AAE + pix2pix

**Gapeng-北京大学-CV 2017-08-23 20:00:29**

今天是论文讨论，主要讨论两个论文：AAE和pix2pix。

**Gapeng-北京大学-CV 2017-08-23 20:00:38**

我们先讨论第一篇AAE: Adversarial Autoencoders(1511.05644)。毛豆大佬写过一篇专栏，把它跟PM，GAN，InfoGAN做过对比，讨论结束以后大家可以去看看https://zhuanlan.zhihu.com/p/27159510，写得很赞。

**Gapeng-北京大学-CV 2017-08-23 20:00:52**

"AAE我们今天主要围绕以下四个问题展开讨论：  
1. 它的loss和网络是怎样的？  
2. 它和AE有什么不同？  
3. 它和GAN又有什么不同？AAE在编码空间进行对抗，有什么好处？要解决什么问题？  
4. 跟GAN相比，它不太适合图像，但是更加适合文本，对吗？"

**kk+国科大+NLP 2017-08-23 20:00:56**

好

**Gapeng-北京大学-CV 2017-08-23 20:01:10**

第一个问题，AAE的loss和网络是怎样的？

**Gapeng-北京大学-CV 2017-08-23 20:01:19**

这个问题看过文章应该都能回答

**kk+国科大+NLP 2017-08-23 20:02:03**

网络结构就是

**kk+国科大+NLP 2017-08-23 20:02:16**

分成两大部分网络結构：自编码部分（图上半部分）、GAN判别网络(图下半部分)

**kk+国科大+NLP 2017-08-23 20:02:29**

是吗？

**Gapeng-北京大学-CV 2017-08-23 20:03:01**

嗯，网络是这样，那它跟GAN有什么区别？

**kk+国科大+NLP 2017-08-23 20:03:22**

等一下，我翻翻笔记

**ccm-GZ-ML 2017-08-23 20:03:56**

和autoencoder ＋ GAN 单独训练有什么区别？

**Gapeng-北京大学-CV 2017-08-23 20:03:59**

嘿嘿，你还做了笔记

**kk+国科大+NLP 2017-08-23 20:04:27**

GAN被用于在神经网络的输出层上的像素级上强加数据分布。AAE依靠AE训练来捕获数据分布;

**kk+国科大+NLP 2017-08-23 20:04:45**

小白，看论文少

**陆鹏起-hust-gan 2017-08-23 20:04:53**

gan可以看作是分类模型，aae可以看做回归模型

**kk+国科大+NLP 2017-08-23 20:04:59**

只能回答些浅显的[捂脸]

**陆鹏起-hust-gan 2017-08-23 20:05:01**

不知道我这样理解对不对

**kk+国科大+NLP 2017-08-23 20:05:51**

数据分布一个拟合。一个ae编码

**kk+国科大+NLP 2017-08-23 20:05:59**

感觉是最大区别

**kk+国科大+NLP 2017-08-23 20:06:34**

不知道理解对不对

**小虎牙-阿里-相关性 2017-08-23 20:07:03**

z是一个向量？

**Gapeng-北京大学-CV 2017-08-23 20:07:30**

AAE跟GAN其实是相反的一个过程

**kk+国科大+NLP 2017-08-23 20:07:41**

嗯，z是编码向量

**kk+国科大+NLP 2017-08-23 20:08:16**

怎么理解？

**Gapeng-北京大学-CV 2017-08-23 20:08:21**

AAE的encoder是从x -> z，再重构x，而GAN是z -> x

**小虎牙-阿里-相关性 2017-08-23 20:08:45**

嗯 看网络结构是这样子的

**kk+国科大+NLP 2017-08-23 20:08:51**

对

**kk+国科大+NLP 2017-08-23 20:09:02**

但这z含义有点不同

**小虎牙-阿里-相关性 2017-08-23 20:10:19**

D在低位隐空间比较

**kk+国科大+NLP 2017-08-23 20:10:20**

我感觉是不是有种gan中的g加了ae

**安捷 北京大学 深度学习 2017-08-23 20:10:20**

z分别有什么含义呀

**小虎牙-阿里-相关性 2017-08-23 20:10:22**

第一次看这个网络...

**Gapeng-北京大学-CV 2017-08-23 20:10:44**

"z的含义其实没有太大区别吧，AAE的encoder是对p(z|x)建模，GAN的G是对p(x|z)建模"

**kk+国科大+NLP 2017-08-23 20:10:50**

感觉aae是降维下来的，应该含有数据的特征

**kk+国科大+NLP 2017-08-23 20:11:29**

嗯

**Gapeng-北京大学-CV 2017-08-23 20:11:35**

它们对抗的空间不一样

**安捷 北京大学 深度学习 2017-08-23 20:12:00**

我觉得aae除了降维或许也是为了编码

**Gapeng-北京大学-CV 2017-08-23 20:12:02**

一个在隐空间对抗，一个在原始的数据空间对抗

**kk+国科大+NLP 2017-08-23 20:12:22**

嗯，确实这样理解好像也是一种思路

**ccm-GZ-ML 2017-08-23 20:12:24**

有道理

**kk+国科大+NLP 2017-08-23 20:12:28**

受教了

**小虎牙-阿里-相关性 2017-08-23 20:12:40**

隐空间优势？

**Gapeng-北京大学-CV 2017-08-23 20:12:48**

BiGAN也可以做编码，它是双向的

**kk+国科大+NLP 2017-08-23 20:12:50**

有道理，确实感觉还会回去

**Gapeng-北京大学-CV 2017-08-23 20:13:22**

AAE的loss呢？

**黄瑞阳\_郑州大学\_NLP 2017-08-23 20:13:25**

网上看到的，说AAEDecoder是以重构误差为目的进行训练，而非GAN那样以欺骗判别器为目的

**kk+国科大+NLP 2017-08-23 20:13:47**

aae有两个目的

**小虎牙-阿里-相关性 2017-08-23 20:14:00**

隐空间优势？谁说一下

**Gapeng-北京大学-CV 2017-08-23 20:14:14**

如果只有重构误差的话，学到的图像不够sharp

**张佳婕-暨南大学-DL 2017-08-23 20:14:14**

重构误差怎么理解？

**阿来-THU-KBQA 2017-08-23 20:14:46**

是可以理解为gan是从傻白甜的高斯分布变成有用的正确数据，而aae是从正确数据把自己转换成傻白甜？

**kk+国科大+NLP 2017-08-23 20:14:58**

一个是重构误差一个是对抗

**安捷 北京大学 深度学习 2017-08-23 20:15:10**

谁能给我解释下p哪里来的

**潘少-东南大学-干 2017-08-23 20:15:28**

莱昂纳多的解释很有趣

**潘少-东南大学-干 2017-08-23 20:15:34**

😄

**Gapeng-北京大学-CV 2017-08-23 20:16:00**

最喜欢这样的解释了

**Gapeng-北京大学-CV 2017-08-23 20:16:32**

重构误差就是重构误差[捂脸]里面有个autoencoder嘛@张佳婕-暨南大学-DL

**潘少-东南大学-干 2017-08-23 20:16:32**

这句话一定要写进总结文里

**阿来-THU-KBQA 2017-08-23 20:16:37**

我是个完全的外行，来虚心学习的。。。

**kk+国科大+NLP 2017-08-23 20:16:42**

"x是样本输入，z是编码向量，p(z)是隐含变量的先验分布， q(z|x)是編码分布函数，p(x|z)是解码分布函数，pd(x)表示真实的数据分布，p(x)表示模型数椐分布。"

**ccm-GZ-ML 2017-08-23 20:16:58**

这笔记做的...

**kk+国科大+NLP 2017-08-23 20:17:16**

我先给一下变量的含义

**兔子-不存在-ML 2017-08-23 20:17:27**

把正确的变成傻白甜那个是vae吧

**Gapeng-北京大学-CV 2017-08-23 20:17:30**

哈哈

**kk+国科大+NLP 2017-08-23 20:17:33**

不然我就跟不上讨论，手机打字好慢

**Gapeng-北京大学-CV 2017-08-23 20:18:22**

vae跟AAE是有相似支持的

**Gapeng-北京大学-CV 2017-08-23 20:18:57**

两者都有一个过程（encoder）把正确的数据变成傻白甜

**阿来-THU-KBQA 2017-08-23 20:19:26**

怎么训练呢

**Gapeng-北京大学-CV 2017-08-23 20:19:38**

loss呢，谁来说一下AAE的loss是怎样的？

**兔子-不存在-ML 2017-08-23 20:19:46**

但aae的目的不只是为了变成傻白甜呀，如果只需要傻白甜应该不需要引入对抗吧

**kk+国科大+NLP 2017-08-23 20:19:55**

所以其实傻白甜才是核心

**Gapeng-北京大学-CV 2017-08-23 20:20:12**

从网络可以看出分成三个部分：encoder, decoder和D，三者的loss分别是怎样的？

**安捷 北京大学 深度学习 2017-08-23 20:20:49**

encode和decode是有单独的loss？[捂脸]

**Gapeng-北京大学-CV 2017-08-23 20:21:35**

encoder跟D有关，decoder可以说跟D没有直接关系吧

**kk+国科大+NLP 2017-08-23 20:21:46**

应该是算法分两个阶段步了

**ccm-GZ-ML 2017-08-23 20:22:03**

decoder就是一般的重构误差吧 mse

**Gapeng-北京大学-CV 2017-08-23 20:22:09**

先说loss

**Gapeng-北京大学-CV 2017-08-23 20:22:14**

再说训练

**张佳婕-暨南大学-DL 2017-08-23 20:22:20**

😭跟不上了 嘤嘤嘤

**Gapeng-北京大学-CV 2017-08-23 20:22:35**

对，decoder的loss就是mse

**Gapeng-北京大学-CV 2017-08-23 20:22:46**

encoder呢？D呢？

**黄瑞阳\_郑州大学\_NLP 2017-08-23 20:22:57**

在看一篇AAE的笔记，写得不错。http://blog.csdn.net/hjimce/article/details/54411244

**Gapeng-北京大学-CV 2017-08-23 20:24:01**

D比较简单，还是判别真假，只是它是在隐空间上判别

**Gapeng-北京大学-CV 2017-08-23 20:24:22**

encoder的loss比较关键，它是怎样的？

**Gapeng-北京大学-CV 2017-08-23 20:24:47**

我们来分析一下encoder的loss有几项？

**wwwyn\_scut\_nlp 2017-08-23 20:25:18**

既有重构loss又有判别器的loss

**Gapeng-北京大学-CV 2017-08-23 20:26:13**

嗯嗯

**kk+国科大+NLP 2017-08-23 20:26:16**

是不是两项

**anshiquanshu66-师大-医学图像处理 2017-08-23 20:26:26**

默默的在聆听

**Gapeng-北京大学-CV 2017-08-23 20:26:34**

是两项

**kk+国科大+NLP 2017-08-23 20:27:01**

从d传回的还有decoder传回的

**Gapeng-北京大学-CV 2017-08-23 20:27:15**

是的

**Gapeng-北京大学-CV 2017-08-23 20:27:18**

好，那AAE是怎么训练的？

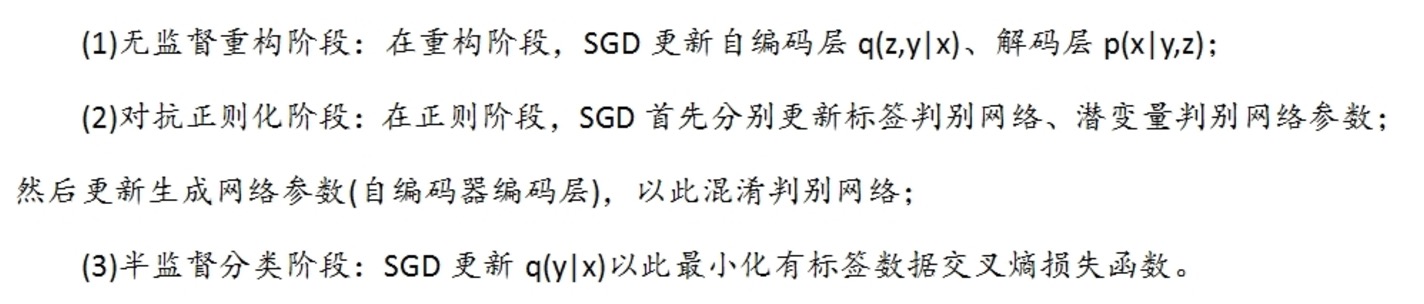
**kk+国科大+NLP 2017-08-23 20:29:28**

这训练是不是一次迭代器要先decoder 训练一下，再训练一个gan

**Gapeng-北京大学-CV 2017-08-23 20:29:34**

要注意总共有三个部分，encoder，decoder和D

**hwang-中南大学-生成模型 2017-08-23 20:30:09**



**Gapeng-北京大学-CV 2017-08-23 20:30:14**

文章说的是分成两个阶段训练，第一阶段训练autoencoder，第二阶段训练GAN

**hwang-中南大学-生成模型 2017-08-23 20:30:20**

群主大大有何看法

**kk+国科大+NLP 2017-08-23 20:30:43**

👍

**安捷 北京大学 深度学习 2017-08-23 20:31:11**

第二阶段能详细说一下吗

**Gapeng-北京大学-CV 2017-08-23 20:31:14**

他说的很对[捂脸]

**安捷 北京大学 深度学习 2017-08-23 20:31:17**

[可怜]

**kk+国科大+NLP 2017-08-23 20:31:20**

瞬间感觉这网络结构很复杂呀

**Gapeng-北京大学-CV 2017-08-23 20:32:07**

第二阶段就是GAN正常训练的阶段，encoder就是GAN的G，D就是GAN里面的D，按照GAN的更新方式去训练就行了

**Gapeng-北京大学-CV 2017-08-23 20:33:21**

上面那张图片是AAE做半监督学习的训练过程，会涉及到其他的loss项，所以看上去复杂一点

**ccm-GZ-ML 2017-08-23 20:33:33**

"\* reconstruction phase  
 1. update encoder and decoder to minimize reconstruction error  
\* regularization phase  
 1. update discriminative network  
 2. update generator(encoder)"

**ccm-GZ-ML 2017-08-23 20:33:41**

paper里是两阶段

**安捷 北京大学 深度学习 2017-08-23 20:33:50**

懂了

**kk+国科大+NLP 2017-08-23 20:33:56**

嗯

**kk+国科大+NLP 2017-08-23 20:34:16**

这模型还是好理解的

**安捷 北京大学 深度学习 2017-08-23 20:34:18**

encoder每一轮训练要被更新两次？

**hwang-中南大学-生成模型 2017-08-23 20:34:36**

感觉切实滴跑一遍就能懂得差不多了

**kk+国科大+NLP 2017-08-23 20:34:47**

我觉得至少两次

**安捷 北京大学 深度学习 2017-08-23 20:35:11**

那我问个问题哈

**kk+国科大+NLP 2017-08-23 20:35:23**

毕竟gan的d可以跑多次吧，

**安捷 北京大学 深度学习 2017-08-23 20:35:35**

aae到底是vae上面用了gan还是gan上面用了vae呀

**starif-西电-ml 2017-08-23 20:35:48**

@hwang-中南大学-生成模型 那个是aae用于半监督，有两个对抗，他们现在在讨论监督学习的

**hwang-中南大学-生成模型 2017-08-23 20:36:10**

嗯嗯[捂脸]

**kk+国科大+NLP 2017-08-23 20:36:17**

我感觉是gan的g用了ae

**Gapeng-北京大学-CV 2017-08-23 20:36:42**

跟vae有一定联系，但是没有太多关系

**ccm-GZ-ML 2017-08-23 20:36:58**

aae和vae有啥区别呀

**starif-西电-ml 2017-08-23 20:37:07**

感觉是aae上加了个gan来对z进行规范

**丁铭 清华 数据挖掘 2017-08-23 20:37:12**

为啥我觉得跟vae关系更密切

**Lynn-中科院计算所-CV 2017-08-23 20:37:22**

@Gapeng-北京大学-CV 为何没有太多关系？

**Gapeng-北京大学-CV 2017-08-23 20:37:24**

优化目标不一样

**Lynn-中科院计算所-CV 2017-08-23 20:37:34**

我也认为关系很密切

**kk+国科大+NLP 2017-08-23 20:37:51**

感觉加了vae有降维，提主要特征的意思

**安捷 北京大学 深度学习 2017-08-23 20:38:19**

那用ae的目的是什么呢？咱们刚才说到了降维，还有什么？是不是跟aae的问题场景有关系？vae和aae分别用来做什么？

**kk+国科大+NLP 2017-08-23 20:38:20**

嗯，怎么理解？

**Gapeng-北京大学-CV 2017-08-23 20:38:52**

两者都有后验概率逼近先验概率的想法，实现的手段不一样

**Lynn-中科院计算所-CV 2017-08-23 20:39:47**

手段不同，但感觉本质上很像

**安捷 北京大学 深度学习 2017-08-23 20:40:10**

vae是用来做nlp编码和降维的吧？我有没有理解错啊[快哭了]

**kk+国科大+NLP 2017-08-23 20:40:15**

我个人感觉是ae本身就有点像svd感觉，留下主要因素

**Gapeng-北京大学-CV 2017-08-23 20:40:25**

嗯，本质上是很像，以至于我一开始就想把D拿掉😂

**Gapeng-北京大学-CV 2017-08-23 20:41:03**

vae也是个生成模型啊

**Lynn-中科院计算所-CV 2017-08-23 20:41:12**

"有点不太一样的就是p(z|x)在vae里面是个分布，在aae里面建模成确定性函数了"

**Gapeng-北京大学-CV 2017-08-23 20:41:52**

它也可以引入随机性

**hwang-中南大学-生成模型 2017-08-23 20:41:54**

可不可以理解成，就是用对抗的方式学习图像的编码的分布吧？

**Gapeng-北京大学-CV 2017-08-23 20:42:17**

嗯，可以这么理解

**安捷 北京大学 深度学习 2017-08-23 20:42:23**

@hwang-中南大学-生成模型 我暂时也是这么理解的

**kk+国科大+NLP 2017-08-23 20:42:31**

"1.VAE使用KL惩罚对自动编码器的hidden code vector施加先验分布，AAE使用对抗训练步将hidden code vector的聚合后验与先验分布匹配；  
2.在VAE中，为了通过蒙特卡洛取样通过相对熵反向传播，就需要先验分布的确切函数形式。 然而，在AAE中，只需要从先验分布中抽样，引导q（z）匹配p（z）即可"

**Lynn-中科院计算所-CV 2017-08-23 20:42:40**

aae怎么引入随机性？引入之后怎么学习，还是reparametirc trick?

**kk+国科大+NLP 2017-08-23 20:42:42**

这两条是原文写的

**Gapeng-北京大学-CV 2017-08-23 20:42:47**

那就有一个问题，为什么要在编码空间进行对抗，有什么好处吗？

**Gapeng-北京大学-CV 2017-08-23 20:43:34**

跟VAE类似，还是得用reparametrization trick@Lynn-中科院计算所-CV

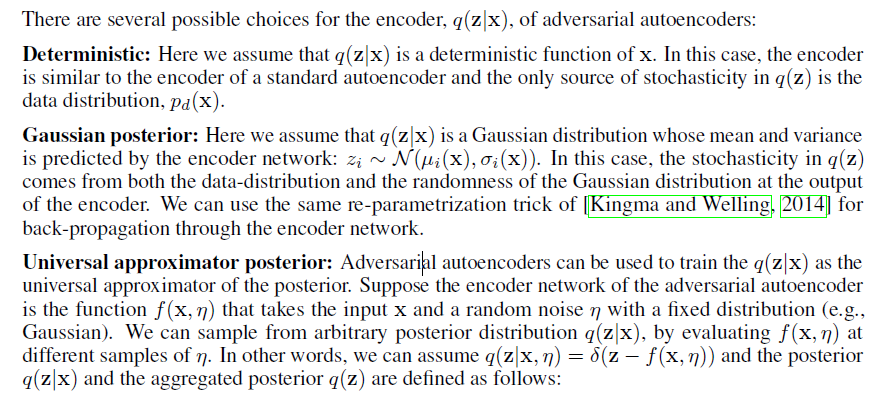
**安捷 北京大学 深度学习 2017-08-23 20:43:40**

维数会更低？

**安捷 北京大学 深度学习 2017-08-23 20:43:49**

更好训练d？

**Gapeng-北京大学-CV 2017-08-23 20:43:51**



**Lynn-中科院计算所-CV 2017-08-23 20:44:58**

那这样说的话aae的功能应该是覆盖vae的

**kk+国科大+NLP 2017-08-23 20:44:58**

隐空间优势

**Gapeng-北京大学-CV 2017-08-23 20:44:59**

但是D太强了是不是encoder就梯度消失了？

**hwang-中南大学-生成模型 2017-08-23 20:45:06**

在编码空间进行对抗就可以把解码器当作一个比较好的生成器来用了，而一般的生成器通常比较弱

**Gapeng-北京大学-CV 2017-08-23 20:45:09**

隐空间有什么优势？

**shawn-网易游戏-GAN&RL 2017-08-23 20:45:21**

D应该本来就挺好训练的

**kk+国科大+NLP 2017-08-23 20:46:35**

我还是感觉他自带特征提取，聚类的那种效果

**安捷 北京大学 深度学习 2017-08-23 20:46:42**

意思是说编码器的存在实际上让解码器能比gan中的g更好的生成？

**kk+国科大+NLP 2017-08-23 20:46:52**

使得噪声减少

**hwang-中南大学-生成模型 2017-08-23 20:47:23**

感觉上应该还是有这方面的优势的

**Gapeng-北京大学-CV 2017-08-23 20:47:25**

嗯，这算一个

**Gapeng-北京大学-CV 2017-08-23 20:47:42**

嗯，其他方面的优势呢？

**hwang-中南大学-生成模型 2017-08-23 20:47:45**

把自编码器完全训练好，然后直接学习所有图像的编码分布，应该也不会太差

**郑华滨-中山大学-深度学习 2017-08-23 20:47:47**

在图像上其实应该反而是没有好处的，因为AE的mse loss会使得图像模糊，但是AAE这个架构在文本上就有很大优势了，巧妙绕开了文本离散不可微的问题

**Gapeng-北京大学-CV 2017-08-23 20:47:54**

毛豆大佬曾经说过哇

**安捷 北京大学 深度学习 2017-08-23 20:48:05**

赞

**kk+国科大+NLP 2017-08-23 20:48:23**

👍

**Gapeng-北京大学-CV 2017-08-23 20:48:46**

等了好久，毛豆大佬终于出来了

**安捷 北京大学 深度学习 2017-08-23 20:48:50**

突然醒悟前几天看到有人用来生成歌词的貌似就是这个模型

**kk+国科大+NLP 2017-08-23 20:48:53**

但感觉实验上，他在图像的效果也很好呀

**Gapeng-北京大学-CV 2017-08-23 20:49:15**

其实有点模糊的，你看mnist

**kk+国科大+NLP 2017-08-23 20:49:32**

厉害了，我要用这个模型，我就可以毕业了

**Gapeng-北京大学-CV 2017-08-23 20:49:38**

细节上能分辨出来，就是因为mse项

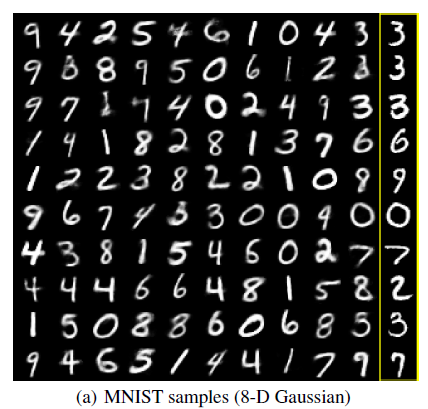
**安捷 北京大学 深度学习 2017-08-23 20:49:58**

mse为什么会造成模糊？

**starif-西电-ml 2017-08-23 20:50:04**

嗯，如果用照片一定就跪了

**Gapeng-北京大学-CV 2017-08-23 20:50:16**



**kk+国科大+NLP 2017-08-23 20:50:20**

但它实验数据的并不比gan低

**Gapeng-北京大学-CV 2017-08-23 20:50:29**

跟最右边那一列对比

**kk+国科大+NLP 2017-08-23 20:50:37**

嗯，这确实

**陆鹏起-hust-gan 2017-08-23 20:51:47**

mse这种pixel-wise的方法检验不了图像的结构化特征

**Gapeng-北京大学-CV 2017-08-23 20:52:00**

mse偏向于学所有可行解的平均结果，而不是任意一个可行解

**安捷 北京大学 深度学习 2017-08-23 20:52:35**

不够直观啊

**陆鹏起-hust-gan 2017-08-23 20:53:02**

把图像往任意方向平移一个距离mse又不一样了

**kk+国科大+NLP 2017-08-23 20:53:49**

所以mse在离散上会更具优势，是吗？

**Gapeng-北京大学-CV 2017-08-23 20:54:57**

也不是吧

**郑华滨-中山大学-深度学习 2017-08-23 20:54:59**

我记得一篇论文里面分析mse不好的原因是，没有考虑同一偏差值在不同context下对于人类的“显眼程度”

**安捷 北京大学 深度学习 2017-08-23 20:55:28**

换言之，mse模糊，不好，那什么好？

**郑华滨-中山大学-深度学习 2017-08-23 20:55:44**

比如在色彩突变交界处的像素偏差就会造成总体感觉上的模糊，但是同样的偏差值在别的地方可能就很不“显眼”

**kk+国科大+NLP 2017-08-23 20:55:45**

额

**郑华滨-中山大学-深度学习 2017-08-23 20:55:55**

似乎有很多替代方案

**kk+国科大+NLP 2017-08-23 20:56:15**

这个厉害了

**安捷 北京大学 深度学习 2017-08-23 20:56:15**

[可怜]

**kk+国科大+NLP 2017-08-23 20:56:25**

还记得论文吗

**郑华滨-中山大学-深度学习 2017-08-23 20:56:31**

最直接的一个就是改成L1 loss，还有方案是把绝对像素偏差换成相对偏差

**kk+国科大+NLP 2017-08-23 20:56:33**

题目？

**丁铭 清华 数据挖掘 2017-08-23 20:56:36**

简单的用mae，不过稍微难训练点

**郑华滨-中山大学-深度学习 2017-08-23 20:56:36**

有的哈，找找

**郑华滨-中山大学-深度学习 2017-08-23 20:57:00**

Loss Functions for Neural Networks for Image Processing

**kk+国科大+NLP 2017-08-23 20:57:00**

嗯嗯

**安捷 北京大学 深度学习 2017-08-23 20:57:23**

喔喔一说mae就有点明白了

**丁铭 清华 数据挖掘 2017-08-23 20:57:58**

很多地方直接用一个判别器替代

**丁铭 清华 数据挖掘 2017-08-23 20:58:22**

mse改成gan的loss

**Gapeng-北京大学-CV 2017-08-23 20:59:02**

这样结果会更sharp一些

**hwang-中南大学-生成模型 2017-08-23 20:59:38**

模糊地记得DL book里貌似说过这个问题

**Gapeng-北京大学-CV 2017-08-23 21:00:25**

我们进入下一个问题，AAE跟GAN相比，它不太适合图像，但是更加适合文本，对吗？

**丁铭 清华 数据挖掘 2017-08-23 21:00:50**

我不太懂vae，哪位大佬告诉我为啥reconstruction error是mse

**Gapeng-北京大学-CV 2017-08-23 21:00:58**

也就是说，毛豆大佬一开始说的那句话是不是对的

**丁铭 清华 数据挖掘 2017-08-23 21:01:05**

我看到代码都是与原图的mse

**安捷 北京大学 深度学习 2017-08-23 21:01:33**

如果刚才对mse的解释是正确的，我觉得是对的吧？

**安捷 北京大学 深度学习 2017-08-23 21:01:56**

都不sharp了生成啥[捂脸]

**Gapeng-北京大学-CV 2017-08-23 21:02:30**

或许有人有不同意见呢

**kk+国科大+NLP 2017-08-23 21:02:36**

我感觉aae在图像上确实还可以

**安捷 北京大学 深度学习 2017-08-23 21:02:45**

[可怜]

**丁铭 清华 数据挖掘 2017-08-23 21:02:49**

连文本怎么生成我搜不知道

**丁铭 清华 数据挖掘 2017-08-23 21:02:55**

都

**丁铭 清华 数据挖掘 2017-08-23 21:03:21**

不会是生成一列概率分布吧[捂脸]

**Gapeng-北京大学-CV 2017-08-23 21:03:31**

有一篇文章，ICLR 2017，不过被拒了

**Gapeng-北京大学-CV 2017-08-23 21:03:46**

Unsupervised sentence representation learning with adversarial auto-encoder

**Gapeng-北京大学-CV 2017-08-23 21:04:17**

网上现在下不到这篇文章了

**安捷 北京大学 深度学习 2017-08-23 21:04:21**

6666

**Gapeng-北京大学-CV 2017-08-23 21:04:46**

毛豆大佬给我的

**安捷 北京大学 深度学习 2017-08-23 21:05:15**

刚才咱们讨论的mse不sharp好像没有考虑到d的存在

**kk+国科大+NLP 2017-08-23 21:05:30**

厉害

**丁铭 清华 数据挖掘 2017-08-23 21:05:41**

确实是softmax生成一列分布？

**丁铭 清华 数据挖掘 2017-08-23 21:05:58**

那字典大了怎么办(⊙o⊙)

**Gapeng-北京大学-CV 2017-08-23 21:06:11**

D在的话，会有一些矫正作用把，相当于两者的一个加权

**安捷 北京大学 深度学习 2017-08-23 21:07:16**

d矫正的是encode部分，decode部分和mse好像也没什么影响？

**Gapeng-北京大学-CV 2017-08-23 21:07:32**

字典大了用sampling或者hierarchical？我猜的，那篇文章我没细看

**丁铭 清华 数据挖掘 2017-08-23 21:07:39**

是不是nlp字典大还没太好办法？

**安捷 北京大学 深度学习 2017-08-23 21:07:43**

我理解是不是d让encode decode的权重有了更大的随机性？

**丁铭 清华 数据挖掘 2017-08-23 21:08:12**

sampling不行吧

**丁铭 清华 数据挖掘 2017-08-23 21:08:36**

hierarchical应该还靠谱

**Gapeng-北京大学-CV 2017-08-23 21:09:34**

随机性？@安捷 北京大学 深度学习

**丁铭 清华 数据挖掘 2017-08-23 21:09:37**

再次提问，哪位大佬告诉我为啥reconstruction error是mse😂

**kk+国科大+NLP 2017-08-23 21:09:37**

其实图像和文字的主要区别在哪里？

**kk+国科大+NLP 2017-08-23 21:10:13**

除了一个连续一个离散

**kk+国科大+NLP 2017-08-23 21:10:29**

同问，难道不可以是其他吗？

**丁铭 清华 数据挖掘 2017-08-23 21:10:35**

还是个时序的

**Gapeng-北京大学-CV 2017-08-23 21:11:12**

reconstruction error，大概以前mse用的比较多，习惯性叫法？

**安捷 北京大学 深度学习 2017-08-23 21:11:16**

d训练的时候更新了encoder，这些更新是mse更新之外附加的？

**丁铭 清华 数据挖掘 2017-08-23 21:11:43**

看原文，并没有觉得跟mse有半毛钱关系

**Gapeng-北京大学-CV 2017-08-23 21:11:57**

是的@安捷 北京大学 深度学习

**丁铭 清华 数据挖掘 2017-08-23 21:12:00**

那一项按理说是个log

**丁铭 清华 数据挖掘 2017-08-23 21:12:44**

这个问题纠结了我好多天了

**Gapeng-北京大学-CV 2017-08-23 21:12:47**

你说的是VAE那里？

**丁铭 清华 数据挖掘 2017-08-23 21:12:54**

没找到任何理论依据

**丁铭 清华 数据挖掘 2017-08-23 21:13:41**

autoencoder variational bayes

**丁铭 清华 数据挖掘 2017-08-23 21:13:47**

貌似叫这个

**丁铭 清华 数据挖掘 2017-08-23 21:14:04**

反正就是那篇提出vae的文章

**Lynn-中科院计算所-CV 2017-08-23 21:14:42**

"log(p(x|z))在p(x|z)建模为高斯分布后推出来就是mse"

**kk+国科大+NLP 2017-08-23 21:15:11**

vae就是mse，那篇论文是这样写的

**kk+国科大+NLP 2017-08-23 21:15:38**

我能不能问一个最本质的问题

**丁铭 清华 数据挖掘 2017-08-23 21:15:43**

咦，有推导？

**丁铭 清华 数据挖掘 2017-08-23 21:15:54**

是在附录里吗？

**Gapeng-北京大学-CV 2017-08-23 21:16:02**

附录有推导

**kk+国科大+NLP 2017-08-23 21:16:16**

这个论题是aae更适合NLP还是图像

**kk+国科大+NLP 2017-08-23 21:16:54**

那么是不是首先寻找NLP和图像最大的区别是什么？

**安捷 北京大学 深度学习 2017-08-23 21:17:31**

连续和离散？

**kk+国科大+NLP 2017-08-23 21:17:35**

第二寻找aae的ae这一块不适合图像，说是mse

**kk+国科大+NLP 2017-08-23 21:18:10**

但论文提出的图像效果还行，是什么原因？

**kk+国科大+NLP 2017-08-23 21:18:46**

第三，mse不行，是不是可以改进来适应aae这个模型？

**kk+国科大+NLP 2017-08-23 21:18:54**

求大佬解答

**Gapeng-北京大学-CV 2017-08-23 21:19:06**

要达到论文的效果是需要一些超参数搜索的

**郑华滨-中山大学-深度学习 2017-08-23 21:19:23**

不仅仅有连续和离散，还有局部平滑性

**kk+国科大+NLP 2017-08-23 21:19:23**

这是我当时看到论题的三大疑问

**郑华滨-中山大学-深度学习 2017-08-23 21:19:37**

图像相邻像素有局部平滑性，文本没有

**kk+国科大+NLP 2017-08-23 21:19:42**

嗯嗯，

**郑华滨-中山大学-深度学习 2017-08-23 21:19:48**

感觉有局部平滑性的数据不太适合mse

**Gapeng-北京大学-CV 2017-08-23 21:20:15**

那L1 loss + TV loss能不能解决局部平滑性的问题？

**郑华滨-中山大学-深度学习 2017-08-23 21:20:39**

有可能

**Gapeng-北京大学-CV 2017-08-23 21:21:20**

好像并不能完全解决，我之前的实验发现还是会有问题

**kk+国科大+NLP 2017-08-23 21:21:39**

连续和离散这个鸿沟有什么方案呀

**Gapeng-北京大学-CV 2017-08-23 21:21:48**

checkboard现象似乎解决不了

**郑华滨-中山大学-深度学习 2017-08-23 21:22:04**

所以其实为什么不能把AAE和GAN合到一起呢2333

**安捷 北京大学 深度学习 2017-08-23 21:22:06**

checkboard是啥

**郑华滨-中山大学-深度学习 2017-08-23 21:22:15**

（应该有这样的论文吧？）

**Gapeng-北京大学-CV 2017-08-23 21:22:19**

棋盘效应

**郑华滨-中山大学-深度学习 2017-08-23 21:22:32**

马赛克状的纹理

**安捷 北京大学 深度学习 2017-08-23 21:22:40**

喔喔

**Gapeng-北京大学-CV 2017-08-23 21:22:41**

AAE把重构误差换成adv loss吗？

**Gapeng-北京大学-CV 2017-08-23 21:22:52**

哈哈，我一开始也是这么想的

**郑华滨-中山大学-深度学习 2017-08-23 21:23:22**

对的哈哈

**Gapeng-北京大学-CV 2017-08-23 21:23:50**

课后实验一下[机智]

**kk+国科大+NLP 2017-08-23 21:24:04**

这个可以有

**安捷 北京大学 深度学习 2017-08-23 21:24:04**

开始欣赏两位大佬的表演[奸笑]

**Lynn-中科院计算所-CV 2017-08-23 21:24:15**

aae的主要目的不是为了生成吧，我理解是为了编码，所以结合gan能做什么？

**Gapeng-北京大学-CV 2017-08-23 21:24:21**

但是有个问题，id perserving的问题怎么解决？

**kk+国科大+NLP 2017-08-23 21:24:27**

吃瓜群众已上线

**Lynn-中科院计算所-CV 2017-08-23 21:25:28**

vaegan怎么做的aaegan就怎么做😂

**郑华滨-中山大学-深度学习 2017-08-23 21:25:28**

AAE也可以生成吧

**Gapeng-北京大学-CV 2017-08-23 21:25:35**

嗯，aae更大的作用是在编码上

**安捷 北京大学 深度学习 2017-08-23 21:25:56**

为啥更大作用在编码

**Gapeng-北京大学-CV 2017-08-23 21:26:35**

生成上mse会造成模糊

**Gapeng-北京大学-CV 2017-08-23 21:26:40**

编码才是这个方法的主要目的？

**兔子-不存在-ML 2017-08-23 21:26:46**

因为生成效果比不上GAN，会模糊

**安捷 北京大学 深度学习 2017-08-23 21:26:57**

喔喔

**安捷 北京大学 深度学习 2017-08-23 21:27:03**

我一直有个疑问哈

**兔子-不存在-ML 2017-08-23 21:27:06**

（终于有一个问题我能回答了

**kk+国科大+NLP 2017-08-23 21:27:06**

哦哦

**安捷 北京大学 深度学习 2017-08-23 21:27:23**

为什么知道模糊是mse造成的而不是trans conv呢？

**Gapeng-北京大学-CV 2017-08-23 21:28:22**

transconv会造成checkboard，所以是不是也可以说，模糊有一部分原来也是来自于它

**kk+国科大+NLP 2017-08-23 21:29:14**

transconcv是什么？

**兔子-不存在-ML 2017-08-23 21:29:32**

逆卷积吧

**Gapeng-北京大学-CV 2017-08-23 21:29:32**

convolution transpose

**兔子-不存在-ML 2017-08-23 21:29:50**

这货的翻译千奇百怪

**Gapeng-北京大学-CV 2017-08-23 21:30:18**

“反卷积”

**阮翀-北大-NLP 2017-08-23 21:32:37**

fractional conv

**阮翀-北大-NLP 2017-08-23 21:32:46**

transposed conv

**kk+国科大+NLP 2017-08-23 21:32:57**

哦哦

**kk+国科大+NLP 2017-08-23 21:33:43**

那是不是还有反池化的锅

**Gapeng-北京大学-CV 2017-08-23 21:34:06**

你可以不用呀

**Gapeng-北京大学-CV 2017-08-23 21:34:29**

反卷积也可以不用的

**kk+国科大+NLP 2017-08-23 21:34:32**

嗯

**Lynn-中科院计算所-CV 2017-08-23 21:34:44**

upsampling加conv

**Gapeng-北京大学-CV 2017-08-23 21:35:20**

一直没有尝试，不知道换了效果会怎样

**Gapeng-北京大学-CV 2017-08-23 21:35:31**

@Lynn-中科院计算所-CV 你尝试过吗？

**kk+国科大+NLP 2017-08-23 21:35:47**

可以试试

**kk+国科大+NLP 2017-08-23 21:36:12**

说不定也是一个新思路

**Gapeng-北京大学-CV 2017-08-23 21:36:25**

大家还有什么问题，自由讨论吧，今天的讨论到此结束

**kk+国科大+NLP 2017-08-23 21:36:35**

好

**kk+国科大+NLP 2017-08-23 21:36:47**

谢谢，学习了好多

**Lynn-中科院计算所-CV 2017-08-23 21:36:55**

我试过在wgan-gp里面用，一塌糊涂，可能没调好。但我看到vae里面用有效果