第四次GAN讨论：image to image translation(part 2)

**Gapeng-北京大学-CV 2017-07-12 20:00:35**

好了，我们开始吧，今天的主题是image to image translation part 2

**Gapeng-北京大学-CV 2017-07-12 20:01:12**

第一个话题是GAN的一种应用——超分重建

**大俊-PaperWeekly 2017-07-12 20:01:12**

[强]

**Gapeng-北京大学-CV 2017-07-12 20:01:17**

有没有同学先来谈谈什么是超分重建？

**ZG 中科院计算所 人脸识别 2017-07-12 20:01:26**

@Gapeng-北京大学-CV unpaired image吗？

**Gapeng-北京大学-CV 2017-07-12 20:01:32**

讲讲这个任务是怎样的？数据是怎样的

**James-百度-OCR 2017-07-12 20:01:41**

低分辨率重建会高分辨率嘛

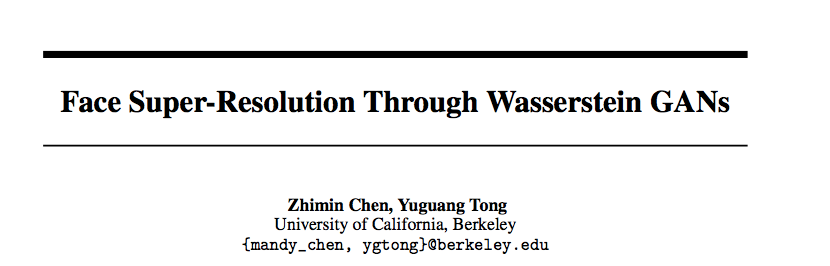
**Meow-XJTU-CV 2017-07-12 20:01:48**

图像低分辨率输入，高分辨率输出？

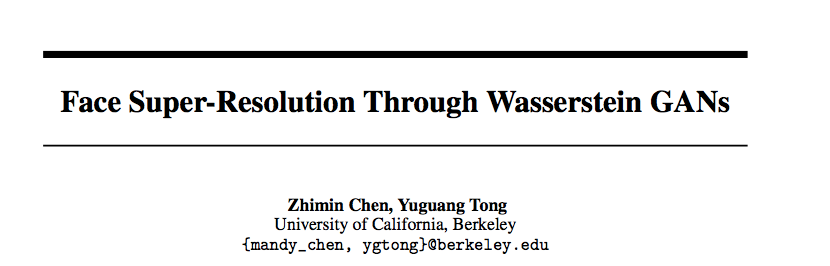
**Gapeng-北京大学-CV 2017-07-12 20:01:51**

unpair data已经讨论过了

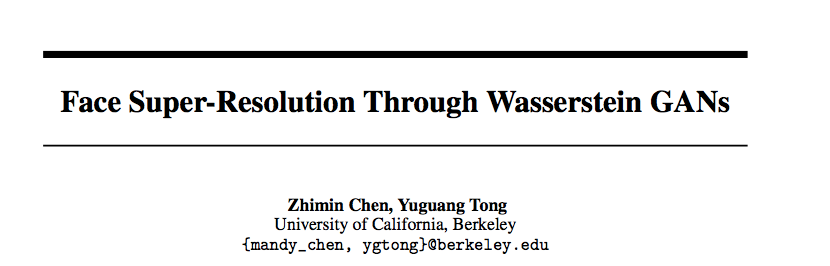
**许哲豪上理图像 2017-07-12 20:01:51**



**许哲豪上理图像 2017-07-12 20:01:51**



**许哲豪上理图像 2017-07-12 20:01:51**



**许哲豪上理图像 2017-07-12 20:02:03**

发现了这几篇论文

**Gapeng-北京大学-CV 2017-07-12 20:02:08**

嗯嗯

**Gapeng-北京大学-CV 2017-07-12 20:02:17**

别急，我们先把问题搞清楚

**Gapeng-北京大学-CV 2017-07-12 20:02:28**

也方便先做这个任务的同学

**Gapeng-北京大学-CV 2017-07-12 20:02:37**

训练数据是怎样的？

**Gapeng-北京大学-CV 2017-07-12 20:02:46**

需要pair data吗？

**Gapeng-北京大学-CV 2017-07-12 20:02:55**

还是unpair data也可以？

**许哲豪上理图像 2017-07-12 20:03:20**

需要pair的吧，

**郑华滨-中山大学-深度学习 2017-07-12 20:03:23**

任务上是把低分辨率图像变成高分辨率图像，也就是去马赛克

**ZG 中科院计算所 人脸识别 2017-07-12 20:03:44**

要是paired data cgan加l1就可以吧

**张治坤 浙大 信息安全 2017-07-12 20:03:45**

给没做过图像的科普一下，pair data是啥？😂

**Meow-XJTU-CV 2017-07-12 20:03:45**

输入数据是图像序列

**Gapeng-北京大学-CV 2017-07-12 20:03:48**

@SURE-天大-超分辨率-三星-人体姿态 你就是做超分重建的吧，来说说

**ZG 中科院计算所 人脸识别 2017-07-12 20:04:15**

其实unpaired data 我觉得cycle gan工作很不错

**郑华滨-中山大学-深度学习 2017-07-12 20:04:18**

pair data就是存在输入和输出样本之间的一一对应

**ZG 中科院计算所 人脸识别 2017-07-12 20:05:00**

直接用l1图像会模糊吧

**Gapeng-北京大学-CV 2017-07-12 20:05:06**

而且这种一一对应在训练的时候是可以让GAN获取的，而不是在学习的时候自己去学那种对应

**ZG 中科院计算所 人脸识别 2017-07-12 20:05:09**

cgan加上l1

**Gapeng-北京大学-CV 2017-07-12 20:05:37**

方法呢？有哪些GAN的方法？

**ZG 中科院计算所 人脸识别 2017-07-12 20:06:01**

2016 cvpr image to image translation

**ZG 中科院计算所 人脸识别 2017-07-12 20:06:10**

这个我觉得算一个

**Gapeng-北京大学-CV 2017-07-12 20:06:24**

能介绍一下思路吗？

**ZG 中科院计算所 人脸识别 2017-07-12 20:06:24**

是cycle gan前作

**SURE-天大-超分辨率-三星-人体姿态 2017-07-12 20:06:40**

在没有标准的数据集的时候一般是用高分辨率图像下采样成对应的低分辨率图像。所以输入的是图像对

**Gapeng-北京大学-CV 2017-07-12 20:06:49**

顺带贴一下GAN的结构图

**Meow-XJTU-CV 2017-07-12 20:07:50**

这个pair指的是低与高分辨率图像对吗？

**Gapeng-北京大学-CV 2017-07-12 20:08:06**

没错

**郑华滨-中山大学-深度学习 2017-07-12 20:08:08**

是的

**老白-中石大-cg 2017-07-12 20:08:17**

生成高分辨率是一步到位还是 逐步生成 ？

**Meow-XJTU-CV 2017-07-12 20:08:17**

谢！

**ZG 中科院计算所 人脸识别 2017-07-12 20:08:20**

这个就是同时用cgan+l1 loss,就是G有两个loss,一个是l1 loss,另一个是D的loss

**郑华滨-中山大学-深度学习 2017-07-12 20:08:36**

但是下午有人提出了，靠自己下采样得到的低分辨率图像可能不能反映真实世界中的模糊情况

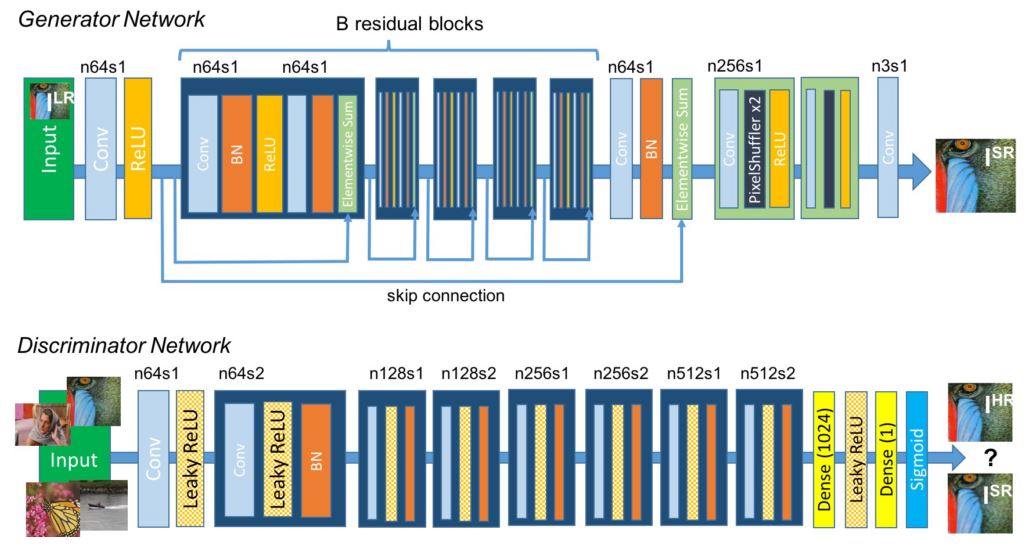
**Gapeng-北京大学-CV 2017-07-12 20:08:42**

我先来吧，SRGAN，用了adversarial loss + content loss (perceptual loss) [+ regularization loss (TV loss)]

**郑华滨-中山大学-深度学习 2017-07-12 20:08:55**

所以自己造高低分辨率paired data不一定好

**Gapeng-北京大学-CV 2017-07-12 20:09:01**



**Gapeng-北京大学-CV 2017-07-12 20:09:37**

自己造的数据确实代表性不足

**Gapeng-北京大学-CV 2017-07-12 20:09:52**

所以有什么解决方法吗？

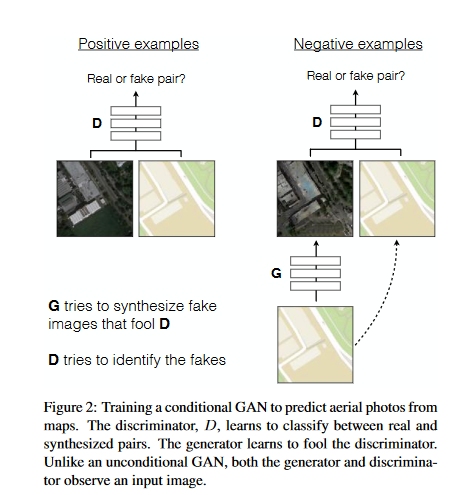
**许哲豪上理图像 2017-07-12 20:09:56**

但unpair的应该学习不到细节，只会学习到风格，不知对不对

**Gapeng-北京大学-CV 2017-07-12 20:10:44**

像cyclegan这类unpair的方法，细节确实学的不好

**ZG 中科院计算所 人脸识别 2017-07-12 20:10:44**



**郑华滨-中山大学-深度学习 2017-07-12 20:11:02**

最好的做法应该是收集真实世界中的pair data，当然这样成本比较高

**Gapeng-北京大学-CV 2017-07-12 20:11:07**

但是我猜测原因可能是没有加perceptual loss

**TM-科大-nlp cv 2017-07-12 20:11:10**

先对怎么造出有代表性的数据做个gan[捂脸]

**ZG 中科院计算所 人脸识别 2017-07-12 20:11:37**

cycle gan加上perceptual loss应该会更好

**Gapeng-北京大学-CV 2017-07-12 20:12:04**

嗯，这个我也是刚想到

**Gapeng-北京大学-CV 2017-07-12 20:12:26**

有没有同学试验过cyclegan加上perceptual loss的？

**Gapeng-北京大学-CV 2017-07-12 20:12:34**

分享一下成果？

**ZG 中科院计算所 人脸识别 2017-07-12 20:12:37**

"我实验过  
"

**尹伟东 复旦 深度学习 2017-07-12 20:12:37**

Conditional CycleGAN for Attribute Guided Face Image Generation这篇有这么干的

**尹伟东 复旦 深度学习 2017-07-12 20:12:56**

identity preserving loss用的是perceptual loss

**郑华滨-中山大学-深度学习 2017-07-12 20:13:02**

我有个问题，perceptual loss究竟为什么叫perceptual loss？虽然我知道它是什么

**Gapeng-北京大学-CV 2017-07-12 20:13:41**

我也不知道，求解答

**尹伟东 复旦 深度学习 2017-07-12 20:14:11**

不懂。。感觉像2016年一篇nips提出来的

**尹伟东 复旦 深度学习 2017-07-12 20:14:17**

忘了叫啥名

**张树-自动化所-机器视觉 2017-07-12 20:14:48**

LiFeiFei

**张树-自动化所-机器视觉 2017-07-12 20:14:56**

的一个学生做的ECCV的文章里面提出来的

**ZG 中科院计算所 人脸识别 2017-07-12 20:15:21**

"好像是neutal style transfer  
"

**张树-自动化所-机器视觉 2017-07-12 20:15:28**

本质上就是让学到的图和GT的图像 经过特征提取器之后，得到的特征尽量相似

**James-百度-OCR 2017-07-12 20:15:34**

perceptral loss是vgg feature map的匹配程度来着？

**Gapeng-北京大学-CV 2017-07-12 20:15:36**

Perceptual Losses for Real-Time Style Transfer and Super-Resolution

**Gapeng-北京大学-CV 2017-07-12 20:15:38**

这篇吗？

**张树-自动化所-机器视觉 2017-07-12 20:16:01**

对的

**Gapeng-北京大学-CV 2017-07-12 20:16:08**

很多paper的perceptual loss都是用VGG

**Gapeng-北京大学-CV 2017-07-12 20:16:32**

就是匹配图像的中间层特征

**张树-自动化所-机器视觉 2017-07-12 20:16:36**

专门做人脸的话 我们会用提人脸特征的CNN

**James-百度-OCR 2017-07-12 20:16:52**

"⌈ 张树-自动化所-机器视觉: 本质上就是让学到的图和GT的图像 经过特征提取器之后，得到的特征尽量相似 ⌋  
- - - - - - - - - - - - - - -  
特征相似的话就是提取器对这两张图的感受相近[奸笑]"

**ZG 中科院计算所 人脸识别 2017-07-12 20:17:27**

"人脸的话，用的是人脸识别器  
"

**郑华滨-中山大学-深度学习 2017-07-12 20:17:31**

perceptual loss都是用了预训练模型的样子

**James-百度-OCR 2017-07-12 20:17:37**

我是这么理解perceptual loss的

**Gapeng-北京大学-CV 2017-07-12 20:17:56**

嗯，都是预训练的

**张树-自动化所-机器视觉 2017-07-12 20:18:47**

这东西还是挺好用的 但是我的经验是，它对提升细节帮助不大

**Gapeng-北京大学-CV 2017-07-12 20:18:56**

实际上就是图像的高阶特征要匹配，所以也适合用来做id perserving

**泛艺术范儿-西电-DL 2017-07-12 20:19:37**

是怎么匹配的？ L2 loss？

**SURE-天大-超分辨率-三星-人体姿态 2017-07-12 20:19:41**

我有个问题

**SURE-天大-超分辨率-三星-人体姿态 2017-07-12 20:19:44**

从srcnn提出之后基本盖过了之前所有的机器学习方法，之后srgan改进了srcnn，但是是从主观评价上改进，psnr指标并不见得提高，也就是加入了人类经验脑补的功能。

**SURE-天大-超分辨率-三星-人体姿态 2017-07-12 20:19:49**

但是在纯提高细节的方向，比如车牌识别，医学图像我觉得不能加入猜的成分啊，有没有什么看法呢

**陆鹏起-hust-gan 2017-07-12 20:20:53**

perceptual loss也是pixel级别的吧

**郑华滨-中山大学-深度学习 2017-07-12 20:20:53**

确实是用了人类经验脑补，但是在有些场景下其实很好用

**郑华滨-中山大学-深度学习 2017-07-12 20:21:03**

不是

**ZG 中科院计算所 人脸识别 2017-07-12 20:21:33**

"perceptual loss 是ferture map级别的  
"

**Gapeng-北京大学-CV 2017-07-12 20:21:40**

脑补适合inpainting吧

**泛艺术范儿-西电-DL 2017-07-12 20:21:47**

主观评价和PSNR值确实不完全一致

**陆鹏起-hust-gan 2017-07-12 20:21:53**

噢

**郑华滨-中山大学-深度学习 2017-07-12 20:22:11**

其实超分辨率重构完全不利用脑补的话，什么都做不了吧

**郑华滨-中山大学-深度学习 2017-07-12 20:22:46**

低分辨率本来就损失了信息，你不可能凭空把那个信息还原出来

**Gapeng-北京大学-CV 2017-07-12 20:23:04**

细节就是算法脑补出来的

**ZG 中科院计算所 人脸识别 2017-07-12 20:23:08**

"应该就是要学习先验知识吧  
"

**郑华滨-中山大学-深度学习 2017-07-12 20:23:10**

只能借由一些别的真实样本引入额外的信息，这个过程其实就是脑补

**郑华滨-中山大学-深度学习 2017-07-12 20:23:16**

对

**ZG 中科院计算所 人脸识别 2017-07-12 20:23:34**

我记得之前有人用cycle gan做的去马赛克

**黄瑞阳\_郑州大学\_NLP 2017-07-12 20:23:43**

有个经典的VDSR，据说是现在超分辨率重建PSNR值最高的，不知道GAN效果怎样

**木羊同学-GDC-SL 2017-07-12 20:24:00**

脑补理解成高概率？

**ZG 中科院计算所 人脸识别 2017-07-12 20:24:09**

"还有人用cgan 做的去雾算法  
"

**泛艺术范儿-西电-DL 2017-07-12 20:24:14**

gan单纯看PSNR的话 效果不怎么样

**郑华滨-中山大学-深度学习 2017-07-12 20:24:23**

高likelihood吧

**郑华滨-中山大学-深度学习 2017-07-12 20:24:32**

求科普PSNR是啥

**尹伟东 复旦 深度学习 2017-07-12 20:24:42**

同求

**Gapeng-北京大学-CV 2017-07-12 20:24:45**

去雾，去雨都有人做了

**ZG 中科院计算所 人脸识别 2017-07-12 20:24:54**

"这个问题本身就是ill-posed problem  
"

**泛艺术范儿-西电-DL 2017-07-12 20:25:20**

"Peak signal-to-noise ratio  
"

**Gapeng-北京大学-CV 2017-07-12 20:25:20**

@泛艺术范儿-西电-DL 你来说说，PSNR是什么指标

**泛艺术范儿-西电-DL 2017-07-12 20:25:46**

就是把MSE换算了一下

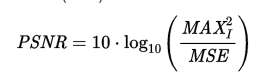
**Gapeng-北京大学-CV 2017-07-12 20:25:50**

这个指标怎么计算？

**Gapeng-北京大学-CV 2017-07-12 20:26:08**

再具体一点呢？

**泛艺术范儿-西电-DL 2017-07-12 20:26:30**



**小虎牙-阿里-相关性 2017-07-12 20:26:57**

有啥优点

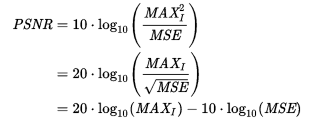
**ZG 中科院计算所 人脸识别 2017-07-12 20:27:00**

"感觉跟l2 loss很相关  
"

**Gapeng-北京大学-CV 2017-07-12 20:27:00**

哈哈

**Gapeng-北京大学-CV 2017-07-12 20:27:03**



**郑华滨-中山大学-深度学习 2017-07-12 20:27:03**

哦哦反正还是个pixel wise的指标

**Gapeng-北京大学-CV 2017-07-12 20:27:09**

摘自维基百科

**泛艺术范儿-西电-DL 2017-07-12 20:27:18**

公式里的MAX表示图像的最大像素值

**Meow-XJTU-CV 2017-07-12 20:27:24**

这样的换算是为了什么？

**小虎牙-阿里-相关性 2017-07-12 20:27:49**

同问

**SURE-天大-超分辨率-三星-人体姿态 2017-07-12 20:27:49**

大概来说就是和ground truth越接近 噪声就小 信噪比就越大

**Gapeng-北京大学-CV 2017-07-12 20:28:04**

对于RGB来说，MAX\_I就是255

**Meow-XJTU-CV 2017-07-12 20:29:02**

这取决于图像深度吧

**Hsu-img reconstruction 2017-07-12 20:29:09**

我做过16位数据

**黄瑞阳\_郑州大学\_NLP 2017-07-12 20:29:09**

PSNR的计算中，一般用的颜色空间有人熟悉吗？

**泛艺术范儿-西电-DL 2017-07-12 20:29:24**

对 就是把MSE换算到对数，比较的时候更直观了。比如MSE从0.1到0.99，和MSE从0.02到0.01虽然都降低了0.01，但作用其实完全不同

**Hsu-img reconstruction 2017-07-12 20:29:48**

先做了归一化

**小虎牙-阿里-相关性 2017-07-12 20:29:52**

感觉是更敏感

**一路顺枫-THU-Hash 2017-07-12 20:29:53**

为什么要用mse？

**Meow-XJTU-CV 2017-07-12 20:30:11**

对数运算可以压缩数据范围

**泛艺术范儿-西电-DL 2017-07-12 20:30:11**

换到对数，算MSE降低的比率更直观

**黄瑞阳\_郑州大学\_NLP 2017-07-12 20:30:26**

YCbCr，一般我看计算PSNR都是用这个颜色空间，核心在于Y的计算。

**一路顺枫-THU-Hash 2017-07-12 20:30:39**

感觉mse很反直觉啊

**泛艺术范儿-西电-DL 2017-07-12 20:31:26**

MSE是传统图像处理领域用的很多的一个指标，对于high level的任务确实没太大意义

**Gapeng-北京大学-CV 2017-07-12 20:31:30**

mse跟脑补有那么一点不和谐在那里

**小虎牙-阿里-相关性 2017-07-12 20:32:09**

脑补总会有点差距。。

**anshiquanshu66-师大-医学图像处理 2017-07-12 20:32:18**

安静的听分享[呲牙]

**郑华滨-中山大学-深度学习 2017-07-12 20:32:24**

那么现在超分辨率+GAN是不是都是条件生成的套路？

**郑华滨-中山大学-深度学习 2017-07-12 20:32:30**

conditional generation

**泛艺术范儿-西电-DL 2017-07-12 20:33:01**

对 就是把低分辨的图送进去

**郑华滨-中山大学-深度学习 2017-07-12 20:33:25**

但是判别器那边应该也需要送低分辨率图对吧

**ZG 中科院计算所 人脸识别 2017-07-12 20:33:30**

"如果是paired data，还可以用l1 loss  
"

**ZG 中科院计算所 人脸识别 2017-07-12 20:33:39**

高分辨率吧

**泛艺术范儿-西电-DL 2017-07-12 20:33:50**

判别器不需要低分变

**一路顺枫-THU-Hash 2017-07-12 20:33:50**

但网络结构的设计就有讲究了

**郑华滨-中山大学-深度学习 2017-07-12 20:34:05**

不，我是说高低分辨率一起送

**ZG 中科院计算所 人脸识别 2017-07-12 20:34:08**

或者用patch gan

**Gapeng-北京大学-CV 2017-07-12 20:34:19**

判别器的任务是什么？

**泛艺术范儿-西电-DL 2017-07-12 20:34:23**

判别器判别生成的高分辨图和原始的高分辨图

**Meow-XJTU-CV 2017-07-12 20:34:43**

努力区分真假

**泛艺术范儿-西电-DL 2017-07-12 20:34:46**

一般都是按照图像块送进去

**ZG 中科院计算所 人脸识别 2017-07-12 20:34:57**

"D只判断一个patch是真实的高分辨率图片，还是生成的  
"

**泛艺术范儿-西电-DL 2017-07-12 20:35:03**

比如65\*65这样

**郑华滨-中山大学-深度学习 2017-07-12 20:35:07**

哦反正只需要做局部判别

**Gapeng-北京大学-CV 2017-07-12 20:35:09**

判别器的loss只有adversarial loss吗？还是要加其他的loss？

**ZG 中科院计算所 人脸识别 2017-07-12 20:35:12**

同意

**ZG 中科院计算所 人脸识别 2017-07-12 20:35:18**

加上l1 loss

**ZG 中科院计算所 人脸识别 2017-07-12 20:35:28**

"  
收敛更快"

**泛艺术范儿-西电-DL 2017-07-12 20:35:41**

l1 loss有人做过了

**Gapeng-北京大学-CV 2017-07-12 20:35:45**

l1 loss是加给G的吧？

**ZG 中科院计算所 人脸识别 2017-07-12 20:35:53**

"  
嗯嗯"

**一路顺枫-THU-Hash 2017-07-12 20:35:57**

gan那篇为啥用的是l2？

**ZG 中科院计算所 人脸识别 2017-07-12 20:36:01**

G有两个loss

**Gapeng-北京大学-CV 2017-07-12 20:36:05**

l1和l2都有人做过了

**ZG 中科院计算所 人脸识别 2017-07-12 20:36:13**

一个是l1 loss，另一个是D的loss

**ZG 中科院计算所 人脸识别 2017-07-12 20:36:38**

或者可以用l1 loss先对G初始化

**Gapeng-北京大学-CV 2017-07-12 20:37:03**

SRGAN用的是TV loss，不是l1或l2吧

**Gapeng-北京大学-CV 2017-07-12 20:37:44**

这些loss有什么作用？

**一路顺枫-THU-Hash 2017-07-12 20:37:48**

在哪里说了srgan用了tvloss？

**ZG 中科院计算所 人脸识别 2017-07-12 20:37:48**

TV Loss是使生成图像更加平滑

**Gapeng-北京大学-CV 2017-07-12 20:38:30**

我看了一个PPT，有人给他加了TV loss，原始论文是没有的

**ZG 中科院计算所 人脸识别 2017-07-12 20:38:30**

l1 Loss的监督作用，比adversial Loss更强，但是会产生blurry

**ZG 中科院计算所 人脸识别 2017-07-12 20:38:40**

TV Loss是使生成图像更加平滑

**一路顺枫-THU-Hash 2017-07-12 20:38:58**

tvloss是什么？

**ZG 中科院计算所 人脸识别 2017-07-12 20:39:01**

TV就是total variance

**Gapeng-北京大学-CV 2017-07-12 20:39:02**

但是有人说L2 loss会产生blurry

**Gapeng-北京大学-CV 2017-07-12 20:39:07**

你们怎么看？

**一路顺枫-THU-Hash 2017-07-12 20:39:18**

具体的形式是？

**ZG 中科院计算所 人脸识别 2017-07-12 20:39:20**

"  
所以加上了adversial loss"

**ZG 中科院计算所 人脸识别 2017-07-12 20:39:43**

"TV Loss就是图像的  
梯度"

**ZG 中科院计算所 人脸识别 2017-07-12 20:39:57**

防止生成图像突变

**一路顺枫-THU-Hash 2017-07-12 20:40:06**

有数学形式吗？

**小虎牙-阿里-相关性 2017-07-12 20:40:09**

l2更模糊？

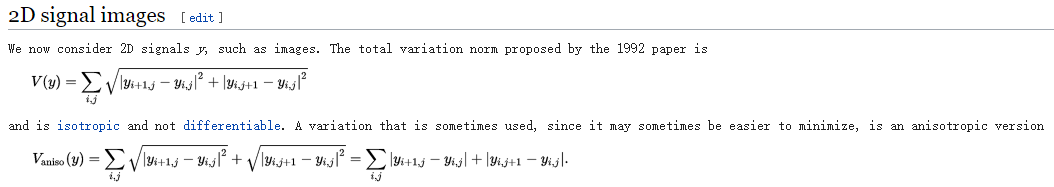
**ZG 中科院计算所 人脸识别 2017-07-12 20:41:09**

我忘了我在哪篇论文里面看到过

**ZG 中科院计算所 人脸识别 2017-07-12 20:41:25**

反正好多篇论文里面都有这个tv loss

**Gapeng-北京大学-CV 2017-07-12 20:41:59**



**全盛-蔚来汽车-NLP 2017-07-12 20:42:07**

可以tvloss l2两个loss都用么

**尹伟东 复旦 深度学习 2017-07-12 20:42:22**

最近fader network貌似用的只是来

**cherish+天大+SR 2017-07-12 20:42:24**

原文加了vgg loss

**尹伟东 复旦 深度学习 2017-07-12 20:42:27**

l2 loss?

**一路顺枫-THU-Hash 2017-07-12 20:42:33**

相当于一个regulizer？

**Gapeng-北京大学-CV 2017-07-12 20:42:50**

嗯，是regularizer

**陆鹏起-hust-gan 2017-07-12 20:43:28**

感觉就是一个衡量函数变化程度的

**一路顺枫-THU-Hash 2017-07-12 20:43:28**

在概率度量里也有一个tv divergence是什么东西？

**泛艺术范儿-西电-DL 2017-07-12 20:43:31**

tv loss算的是图像梯度的L1范数。所以也可以理解为希望图像的梯度是稀疏的。

**ZG 中科院计算所 人脸识别 2017-07-12 20:43:54**

其实就是防止生成图像突变

**cherish+天大+SR 2017-07-12 20:44:55**

也就是说 l1 loss 理论上会比 l2 loss的重建效果更好？

**ZG 中科院计算所 人脸识别 2017-07-12 20:45:08**

对啊

**Eric-SCU-CV 2017-07-12 20:45:22**

好像有篇论文是这样说的

**Eric-SCU-CV 2017-07-12 20:45:41**

但是具体有啥依据啊

**Gapeng-北京大学-CV 2017-07-12 20:45:52**

我还看到有篇论文反过来说的

**Gapeng-北京大学-CV 2017-07-12 20:46:22**

这个还是看实验结果吧

**Gapeng-北京大学-CV 2017-07-12 20:46:31**

再来一个问题，超高分辨率的图像重建，有什么特殊的trick可以用吗？

**SURE-天大-超分辨率-三星-人体姿态 2017-07-12 20:47:05**

超高分辨率？

**Gapeng-北京大学-CV 2017-07-12 20:47:08**

我记得github上面有人做了非常高分辨率的重建，好像是4096\*4096的

**Gapeng-北京大学-CV 2017-07-12 20:47:20**

不过细节真的是脑补的。。。

**泛艺术范儿-西电-DL 2017-07-12 20:47:48**

在纯的超分辨领域，评价性能的指标是PSNR（和MSE直接挂钩），所以如果单纯看PSNR值可能还是L2要好。如果考虑主观感受的话估计L1要好

**ZG 中科院计算所 人脸识别 2017-07-12 20:48:46**

本来如果图片很模糊，人也是脑补啊

**Lynn-中科院计算所-CV 2017-07-12 20:48:58**

有的文章说l2比l1模糊，有的说l1梯度特性更好

**cherish+天大+SR 2017-07-12 20:49:25**

但是用了gan的psnr可能会比bicubic插值重建的psnr还差

**Lynn-中科院计算所-CV 2017-07-12 20:50:12**

最近出的PAN，做去雨去雾的，超分辨率应该也是可以做的

**SURE-天大-超分辨率-三星-人体姿态 2017-07-12 20:50:27**

但是人可以主观控制自己脑补的程度 有没有可能让机器也能加进去一个参数控制这个程度

**Gapeng-北京大学-CV 2017-07-12 20:52:44**

这个话题还有要拓展的吗？

**Gapeng-北京大学-CV 2017-07-12 20:52:52**

没有的话我们就进入下一个话题了

**Gapeng-北京大学-CV 2017-07-12 20:53:15**

今天的第二个话题是：图像属性修改

**Gapeng-北京大学-CV 2017-07-12 20:54:10**

这个名字是我自己起的，它包含了一类这样的任务：用GAN对图像内容的某些属性进行修改

**超-UConn-GAN 2017-07-12 20:54:20**

属性是指什么呢 风格？

**ZG 中科院计算所 人脸识别 2017-07-12 20:54:33**

最直接的就是人脸属性控制

**ZG 中科院计算所 人脸识别 2017-07-12 20:54:38**

[捂脸]

**Gapeng-北京大学-CV 2017-07-12 20:54:43**

比如，笑脸和严肃脸转换，戴眼镜和不戴眼镜

**cherish+天大+SR 2017-07-12 20:55:02**

类似风格迁移么

**Meow-XJTU-CV 2017-07-12 20:55:11**

用于人脸表情合成，迁移，交互修改

**超-UConn-GAN 2017-07-12 20:55:20**

感觉这个意思和风格迁移还不太一样

**超-UConn-GAN 2017-07-12 20:55:26**

是局部信息

**泛艺术范儿-西电-DL 2017-07-12 20:55:38**

是在已有的图片上进行修改 还是说按照属性直接要求生成一张图片？

**Meow-XJTU-CV 2017-07-12 20:55:38**

“修改”

**Meow-XJTU-CV 2017-07-12 20:56:00**

对于人脸，例如表情变化

**泛艺术范儿-西电-DL 2017-07-12 20:56:00**

哦

**一路顺枫-THU-Hash 2017-07-12 20:56:06**

侧脸变正算吗？

**Gapeng-北京大学-CV 2017-07-12 20:56:12**

再比如，物体的色彩、形状的修改

**Gapeng-北京大学-CV 2017-07-12 20:56:29**

正侧面也算吧，这个没有严格定义

**anshiquanshu66-师大-医学图像处理 2017-07-12 20:56:36**

根据侧脸生成正脸

**Meow-XJTU-CV 2017-07-12 20:56:39**

我现在做的就是针对人脸的表情合成，修改，问题范围更小了

**一路顺枫-THU-Hash 2017-07-12 20:56:54**

+1

**Gapeng-北京大学-CV 2017-07-12 20:57:43**

我们可以把问题先局限在人脸的属性修改上，这样方便讨论，交互修改也是一种，但是跟人脸的又有区别

**Meow-XJTU-CV 2017-07-12 20:57:48**

语义属性比如：肤色，圆脸，高鼻梁，戴不戴眼镜这些？有别的定义吗？

**郑华滨-中山大学-深度学习 2017-07-12 20:57:51**

属性修改相当于要细粒度地做风格转换了

**郑华滨-中山大学-深度学习 2017-07-12 20:58:03**

这些都可以

**Gapeng-北京大学-CV 2017-07-12 20:58:09**

可以这么说

**郑华滨-中山大学-深度学习 2017-07-12 20:58:30**

我觉得问题定义应该是你想转什么就转什么，高度受控

**Meow-XJTU-CV 2017-07-12 20:58:36**



**James-百度-OCR 2017-07-12 20:59:09**

正脸到侧脸这个工作是用的unpair的数据吗

**James-百度-OCR 2017-07-12 20:59:27**

\*侧脸到正脸

**郑华滨-中山大学-深度学习 2017-07-12 20:59:30**

问题是一般有什么数据可以用，是不是肯定要有label，比如我要在表情上做转换，是不是一定要有每一张人脸的表情lable

**一路顺枫-THU-Hash 2017-07-12 20:59:40**

有pair的数据

**Meow-XJTU-CV 2017-07-12 21:00:04**

我记得中科院自动化所有一个侧脸转正脸

**Gapeng-北京大学-CV 2017-07-12 21:00:07**

label是需要的，可以是weak label

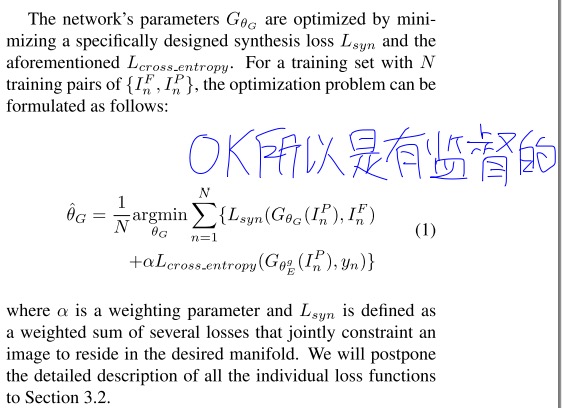
**一路顺枫-THU-Hash 2017-07-12 21:00:13**

最经典是是celeba

**Gapeng-北京大学-CV 2017-07-12 21:00:24**

就是with/without attribute这种

**郑华滨-中山大学-深度学习 2017-07-12 21:00:27**



**郑华滨-中山大学-深度学习 2017-07-12 21:00:34**

paired

**Gapeng-北京大学-CV 2017-07-12 21:00:49**

正侧脸的转换也不太接近今天要讨论的话题

**James-百度-OCR 2017-07-12 21:00:59**

酷

**一路顺枫-THU-Hash 2017-07-12 21:01:04**

multipie数据集是有标签的，所以可以做pair生成

**Meow-XJTU-CV 2017-07-12 21:01:15**

"Beyond Face Rotation: Global and Local Perception GAN for Photorealistic and Identity Preserving Frontal View Synthesis  
[捂脸]"

**Gapeng-北京大学-CV 2017-07-12 21:01:21**

还是回到人脸语义属性的修改上面来吧

**James-百度-OCR 2017-07-12 21:01:23**

属性修改用的数据需要时pair的么 0 0 这个没想明白unpair的怎么做。。

**尹伟东 复旦 深度学习 2017-07-12 21:01:38**

可以不用paired

**郑华滨-中山大学-深度学习 2017-07-12 21:01:44**

我觉得不一定需要pair，但是要有label

**Gapeng-北京大学-CV 2017-07-12 21:01:44**

可以不用pair

**一路顺枫-THU-Hash 2017-07-12 21:01:50**

最简单就是cyclegan

**郑华滨-中山大学-深度学习 2017-07-12 21:01:58**

我要转表情，至少要有各种表情的lable

**Gapeng-北京大学-CV 2017-07-12 21:02:05**

GeneGAN就不用pair，不过label，至少weak label是需要的

**郑华滨-中山大学-深度学习 2017-07-12 21:02:34**

cyclegan能做，也是因为有label，使你知道左边是一类图片，右边是另一类图片

**郑华滨-中山大学-深度学习 2017-07-12 21:02:46**

没label你是分不开两类图片的

**一路顺枫-THU-Hash 2017-07-12 21:03:01**

是的

**James-百度-OCR 2017-07-12 21:03:04**

有道理

**Gapeng-北京大学-CV 2017-07-12 21:03:18**

这个任务有什么实际应用场景吗？

**郑华滨-中山大学-深度学习 2017-07-12 21:03:27**

genegan怎么做呢？

**郑华滨-中山大学-深度学习 2017-07-12 21:03:38**

实际场景就是美图了吧

**郑华滨-中山大学-深度学习 2017-07-12 21:03:47**

定向美颜？

**Gapeng-北京大学-CV 2017-07-12 21:03:59**

GeneGAN我们一会儿邀请@Xiao-pku-math 来介绍

**一路顺枫-THU-Hash 2017-07-12 21:04:04**

还可以辅助人脸识别

**Gapeng-北京大学-CV 2017-07-12 21:04:04**

嗯，算一个

**郑华滨-中山大学-深度学习 2017-07-12 21:04:04**

吼

**Meow-XJTU-CV 2017-07-12 21:04:16**

哇

**Gapeng-北京大学-CV 2017-07-12 21:04:29**

怎么辅助@一路顺枫-THU-Hash ？

**一路顺枫-THU-Hash 2017-07-12 21:05:10**

实际遇到的，比如表情诡异的时候经常会导致人脸识别失效

**尹伟东 复旦 深度学习 2017-07-12 21:05:45**

是类似于把眼镜去掉这样的工作么

**Gapeng-北京大学-CV 2017-07-12 21:05:57**

你需要一类表情诡异（但相近）的图像用来当训练集。。。

**一路顺枫-THU-Hash 2017-07-12 21:05:59**

所以或许生成无表情图片可以提高准确率，当然前提是声称质量高

**Gapeng-北京大学-CV 2017-07-12 21:06:07**

嗯，去眼镜算一个

**一路顺枫-THU-Hash 2017-07-12 21:06:27**

眼镜也是很蛋疼的问题

**一路顺枫-THU-Hash 2017-07-12 21:06:47**

网上有关眼镜对人脸识别影响的报道挺多的

**Gapeng-北京大学-CV 2017-07-12 21:07:39**

还有应用吗？我想不到更多了

**Gapeng-北京大学-CV 2017-07-12 21:08:11**

没有的话，我们邀请两位同学来分享他们的工作

**anshiquanshu66-师大-医学图像处理 2017-07-12 21:08:15**

缺脸补全呢？

**Meow-XJTU-CV 2017-07-12 21:08:44**

视点变换

**SHF同学-同济-GAN 2017-07-12 21:08:50**

人脸遮挡修补

**郑华滨-中山大学-深度学习 2017-07-12 21:08:50**

遮挡？

**郑华滨-中山大学-深度学习 2017-07-12 21:08:58**

对

**Meow-XJTU-CV 2017-07-12 21:09:12**

很像去马赛克……

**SHF同学-同济-GAN 2017-07-12 21:09:20**

部分遮挡的那种

**bearbee 清华大学 GAN 2017-07-12 21:09:32**

今年cvpr不也有一篇相关论文吗

**郑华滨-中山大学-深度学习 2017-07-12 21:09:50**

不过遮挡修补应该算inpaiting了

**Gapeng-北京大学-CV 2017-07-12 21:09:57**

放个标题上来

**泛艺术范儿-西电-DL 2017-07-12 21:10:01**

有木有素颜生成化妆的？

**bearbee 清华大学 GAN 2017-07-12 21:10:06**

通过gan提高图片遮挡识别鲁棒性的

**Meow-XJTU-CV 2017-07-12 21:10:09**

传统的用RBF可以实现变形，这也算一种属性修改吧

**Gapeng-北京大学-CV 2017-07-12 21:10:16**

要是能大概介绍一下cvpr那篇怎么做的就更好了

**一路顺枫-THU-Hash 2017-07-12 21:11:00**

两个gan，一个gan生成遮挡部分，一个生成全图

**Gapeng-北京大学-CV 2017-07-12 21:11:54**

有文章id吗？

**一路顺枫-THU-Hash 2017-07-12 21:12:33**

generative face completion

**SHF同学-同济-GAN 2017-07-12 21:12:33**

@泛艺术范儿-西电-DL 隐约记得有化妆后的转换为卸妆素颜的，记不得文章名了

**Gapeng-北京大学-CV 2017-07-12 21:12:41**

我们有点跑偏了哈哈，这个算是inpainting了

**bearbee 清华大学 GAN 2017-07-12 21:12:41**

A-Fast-RCNN: Hard positive generation via adversary for object detection

**Gapeng-北京大学-CV 2017-07-12 21:13:31**

哦，这篇啊，这个跟今天讨论的话题偏的有点多了

**Gapeng-北京大学-CV 2017-07-12 21:14:02**

好了，下面进入分享环节。首先欢迎@Xiao-pku-math 给我们分享GeneGAN

**上海滩文哥-复旦-CV 2017-07-12 21:14:26**

gan用来生成图片在精度上的问题怎么解决？

**黄瑞阳\_郑州大学\_NLP 2017-07-12 21:14:29**

[强]

**上海滩文哥-复旦-CV 2017-07-12 21:14:40**

generative face completion

**郑华滨-中山大学-深度学习 2017-07-12 21:14:55**

欢迎[强]

**anshiquanshu66-师大-医学图像处理 2017-07-12 21:15:04**

欢迎

**Xiao-pku-math 2017-07-12 21:15:14**

其实想法很简单

**SHF同学-同济-GAN 2017-07-12 21:15:16**

欢迎

**Gapeng-北京大学-CV 2017-07-12 21:15:18**

GeneGAN是干什么的？

**上海滩文哥-复旦-CV 2017-07-12 21:15:18**

刚刚把前面的问题看完

**Xiao-pku-math 2017-07-12 21:16:13**

GeneGAN也是一种做object transfiguration的东西

**Xiao-pku-math 2017-07-12 21:16:28**

但是和cyclegan这些不同

**Gapeng-北京大学-CV 2017-07-12 21:16:41**

不同在哪里？

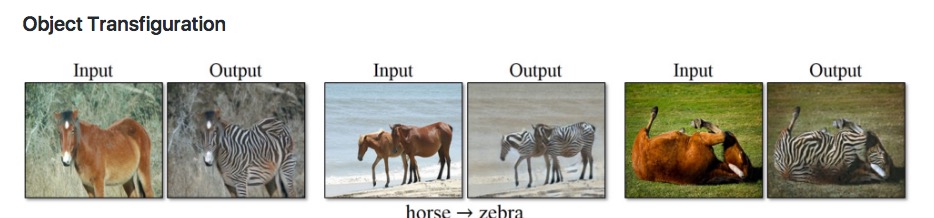
**Xiao-pku-math 2017-07-12 21:16:50**

大家有没有注意到cyclegan的效果图

**ZG 中科院计算所 人脸识别 2017-07-12 21:17:11**

确实不怎么好

**Xiao-pku-math 2017-07-12 21:17:11**



**Xiao-pku-math 2017-07-12 21:17:16**

从github上截图的

**上海滩文哥-复旦-CV 2017-07-12 21:17:36**

效果很一般cyclegan

**Xiao-pku-math 2017-07-12 21:18:00**

不但把马变成斑马，还把背景也变了

**Gapeng-北京大学-CV 2017-07-12 21:18:03**

嗯嗯，背景也改变了

**Xiao-pku-math 2017-07-12 21:18:05**

可能很多人没有注意到

**木羊同学-GDC-SL 2017-07-12 21:18:22**

正想说 背景花了

**阮翀-北大-NLP 2017-07-12 21:18:26**

我感觉变背景没什么问题……

**郑华滨-中山大学-深度学习 2017-07-12 21:18:33**

嗯背景是style的一部分，不爽不要玩[奸笑]

**Xiao-pku-math 2017-07-12 21:18:37**



**Xiao-pku-math 2017-07-12 21:18:40**

how about this

**阮翀-北大-NLP 2017-07-12 21:18:55**

因为 context 对物体判别也很重要

**阮翀-北大-NLP 2017-07-12 21:19:09**

你让一匹普通的马在沙漠里跑不太好吧

**郑华滨-中山大学-深度学习 2017-07-12 21:19:14**

那其实你是想控制有的地方变有的地方不变？

**此用户不存在–ustc–utokyo–cv 2017-07-12 21:19:17**

cyclegan马变斑马的背景色差很明显

**天清-北二外-信息安全 2017-07-12 21:19:17**

Putin。。。

**阮翀-北大-NLP 2017-07-12 21:19:19**

判别器会觉得你智障……

**泛艺术范儿-西电-DL 2017-07-12 21:19:22**

而且背景和前景的变换方向是一致的

**此用户不存在–ustc–utokyo–cv 2017-07-12 21:19:36**

猜是斑马的黑白和马的棕色色差太大？

**泛艺术范儿-西电-DL 2017-07-12 21:19:47**

说明根本没有分开前景和背景

**上海滩文哥-复旦-CV 2017-07-12 21:19:56**

生成时候无法区分前景和背景

**Xiao-pku-math 2017-07-12 21:20:02**

对

**Gapeng-北京大学-CV 2017-07-12 21:20:18**

那GeneGAN怎么做到区分前景和背景的呢？

**Meow-XJTU-CV 2017-07-12 21:20:26**

那么genegan呢

**木羊同学-GDC-SL 2017-07-12 21:20:32**

骑马的人也算背景么

**小虎牙-阿里-相关性 2017-07-12 21:20:37**

我看过一个苹果变橘子的..

**Xiao-pku-math 2017-07-12 21:20:44**

按我的理解，从cyclegan的设计中来看，他是一种全局的修改，对前景和背景并没有很好的区分

**一路顺枫-THU-Hash 2017-07-12 21:20:52**

在cyclegan加identity loss可以解决吗？

**郑华滨-中山大学-深度学习 2017-07-12 21:21:12**

同意，保留背景有时候是需求之一

**郑华滨-中山大学-深度学习 2017-07-12 21:21:41**

反过来，保留前景有时候也是需求，比如那种说”给我P个霸气的背景“[奸笑]

**Xiao-pku-math 2017-07-12 21:22:08**

所以GeneGAN的设计上是想区分出前景和背景的

**一路顺枫-THU-Hash 2017-07-12 21:22:27**

原来如此

**上海滩文哥-复旦-CV 2017-07-12 21:22:37**

可以借鉴faster-rcnn的结构来区分前景核背景么

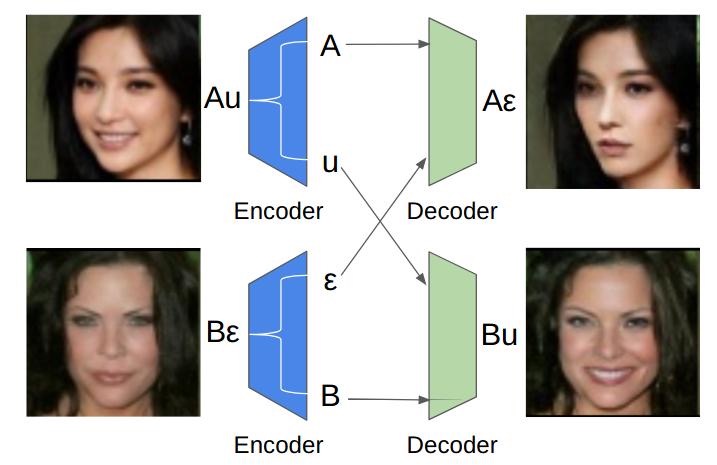
**Meow-XJTU-CV 2017-07-12 21:22:47**

方法？

**上海滩文哥-复旦-CV 2017-07-12 21:23:01**

在鉴别器端

**Xiao-pku-math 2017-07-12 21:23:07**



**Xiao-pku-math 2017-07-12 21:23:10**

参考这个图

**Xiao-pku-math 2017-07-12 21:23:56**

就是前景和背景都是在图像中的，我们相用一个encoder，把前景和背景分开

**此用户不存在–ustc–utokyo–cv 2017-07-12 21:24:04**

这图让我想起一个换脸的论文

**此用户不存在–ustc–utokyo–cv 2017-07-12 21:24:13**

男的全换成尼古拉斯凯奇

**此用户不存在–ustc–utokyo–cv 2017-07-12 21:24:22**

女的全换成泰勒斯威夫特

**Gapeng-北京大学-CV 2017-07-12 21:24:26**

怎么确保它能把前景和背景分开呢？

**此用户不存在–ustc–utokyo–cv 2017-07-12 21:24:52**

但脸部用人脸探测就行

**Xiao-pku-math 2017-07-12 21:25:00**

然后交换图片的属性

**bearbee 清华大学 GAN 2017-07-12 21:25:06**

剔除背景后，前景的分布也会变吧

**Xiao-pku-math 2017-07-12 21:25:15**

再用decoder重新恢复出原图

**郑华滨-中山大学-深度学习 2017-07-12 21:25:23**

前景不一定只是脸

**Eric-SCU-CV 2017-07-12 21:25:30**

@Xiao-pku-math 这是哪个论文？

**Gapeng-北京大学-CV 2017-07-12 21:25:55**

GeneGAN: Learning Object Transfiguration and Attribute Subspace from Unpaired Data

**Xiao-pku-math 2017-07-12 21:25:55**

http://in.arxiv.org/pdf/1705.04932v1

**Gapeng-北京大学-CV 2017-07-12 21:26:20**

@Xiao-pku-math 你继续，先把思路讲完

**Xiao-pku-math 2017-07-12 21:26:26**

我们用的是无配对的数据

**Gapeng-北京大学-CV 2017-07-12 21:26:55**

怎么确保能够区分前景和背景的

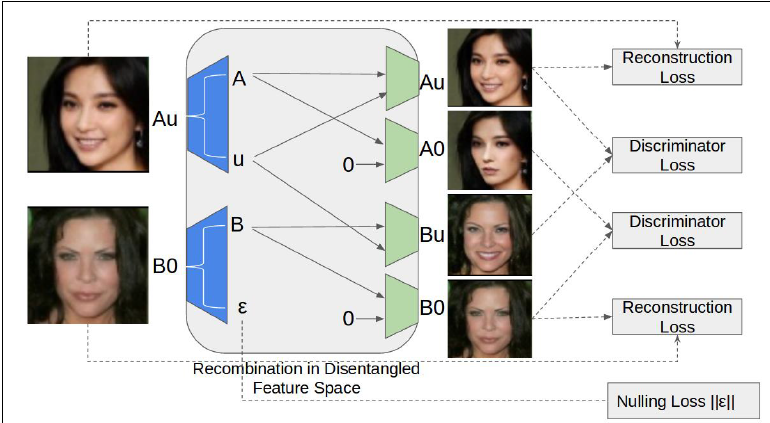
**Xiao-pku-math 2017-07-12 21:27:16**

能让网络能把有属性的部分塞在u中，关键的地方在Nulling loss

**郑华滨-中山大学-深度学习 2017-07-12 21:27:33**

同问，以及背景怎么保留

**Gapeng-北京大学-CV 2017-07-12 21:28:24**



**Gapeng-北京大学-CV 2017-07-12 21:28:34**

放个图上来，方便理解

**Xiao-pku-math 2017-07-12 21:28:41**

因为一张图片是没有我们要的属性的，所以我们想把它经过encoder出来的和u对应的那部分尽量接近0

**Xiao-pku-math 2017-07-12 21:29:47**

换句话说，我们把这个没有属性的图看成一个原点一样的东西

**Gapeng-北京大学-CV 2017-07-12 21:30:59**

你介绍一下loss是怎么设计的吧

**Gapeng-北京大学-CV 2017-07-12 21:31:04**

这样比较好理解

**Xiao-pku-math 2017-07-12 21:31:18**

其他的loss比较自然

**郑华滨-中山大学-深度学习 2017-07-12 21:31:35**

上图的属性是”笑“吗？

**郑华滨-中山大学-深度学习 2017-07-12 21:31:45**

我想确认一下

**Meow-XJTU-CV 2017-07-12 21:31:49**

直观是嘴巴部分

**Xiao-pku-math 2017-07-12 21:31:49**

reconstruction loss就是我们需要把Au拆分开成A和u，然后再把A和u组装回去

**Xiao-pku-math 2017-07-12 21:32:02**

是的

**Xiao-pku-math 2017-07-12 21:32:15**

计算重构误差

**Meow-XJTU-CV 2017-07-12 21:32:43**

A是背景，u是前景，反正是分开了我们想分开的部分

**Gapeng-北京大学-CV 2017-07-12 21:33:19**

我记得文章里面还有个平行四边形loss

**上海滩文哥-复旦-CV 2017-07-12 21:33:21**

试过背景比这个celeba更复杂的图么？

**Gapeng-北京大学-CV 2017-07-12 21:33:25**

这个是什么？

**Xiao-pku-math 2017-07-12 21:33:26**

交换之后得到的A0和B0用一个D来判别

**一路顺枫-THU-Hash 2017-07-12 21:33:48**

我觉得平行四边形loss是区分前景背景的关键。。。

**一路顺枫-THU-Hash 2017-07-12 21:33:57**

不知是不是？

**Xiao-pku-math 2017-07-12 21:34:03**

就是说生成出来的A0应该尽可能是在不笑的那个domain里

**Xiao-pku-math 2017-07-12 21:34:13**

平行四边形loss is optional

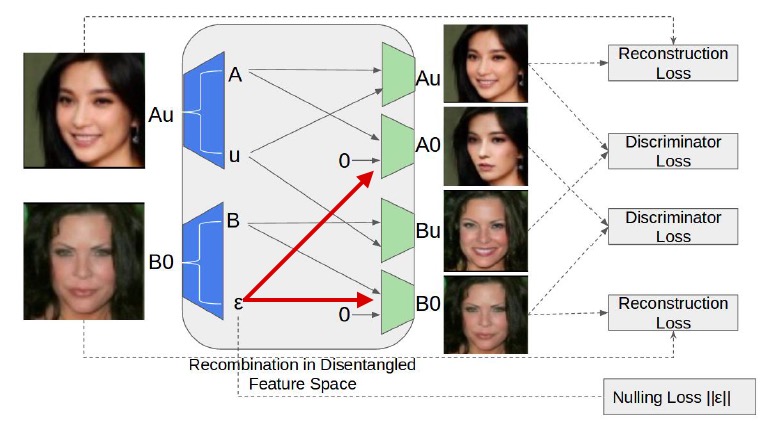
**Gapeng-北京大学-CV 2017-07-12 21:34:22**

我觉得nulling loss才是区分前景和背景的关键

**泛艺术范儿-西电-DL 2017-07-12 21:34:45**

对啊 啥事nulling loss？

**郑华滨-中山大学-深度学习 2017-07-12 21:34:57**



**郑华滨-中山大学-深度学习 2017-07-12 21:35:03**

为什么不这样设计呢？

**Xiao-pku-math 2017-07-12 21:35:22**

其实一开始就是这样，github上的图就是这样

**郑华滨-中山大学-深度学习 2017-07-12 21:35:29**

是实际训练时效果不好吗

**Gapeng-北京大学-CV 2017-07-12 21:35:37**

一张图有某种属性，一张没有某种属性，所以让没有那种属性的前景趋于0

**尹伟东 复旦 深度学习 2017-07-12 21:35:37**

这样test时候会需要两张图片才能交换属性？

**郑华滨-中山大学-深度学习 2017-07-12 21:35:47**

嗯嗯，因为这样设计应该是更直接的

**Xiao-pku-math 2017-07-12 21:35:50**

其实效果和这个是差不多的，因为学好了之后epsilon非常小，和直接加0没有什么太大区别

**bearbee 清华大学 GAN 2017-07-12 21:35:50**

感觉可能不是区分前景还是背景，是不是可能这样，只是通过最后的重构误差判断，如果重构后的背景还是原来的，那么重构误差比较小？

**郑华滨-中山大学-深度学习 2017-07-12 21:36:34**

那为何还是分开了epsilon和0，有什么讲究吗

**上海滩文哥-复旦-CV 2017-07-12 21:36:55**

@bearbee 清华大学 GAN 我同意这种说法

**尹伟东 复旦 深度学习 2017-07-12 21:37:46**

我还有个问题，那两个decoder是share weights么

**Xiao-pku-math 2017-07-12 21:37:46**

重构误差是就想cyclegan一样，保证能够复原原图，这样会是训练更快

**尹伟东 复旦 深度学习 2017-07-12 21:38:00**

说错了，两个encoder

**Xiao-pku-math 2017-07-12 21:38:08**

对，encoder和decoder都是

**Xiao-pku-math 2017-07-12 21:38:27**

只有一个encoder和一个decoder

**Gapeng-北京大学-CV 2017-07-12 21:38:27**

实际上只有一个encoder和decoder吧

**Xiao-pku-math 2017-07-12 21:38:40**

同时对两边的图操作

**泛艺术范儿-西电-DL 2017-07-12 21:38:50**

如何保证encoder出来的是distangled feature space?

**尹伟东 复旦 深度学习 2017-07-12 21:38:55**

哦哦，这样能同时改很多属性吧

**bearbee 清华大学 GAN 2017-07-12 21:39:18**

这种训练时需要两张图片一起是吧？

**Gapeng-北京大学-CV 2017-07-12 21:39:30**

没有要求disentangle吧，只是把一种属性给分离出来

**郑华滨-中山大学-深度学习 2017-07-12 21:39:35**

感觉应该是靠这里面各种loss把两组向量disentangle的

**bearbee 清华大学 GAN 2017-07-12 21:39:41**

如果两张图片的背景不一样怎么办？

**郑华滨-中山大学-深度学习 2017-07-12 21:39:44**

就是指分离

**Meow-XJTU-CV 2017-07-12 21:39:53**

这个属性有标签吗

**尹伟东 复旦 深度学习 2017-07-12 21:40:05**

感觉有点类似fader里面的做法

**郑华滨-中山大学-深度学习 2017-07-12 21:40:08**

肯定有吧

**泛艺术范儿-西电-DL 2017-07-12 21:40:53**

encoder出来的维度是多少？

**Meow-XJTU-CV 2017-07-12 21:40:53**

我不太理解实际用标签是怎么划定某个属性（部位）的

**Xiao-pku-math 2017-07-12 21:40:58**

没要求配对的数据，只要你想换的属性在两组图中是不一样的，只有0/1标签

**郑华滨-中山大学-深度学习 2017-07-12 21:42:12**

我有个很关键的问题，这里面u和epsilon代表的是不是有 / 没有目标属性

**Gapeng-北京大学-CV 2017-07-12 21:42:20**

celeba的标签非常多，而且都是连续值，按照阈值划分成两类数据就行了吧

**Xiao-pku-math 2017-07-12 21:42:21**

当然实际上要做的效果比较好需要做一些对齐的操作

**Xiao-pku-math 2017-07-12 21:42:40**

是的@郑华滨-中山大学-深度学习

**Gapeng-北京大学-CV 2017-07-12 21:43:19**

还有问题吗？有问题快提问

**郑华滨-中山大学-深度学习 2017-07-12 21:43:43**

那么之前cyclegan会不小心把背景也给变了，其实是不是可以认为，是它不小心把背景也认为是”斑马“这个属性的特征之一了

**Gapeng-北京大学-CV 2017-07-12 21:43:59**

没看过论文的建议先看一遍论文，这个文章没看过还是比较难参与讨论的

**郑华滨-中山大学-深度学习 2017-07-12 21:44:06**

那为什么这里不会也不小心把背景认为是”有在笑“这个属性的特征之一呢

**郑华滨-中山大学-深度学习 2017-07-12 21:45:00**

关键在于我们没告诉它究竟这个图何以成为”在笑”，那个图何以成为“没笑”，那它什么都可能学到

**Gapeng-北京大学-CV 2017-07-12 21:45:15**

我觉得原因在于对齐操作

**郑华滨-中山大学-深度学习 2017-07-12 21:45:20**

万一“在笑“的图片集背景全部蓝色，”没笑“的图片集背景全部红色

**郑华滨-中山大学-深度学习 2017-07-12 21:45:35**

那会不会把颜色认为是”在笑“的特征之一？

**Meow-XJTU-CV 2017-07-12 21:45:35**

我刚也是这个问题，如何划定属性所描述的范围

**上海滩文哥-复旦-CV 2017-07-12 21:45:41**

celeba的数据背景都比较干净，在imagenet上的效果咋样？

**一路顺枫-THU-Hash 2017-07-12 21:46:08**

imagenet有属性标签吗？

**上海滩文哥-复旦-CV 2017-07-12 21:46:09**

还能有效的区分前后背景么？

**上海滩文哥-复旦-CV 2017-07-12 21:46:18**

或者VOC

**郑华滨-中山大学-深度学习 2017-07-12 21:46:29**

当然看论文结果是有效果的，但是我很好奇为什么可以做到[捂脸]

**泛艺术范儿-西电-DL 2017-07-12 21:46:41**

celeba这个数据是已经crop和对齐过的吧

**Xiao-pku-math 2017-07-12 21:46:43**

你说的有道理，数据背景各式各样，相当于有相互抵消的作用

**Gapeng-北京大学-CV 2017-07-12 21:46:57**

@郑华滨-中山大学-深度学习 要是这样的话，最优的解肯定会把蓝色也当成需要交换的属性的一部分

**郑华滨-中山大学-深度学习 2017-07-12 21:47:10**

嗯，如果各式各样那就不会被认为是属性之一

**郑华滨-中山大学-深度学习 2017-07-12 21:47:39**

但普京和斑马那两个case中，图片集应该也是有各种各样背景的吧

**郑华滨-中山大学-深度学习 2017-07-12 21:47:46**

为什么它们那两个case就会有问题

**Xiao-pku-math 2017-07-12 21:47:58**

我认为是设计的问题

**郑华滨-中山大学-深度学习 2017-07-12 21:48:11**

嗯嗯怎么说

**Xiao-pku-math 2017-07-12 21:48:38**

cyclegan的设计看出他比较适合做全局的某种属性或者风格的迁移

**Xiao-pku-math 2017-07-12 21:49:23**

而genegan之所以能学到不同的属性是因为从两组不同属性的数据中分离开的

**Xiao-pku-math 2017-07-12 21:50:10**

genegan的encoder设计就是在channel中强制把前景和背景分开

**Gapeng-北京大学-CV 2017-07-12 21:51:30**

大家还有问题吗？

**Gapeng-北京大学-CV 2017-07-12 21:51:54**

没有问题的话，我们邀请另一位同学来分享他的工作

**郑华滨-中山大学-深度学习 2017-07-12 21:52:05**

等下哈

**郑华滨-中山大学-深度学习 2017-07-12 21:52:14**

论文中的u的向量维度是不是很小

**Meow-XJTU-CV 2017-07-12 21:52:25**

我才注意到xiao是GeneGan的作者[捂脸]

**郑华滨-中山大学-深度学习 2017-07-12 21:52:31**

@Xiao-pku-math

**Gapeng-北京大学-CV 2017-07-12 21:52:44**

他说了是channel

**一路顺枫-THU-Hash 2017-07-12 21:52:51**

我有两个问题，实验只做了64\*64的，如果做128\*128的呢？

**郑华滨-中山大学-深度学习 2017-07-12 21:52:52**

如果u的维度向量很小，那就可以解释为什么背景保留了

**郑华滨-中山大学-深度学习 2017-07-12 21:53:04**

或者说u的channel数目很小

**Xiao-pku-math 2017-07-12 21:53:04**

u实际中是取了那一层feature maps中的1/4的channel

**一路顺枫-THU-Hash 2017-07-12 21:53:20**

第二，为什么总wgan，用dcgan或者lsgan或者improved wgan其他方法可以吗？

**一路顺枫-THU-Hash 2017-07-12 21:53:30**

为什么用wgan

**郑华滨-中山大学-深度学习 2017-07-12 21:53:44**

那我觉得这可能是原因，因为u容量小，迫使它只能包含最salient的信息

**一路顺枫-THU-Hash 2017-07-12 21:54:08**

另外，我感觉运行出来的图片质量依然不是太好。。。

**郑华滨-中山大学-深度学习 2017-07-12 21:54:23**

如果把u容量增大，就有可能把其他次salient的特征（如背景）给encode进来

**郑华滨-中山大学-深度学习 2017-07-12 21:54:47**

因为你刚刚提到说在channel中强制分开，我就在想是不是这个分流的channel数目大小有影响

**Xiao-pku-math 2017-07-12 21:55:02**

我觉得是会有这种可能的，不过我没有实验

**郑华滨-中山大学-深度学习 2017-07-12 21:55:08**

不知道你们有没有做这个实验，就是让u的channel数目大得多，看看还能不能保留背景

**Gapeng-北京大学-CV 2017-07-12 21:55:14**

嗯嗯，这也是一个trick

**郑华滨-中山大学-深度学习 2017-07-12 21:55:20**

值得一试[奸笑]

**Xiao-pku-math 2017-07-12 21:55:30**

回头可以试试，谢谢建议

**Gapeng-北京大学-CV 2017-07-12 21:55:36**

回头试试看

**郑华滨-中山大学-深度学习 2017-07-12 21:55:38**

因为就好像PCA一样，容量不够的话，就只能保留下主成分了

**郑华滨-中山大学-深度学习 2017-07-12 21:55:45**

嗯嗯

**Xiao-pku-math 2017-07-12 21:55:58**

讲到维度，其实还有一点没有说

**一路顺枫-THU-Hash 2017-07-12 21:56:02**

想问下你们做没做128\*128的图片呢？

**Xiao-pku-math 2017-07-12 21:56:35**

有做，不过不是用的tf

**泛艺术范儿-西电-DL 2017-07-12 21:56:43**

A和u的通道数目前是怎么分的呢

**郑华滨-中山大学-深度学习 2017-07-12 21:57:03**

3：1吧，上面说了

**泛艺术范儿-西电-DL 2017-07-12 21:57:26**

哦， sorry 没看到

**Xiao-pku-math 2017-07-12 21:57:31**

接着说刚才没说的

**一路顺枫-THU-Hash 2017-07-12 21:58:00**

图片像素增大对生成结果有多大影响？

**Xiao-pku-math 2017-07-12 21:58:21**

就是还有重要的一点，其他一些换属性的方法，生成出来的比如说眼镜是只有同一个，缺少diversity

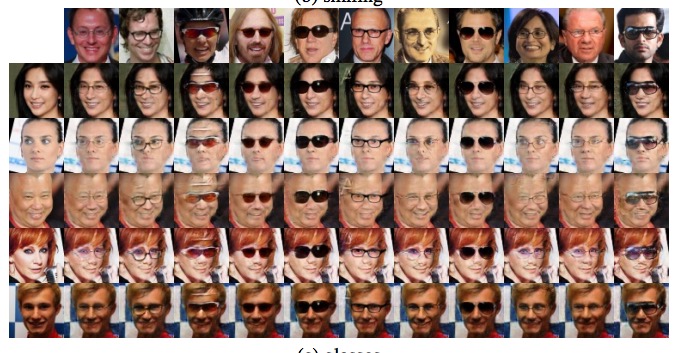
**Gapeng-北京大学-CV 2017-07-12 21:59:26**

像cyclegan，其实加入noise是没有用的，最后实验的结果，noise几乎被忽略了

**Gapeng-北京大学-CV 2017-07-12 21:59:54**

这个我做实验发现是这样

**Xiao-pku-math 2017-07-12 22:00:32**



**一路顺枫-THU-Hash 2017-07-12 22:00:36**

什么叫加入noise？

**泛艺术范儿-西电-DL 2017-07-12 22:00:41**

gan生成的结果是不是普遍缺乏多样性

**郑华滨-中山大学-深度学习 2017-07-12 22:01:06**

很赞的效果

**郑华滨-中山大学-深度学习 2017-07-12 22:01:15**

不过多样性的随机性来源是什么？

**郑华滨-中山大学-深度学习 2017-07-12 22:01:30**

从上面那副架构图看不出来，好像都是确定性映射？

**Xiao-pku-math 2017-07-12 22:01:30**

我们把交换的眼镜也保留下来了

**Lynn-中科院计算所-CV 2017-07-12 22:01:51**

这个眼镜保留特别喜欢

**Xiao-pku-math 2017-07-12 22:03:15**

其实靠的是从图像内部分离出的属性，可能某种意义上有id信息

**Lynn-中科院计算所-CV 2017-07-12 22:03:18**

保留眼镜设计的loss也简洁，666

**郑华滨-中山大学-深度学习 2017-07-12 22:03:48**

其实好像除了眼镜还是transfer了一些别的东西

**Xiao-pku-math 2017-07-12 22:03:50**

而其他的一些方法是靠G生成的，不太容易做到这点

**郑华滨-中山大学-深度学习 2017-07-12 22:03:54**

虽然很不明显

**Guo-Jun Qi 2017-07-12 22:04:09**

@Gapeng-北京大学-CV cyclegan这样操作attributes 的GAN只是对和attributes 相关的部分进行操作，所以输入noise对图像大部分是没有影响的

**Gapeng-北京大学-CV 2017-07-12 22:04:30**

是的

**Guo-Jun Qi 2017-07-12 22:04:34**

这是我的理解

**Xiao-pku-math 2017-07-12 22:04:50**

@郑华滨-中山大学-深度学习 ，比方你要换胡子那么和男变女就扯得很近了

**郑华滨-中山大学-深度学习 2017-07-12 22:05:01**

仔细看上面那幅图，其实也有一部分脸型被transfer过去了

**Xiao-pku-math 2017-07-12 22:05:10**

有些属性还是耦合的比较厉害

**Xiao-pku-math 2017-07-12 22:05:49**

对，这个是值得改进的地方

**郑华滨-中山大学-深度学习 2017-07-12 22:05:49**

不过感觉还好，只是一部分信息串channel了

**Gapeng-北京大学-CV 2017-07-12 22:06:17**

嗯嗯，达到这个效果已经挺好了

**Xiao-pku-math 2017-07-12 22:06:32**

我们实验中还发现黑框眼镜（glasses）和长者（senior）也耦合的比较厉害

**Xiao-pku-math 2017-07-12 22:06:36**

你懂的

**郑华滨-中山大学-深度学习 2017-07-12 22:06:38**

直觉还是告诉我channel数目容量的影响很大[捂脸]

**郑华滨-中山大学-深度学习 2017-07-12 22:06:54**

蛤？

**Gapeng-北京大学-CV 2017-07-12 22:07:15**

。。。。。。好了，我们抓紧时间，GeneGAN的讨论暂且到这里，下面我们有请@尹伟东 复旦 深度学习 跟我们分享Semi-GAN

**郑华滨-中山大学-深度学习 2017-07-12 22:07:43**

感谢@Xiao-pku-math 分享！

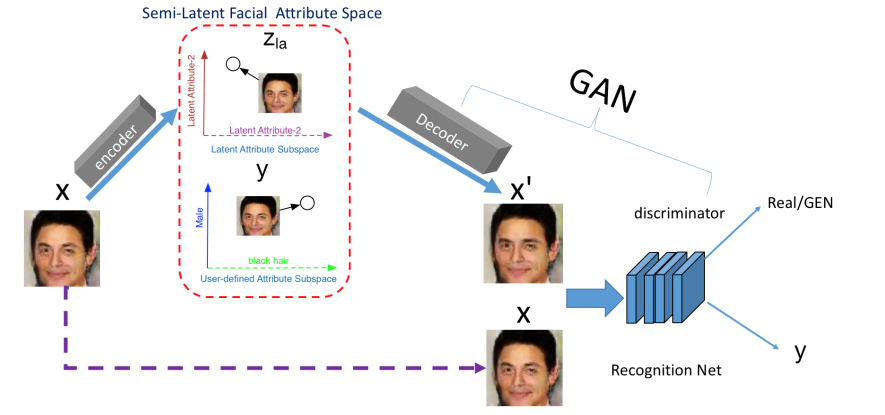
**尹伟东 复旦 深度学习 2017-07-12 22:07:46**

时间也不多了，我就简单讲一下我们的idea。。

**Gapeng-北京大学-CV 2017-07-12 22:07:49**

@尹伟东 复旦 深度学习 Semi-GAN是什么？它是用来做一个什么样的任务？

**尹伟东 复旦 深度学习 2017-07-12 22:08:19**



**尹伟东 复旦 深度学习 2017-07-12 22:09:12**

我们是提出这样一个模型，用来从attribute同时生成人脸和修改人脸

**尹伟东 复旦 深度学习 2017-07-12 22:09:57**

基本思路是把人脸图片encode到一个latent space z, 然后修改属性y就可以修改原来的图片

**Gapeng-北京大学-CV 2017-07-12 22:10:01**

看图我不太理解怎么做到属性提取的

**Gapeng-北京大学-CV 2017-07-12 22:10:43**

需要怎么设计loss才能把属性y给分离出来

**尹伟东 复旦 深度学习 2017-07-12 22:10:49**

我们没有提取属性，是直接加在外面的

**郑华滨-中山大学-深度学习 2017-07-12 22:11:01**

群主求paper链接

**尹伟东 复旦 深度学习 2017-07-12 22:11:40**

属性y是通过一个recognition network加上去

**尹伟东 复旦 深度学习 2017-07-12 22:11:47**

类似infoGAN

**Guo-Jun Qi 2017-07-12 22:11:49**

"用one-hot 编码属性？  
"

**郑华滨-中山大学-深度学习 2017-07-12 22:11:53**

concat？

**尹伟东 复旦 深度学习 2017-07-12 22:11:56**

是的

**Toxic-sjtu-CV 2017-07-12 22:11:59**

前半部分看图感觉有点像VAE？

**尹伟东 复旦 深度学习 2017-07-12 22:12:21**

嗯嗯，我们用的VAE，这样就能同时生成和修改图片

**尹伟东 复旦 深度学习 2017-07-12 22:12:36**

因为latent space上面是一个gaussian分布

**尹伟东 复旦 深度学习 2017-07-12 22:13:01**

我介绍一下各个网络的loss function吧

**尹伟东 复旦 深度学习 2017-07-12 22:13:29**

encoder 是用的vae，kl loss + reconstruction loss

**Gapeng-北京大学-CV 2017-07-12 22:13:29**

arxiv id

**泛艺术范儿-西电-DL 2017-07-12 22:13:31**

VAE有助于产生distangled的编码

**Gapeng-北京大学-CV 2017-07-12 22:13:35**

@郑华滨-中山大学-深度学习 1704.02166

**尹伟东 复旦 深度学习 2017-07-12 22:14:08**

这里的reconstruction loss是用的discriminator的feature mao

**尹伟东 复旦 深度学习 2017-07-12 22:14:11**

map

**Toxic-sjtu-CV 2017-07-12 22:14:23**

嗯嗯

**尹伟东 复旦 深度学习 2017-07-12 22:14:44**

然后decoder loss是 reconstruction + advsersarial + recognition loss

**尹伟东 复旦 深度学习 2017-07-12 22:15:06**

discriminator loss 是 adversarial loss

**尹伟东 复旦 深度学习 2017-07-12 22:15:47**

recognition network 是 discriminator 后面接的一层全连接 ,

**尹伟东 复旦 深度学习 2017-07-12 22:16:05**

后面的loss就是true imagerecognition loss

**Gapeng-北京大学-CV 2017-07-12 22:16:08**

跟infogan的设计很像

**尹伟东 复旦 深度学习 2017-07-12 22:16:17**

对的

**尹伟东 复旦 深度学习 2017-07-12 22:16:41**

这样我们想把y map到原图的属性上去

**尹伟东 复旦 深度学习 2017-07-12 22:16:51**

然后z代表原图的背景

**尹伟东 复旦 深度学习 2017-07-12 22:17:03**

修改y就可以改原来的图片了

**尹伟东 复旦 深度学习 2017-07-12 22:17:22**

从gaussian里面 sample z就能生成图片

**尹伟东 复旦 深度学习 2017-07-12 22:17:43**

模型差不多就是这样

**Gapeng-北京大学-CV 2017-07-12 22:17:47**

y任意修改会有平滑的变换效果吗？

**陆鹏起-hust-gan 2017-07-12 22:18:50**

可以interpolation吗

**尹伟东 复旦 深度学习 2017-07-12 22:19:27**

理论上是有的，但是我们实验做的效果不太好，可能需要修改一些网络结构

**陆鹏起-hust-gan 2017-07-12 22:19:27**

类似fader network这样的

**Gapeng-北京大学-CV 2017-07-12 22:21:47**

大家还有什么问题吗？

**上海滩文哥-复旦-CV 2017-07-12 22:23:50**

还没看过paper，看完之后再问可以么

**Gapeng-北京大学-CV 2017-07-12 22:24:10**

可以

**Gapeng-北京大学-CV 2017-07-12 22:24:25**

作者在群里，看完文章有问题随时可以问

**Meow-XJTU-CV 2017-07-12 22:24:43**

谢谢各位大佬分享

**Jeffrey - ANU - 深度贝叶斯 2017-07-12 22:24:45**

讨论氛围真好，[嘴唇][嘴唇]

**Gapeng-北京大学-CV 2017-07-12 22:24:48**

so long，感谢大家的参与

**Gapeng-北京大学-CV 2017-07-12 22:25:47**

好了，又到了抓同学写总结（提升姿势水平）的时候了

**Gapeng-北京大学-CV 2017-07-12 22:26:14**

今天的讨论分成两个主题，我们按照两个主题来找同学写总结

**Gapeng-北京大学-CV 2017-07-12 22:27:09**

@SURE-天大-超分辨率-三星-人体姿态 你来写超分重建那部分总结，可以么？

**SURE-天大-超分辨率-三星-人体姿态 2017-07-12 22:28:12**

急么 六日有时间写

**Gapeng-北京大学-CV 2017-07-12 22:28:37**

可以，周日前给我吧

**Gapeng-北京大学-CV 2017-07-12 22:28:45**

我一会儿把讨论的记录导出来

**Gapeng-北京大学-CV 2017-07-12 22:32:00**

图像属性修改那部分的总结，交给两位分享的作者来写