Pix2pix总结

姓名：张溢炉 学号：1912392039 专业班级：控制工程194班

1. **论文总结**
2. 摘要

在基于CNN的pix2pix任务中，一般不同的任务需要设置不同的损失函数。而损失函数是将期望的成像效果以数学的方式表示出来，往往很难找到合适的损失函数准确的表示成像效果的期望。本文提出了利用cGANs（conditional Generative Adversarial Networks）的方法能够自动设置损失函数，且适用于多种任务。

1. 方法总结

在已有的文献中，解决pix2pix问题主要存在3种方法。

1. 基于CNNs(convolutional neural nets)的方法：与一般的深度学习方法一样，它需要自己设置合适的损失函数，然后通过梯度下降等方法训练不断减小损失函数的损失。其存在一个开放性的问题是如何设置合适的损失函数表达真实的期望，比如使用欧几里德距离比较输出图片和ground truth图片的像素距离，看似合理，但是它忽略了各像素之间的关联性，会造成模糊。
2. GANs（Generative Adversarial Networks）：GAN结构如图1-1，整个网络由生成器和判别器组成。生成器输入是随机噪声z，生成器输出G(z)；判别器输入合成数据、G(z)或真实数据x，判别器输出判断结果。整个网络的目的是让判别器分不清G(z)和真实数据x，判别器类似于图灵测试，当G(z)和真实数据x真假难辨时说明生成结果G(z)已经很好了。最后使用生成器完成pix2pix任务。

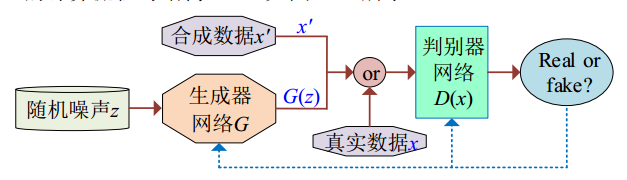


图1- 1 GANS结构图

1. cGANs（conditional Generative Adversarial Networks）：cGANS结构如图1-2。其与GANs的不同之处在其输入是勾画好的图片（如物体的轮廓）加上一定噪声。显然对于文中提到输入不完全是噪声的任务,cGANS更合适。

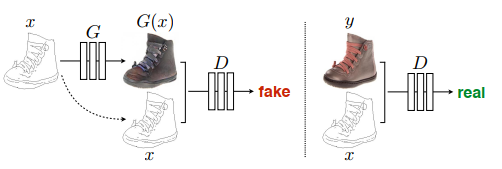
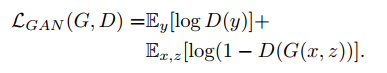


图1- 2 cGANs结构图

1. 本文主要贡献
2. 提出了适合多种任务的cGANs网络；
3. 网络结构简单且能获得好结果，分析了几个重要结构选择的方法。
4. 本文方法
5. 目标函数如下，其中加入L1约束项能使结果更好，又不至于像L2一样产生太多模糊。

其中，



1. 生成器和判别器使用文献[44]提出的结构。其中，Encoder-decoder中采用了像文献[50]一样跳过正中间层，如图1-3，这有利于低层信息直接在相连网络层上传递。另外，因为L1不能很好建模高频结构，故增加PatchGAN来建模高频结构。

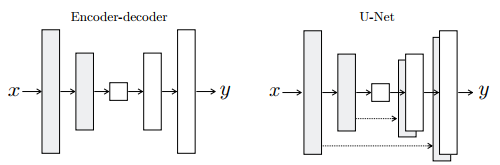
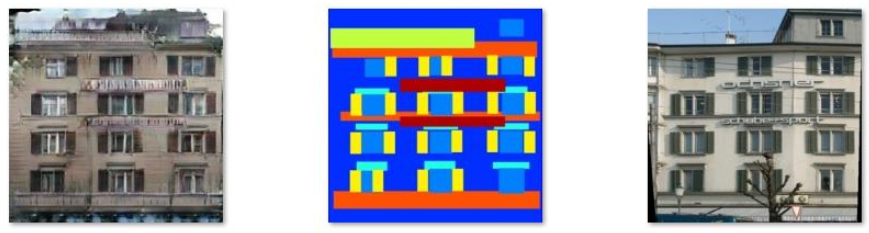
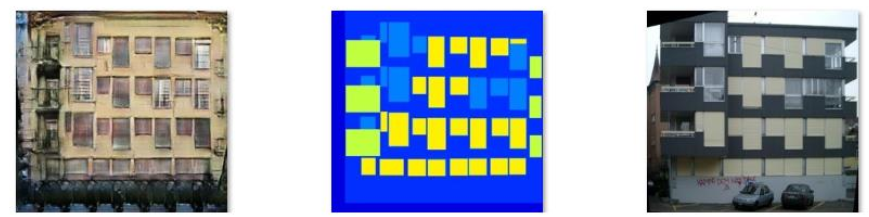


图1- 3 跳线结构图

1. 对整个网络的优化使用文献[24]和[32]提到的方法。结果评价方式采用文献[52]中方法。
2. **实验总结**
3. 实验结果

生成器输出图片 生成器输入图片 Ground truth图片

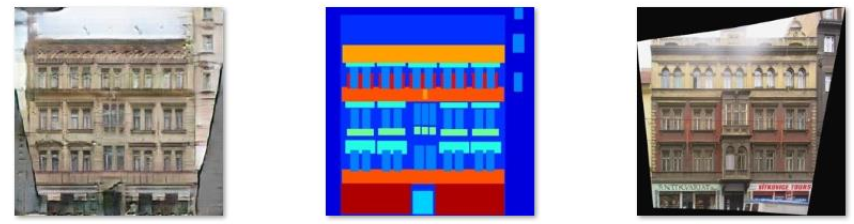


图2- 1

图2- 2

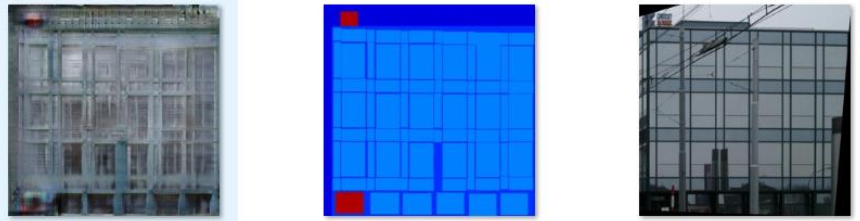


图2- 3

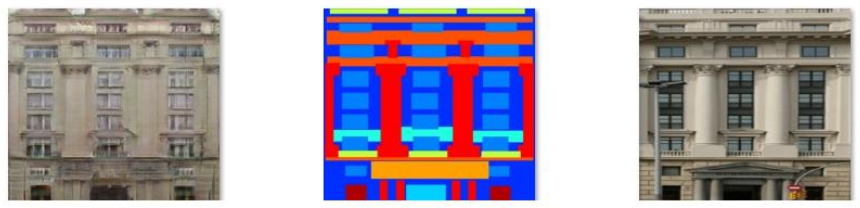


图2- 5

图2- 4

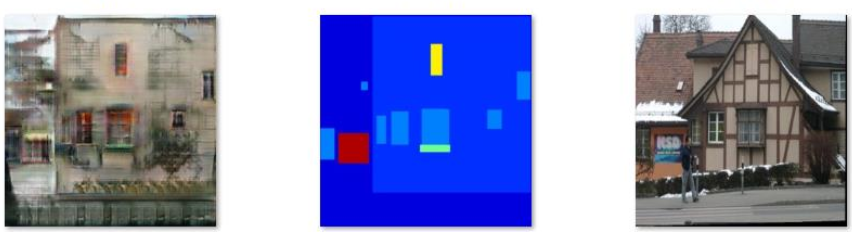
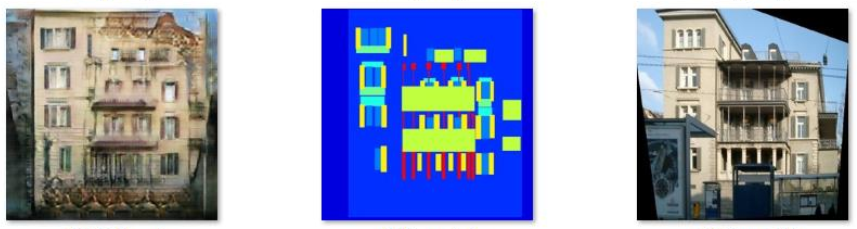


图2- 6

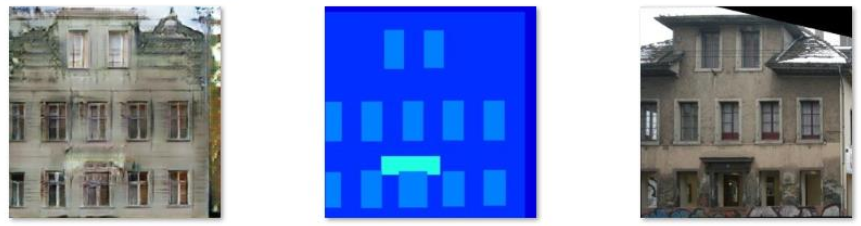


图2- 7

图2- 8

1. 结果分析

通过提取实验中8组图片进行简单分析总结可知，生成图片效果较好，很少模糊，模型比较适合生成规则的房体结构。将生成图片与ground truth图片对比可得以下4点。

1. 从所有图片来看，生成图片和ground truth图片颜色有明显差别；
2. 生成图片对ground truth图片细节恢复得不太好，如图2-1、2-3中得英文字母不能恢复；
3. 该网络比较适合恢复整体和局部是长方形的规则物体，如图2-6、2-7、2-8，特别是图2-6，直接将三角形变成了长方形，且存在严重模糊；
4. 比较适合恢复玻璃和混泥土的结构，如图2-4中的全玻璃结构变成了玻璃和混泥土的结构。

对于以上总结的1）、2）点，可能需要进一步完善模型结构解决，而对于以上总结的3）、4）点，可能是数据集中该类样本不足造成的。