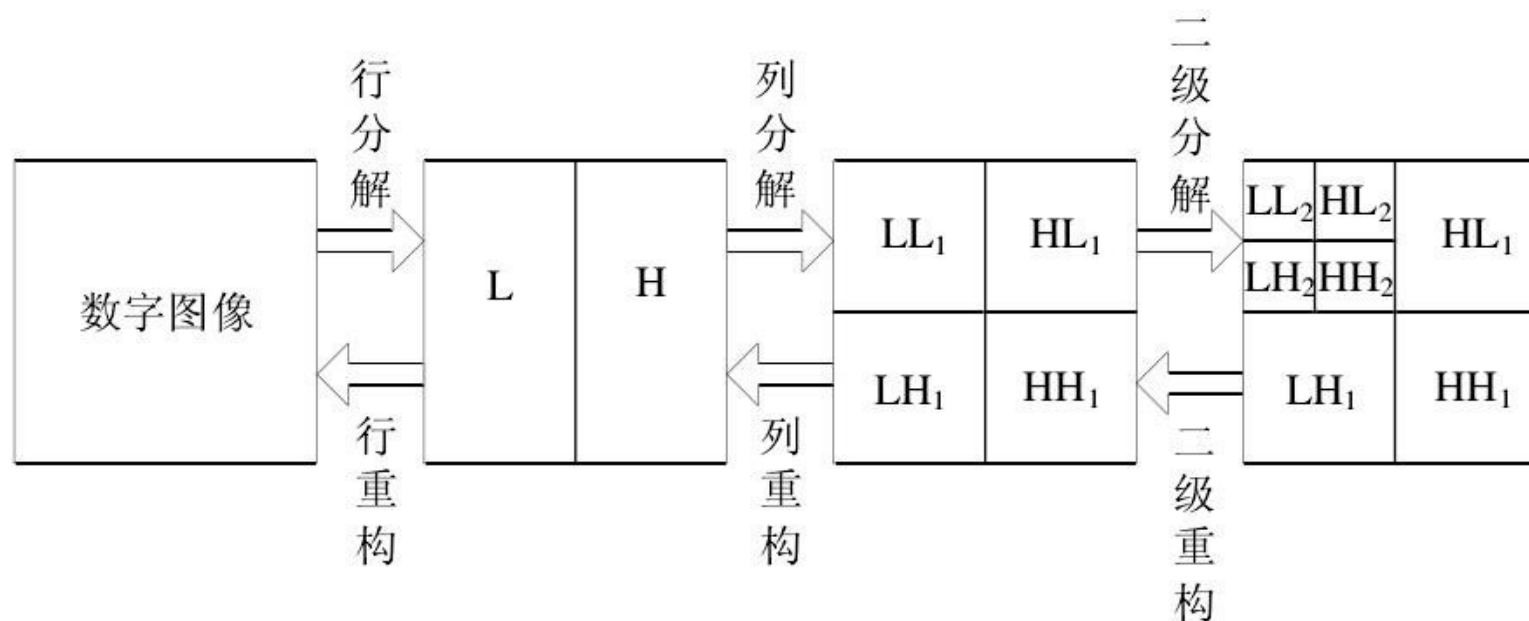
A: 1111  
B: 01  
C: 1110  
D: 00  
E: 0  
F: 11110  
G: 110

可是,我们日常生活中,视觉信息>>听觉信息

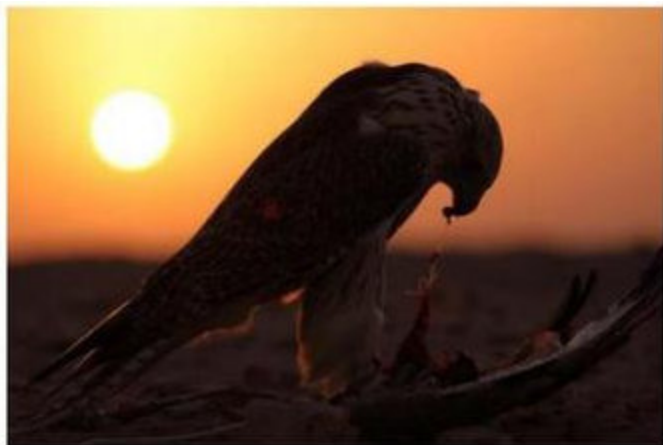
## 图像压缩的传统方法:

Gif, PNG, JPG,TIF,TIFF,WEBP.....

DCT 离散小波变换



换个思路,如果我们能够抽离出图像的更高维信息,比如将图像表示成语义信息?



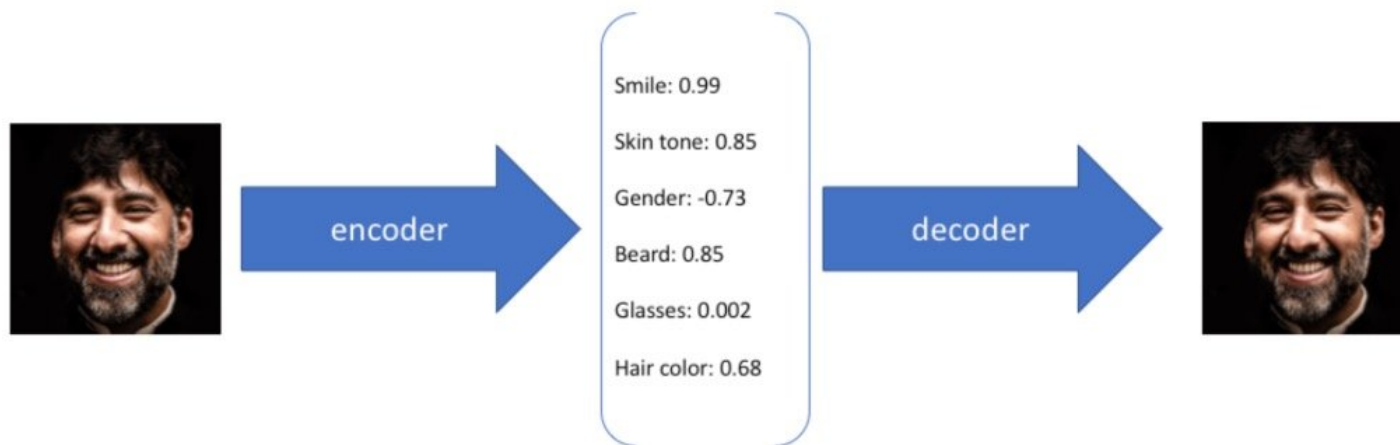
描述:

一只猎鹰在日落时进食。猎鹰站在地上。

诗:

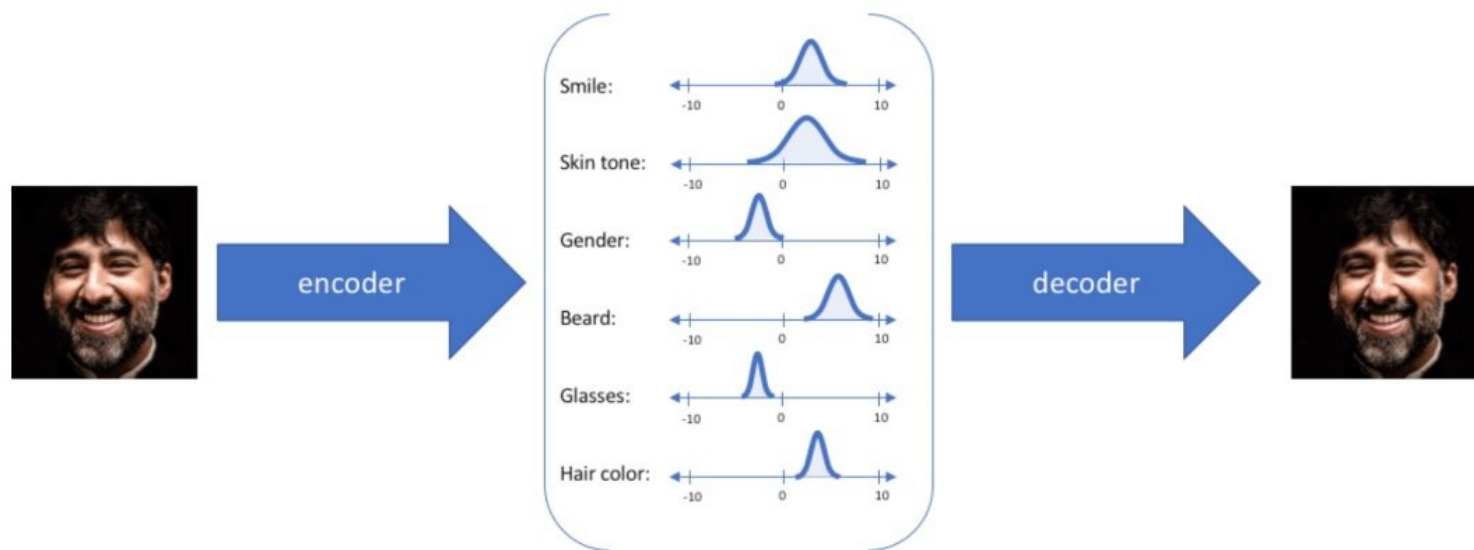
动如黑夜猎鹰  
猎似黑暗骑士  
蓄势待发  
全力以赴

# 有了编码,我们再加上解码呢?



Latent attributes

[https://blog.csdn.net/weixin\\_40955254](https://blog.csdn.net/weixin_40955254)

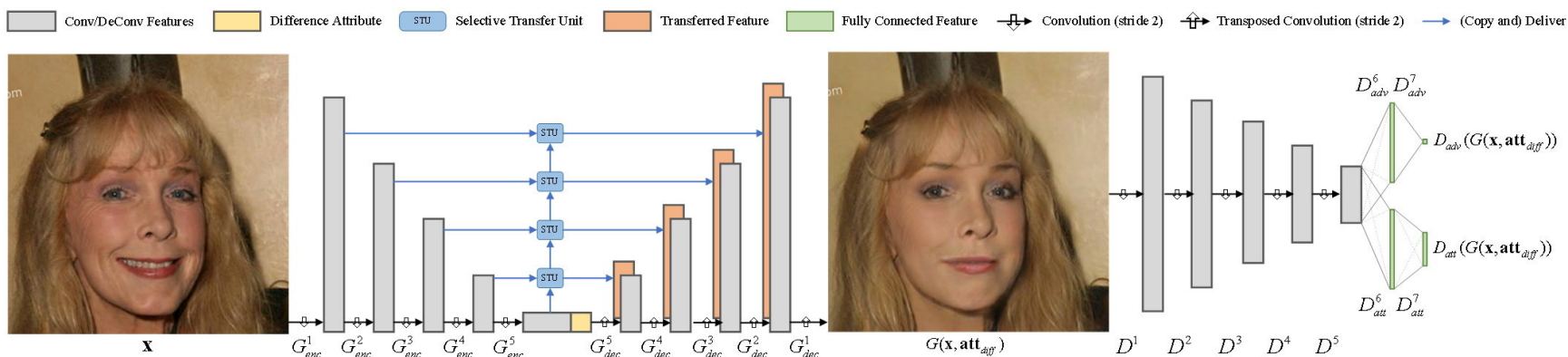


Latent attributes

[https://blog.csdn.net/weixin\\_40955254](https://blog.csdn.net/weixin_40955254)

# 前沿技术的应用

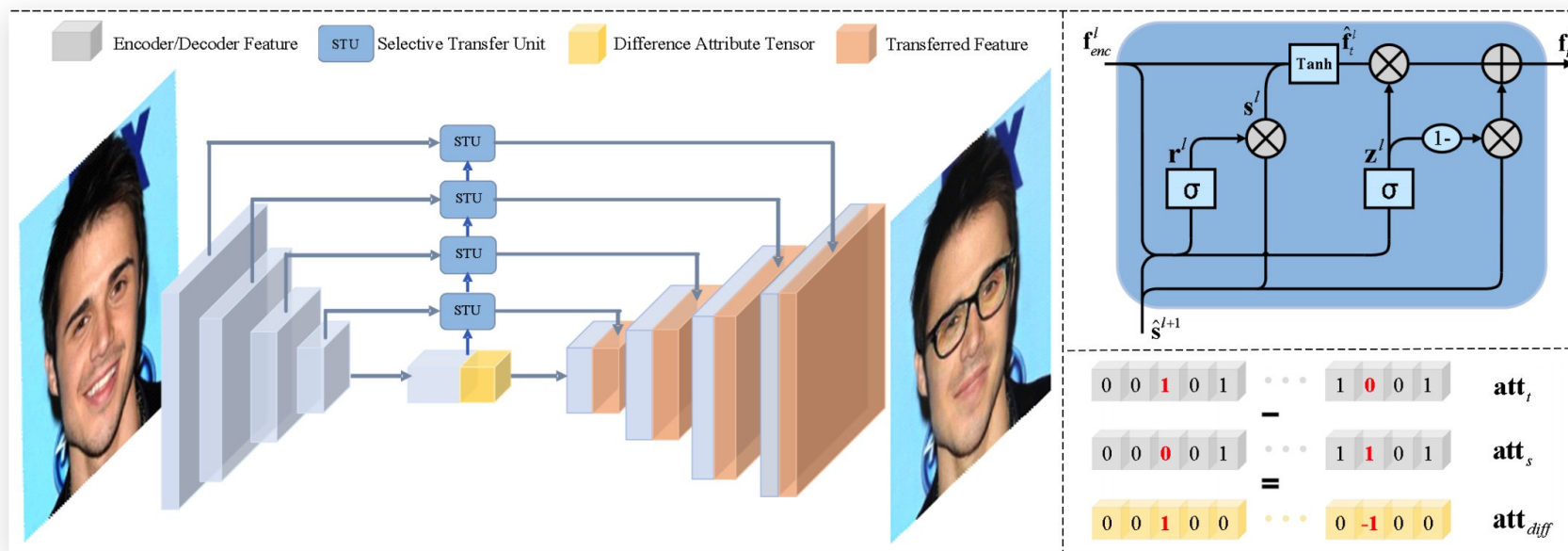
## STGAN (CVPR 2019)



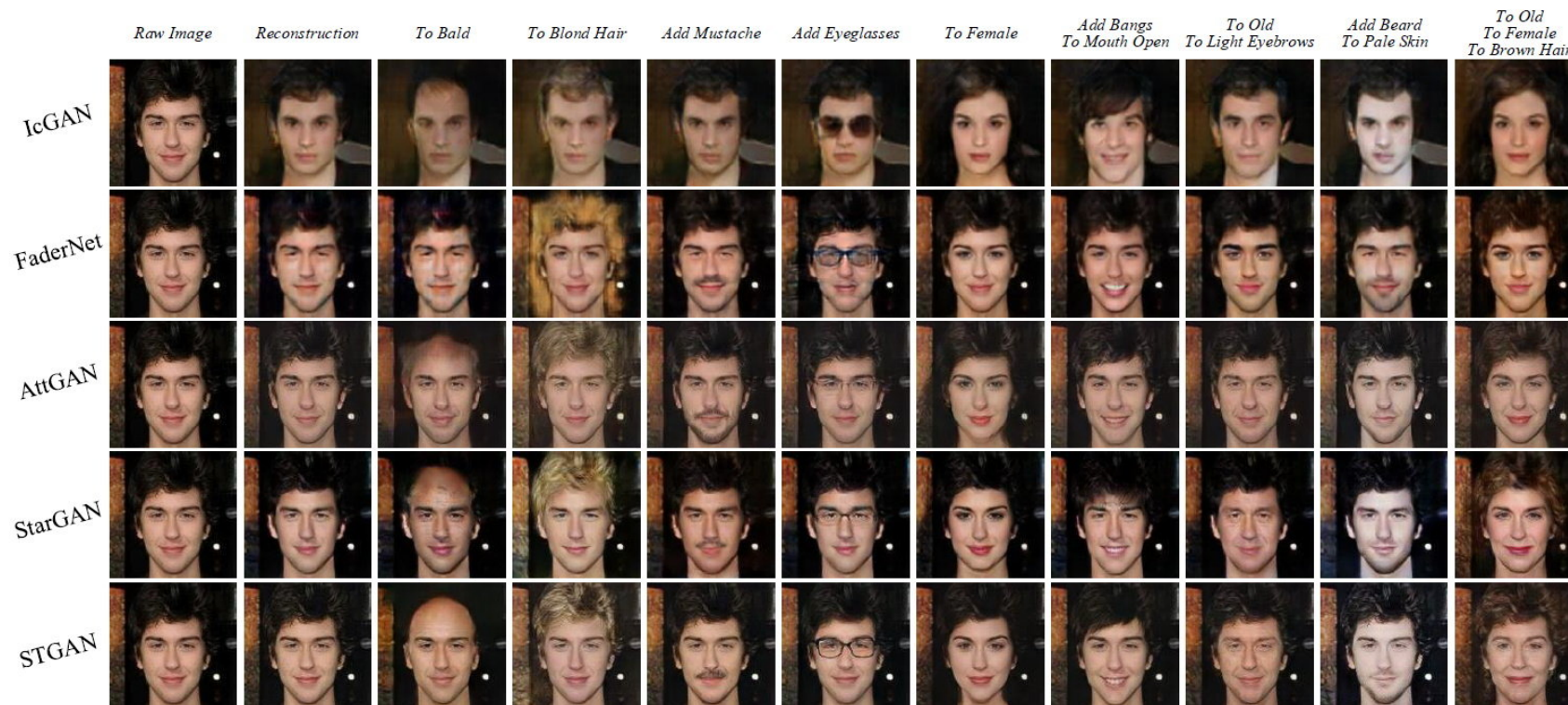
Overall architecture of our STGAN. Taking the image above as an example, in the difference attribute vector  $att_{diff}$ , *Young* is set to 1, *Mouth Open* is set to -1, and others are set to zeros. The outputs of  $D_{att}$  and  $D_{adv}$  are the scalar  $D_{adv}(G(x, att_{diff}))$  and the vector  $D_{att}(G(x, att_{diff}))$ , respectively



假如我们能够把图像的高维信息提取出来表层上可以被人们予以理解的信息的话,通过这个 ENcode--DEcode 的结构,我们就可以实现图像翻译的工作.



## 类似的工作:



感兴趣:<http://www.twistedwg.com/2019/04/29/STGAN.html>

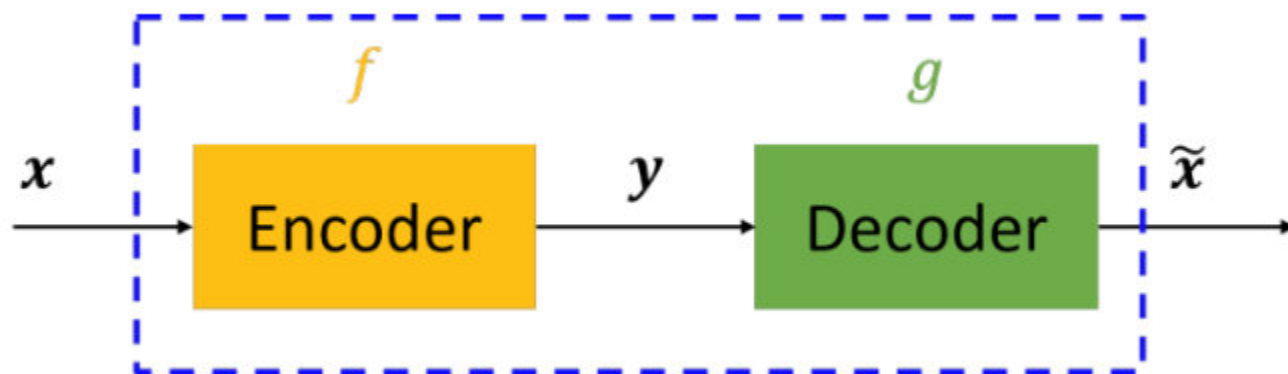
## 我们要做的 demo

可是我并不打算在这里讲解或者说演示怎么复杂的 demo,我为大家准备了一个最简单的也是其中最核心的原理部分—**自编码器**

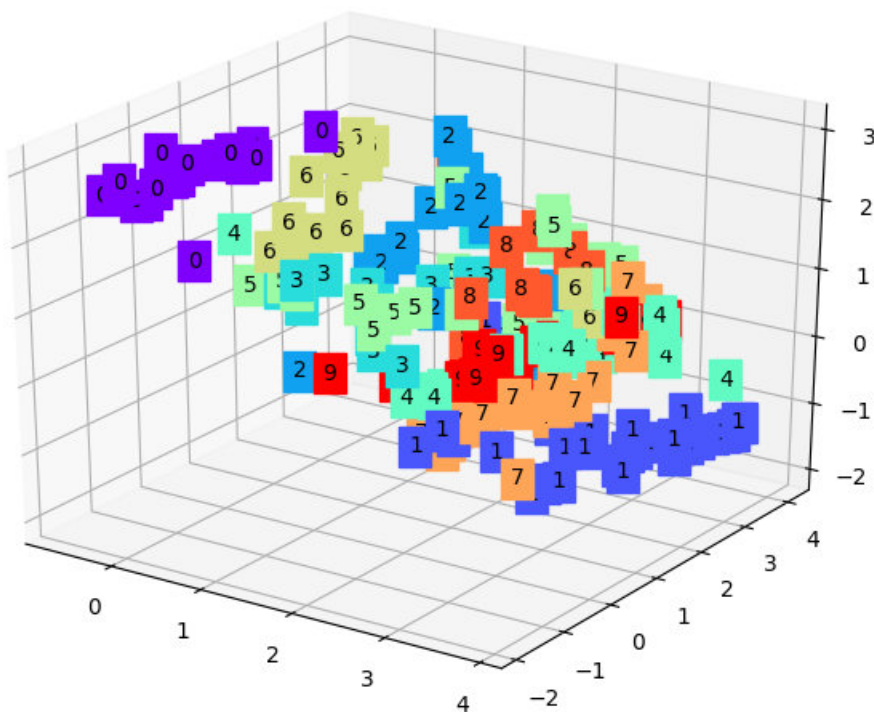
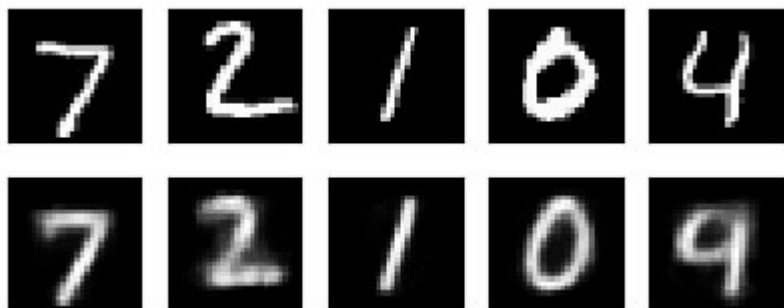
# VAE:

变分自动编码器，而其中利用到的自编码器(autoencoder)是机器学习的一个基础知识。为了更好的理解VAE先要把autoencoder部分又进行一下消化。

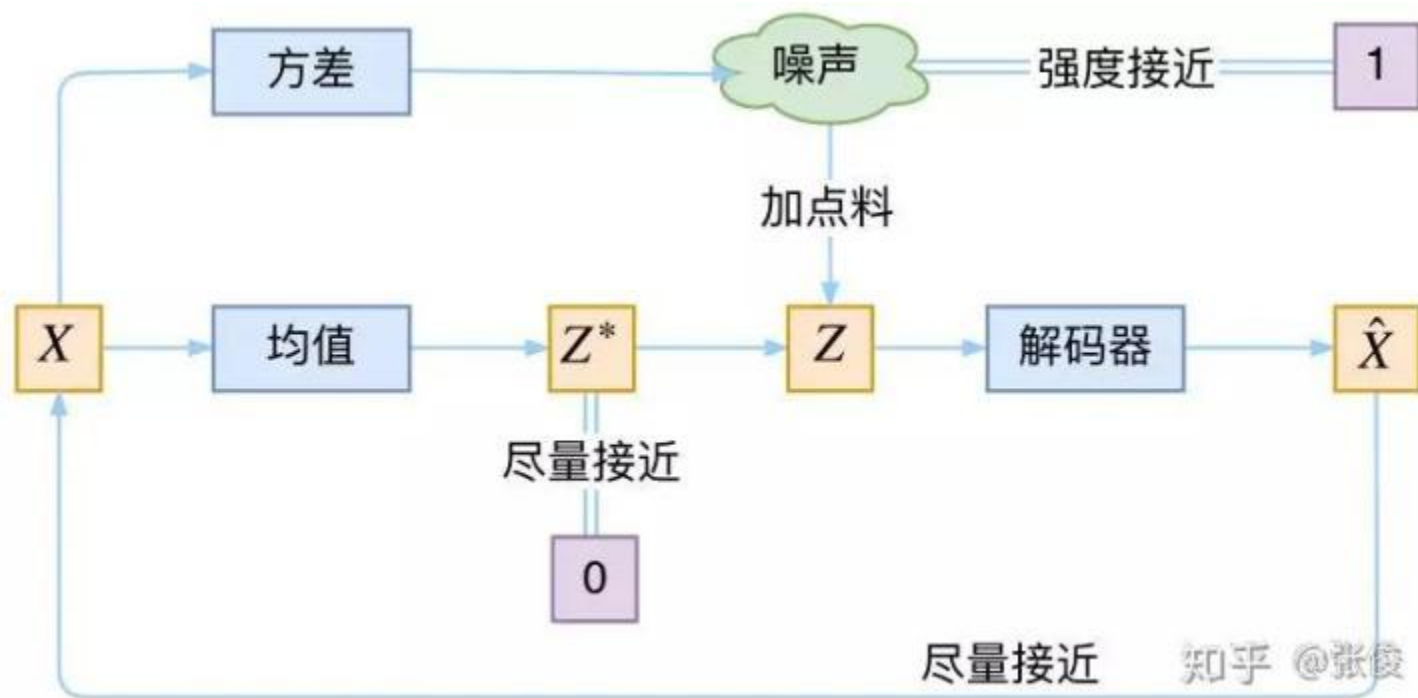
自编码器是通过对输入 $x$ 进行编码后得到一个低维的向量 $y$ ，然后根据这个向量还原出输入 $x$ ，。通过对比 $x$ 与 $\tilde{x}$ 的误差，再利用神经网络去训练使得误差逐渐减小从而达到非监督学习的目的，图1展示autoencoder过程。



举个例子:



详情:<https://zhuanlan.zhihu.com/p/34998569>



$$p(Z) = \sum_X p(Z|X)p(X) = \sum_X N(0, 1)p(X) = N(0, 1) \sum_X p(X) = N(0, 1)$$

$$L(\theta, \Phi; x^{(i)}) = -KL(q_{\Phi}(z|x^{(i)})||p_{\theta}(z)) + E_{q_{\Phi}(z|x)}[\log p_{\theta}(x^{(i)}|z)]$$

$p(Z)$ 是标准正态分布，这就意味着，隐参量符合了正态分布，这样就可以放心地从这个分布中随机采样，保证了生成能力。

**那么打住!到此为止我不打算再进一步深入的来讲了,我们来看演示的 DEMO**

# THANK YOU

## Q&A