实验报告

1.任务说明

把手绘转化成实物。即输入黑白的手绘，输出彩色的实物图，类似于风景图片填色。

项目的Github地址：

2.数据处理

由于我们使用生成式对抗网络GAN，故我们只需收集数据集，无需对数据进行标注。

由 于目标是训练一个能从手绘的机车图生成实际机车图的网络。我们广泛寻找了网上的汽车实物图，总计共2000张。用photoshop的素描滤镜功能生成对应的手绘图并拼接在一起，作为训练集和测试集。

我们输入256\*256的手绘图，也输出256\*256的实物图。

3.算法原理

我们使用了生成式对抗网络GAN（Generative Adversarial Networks ）。

模型通过框架中（至少）两个模块：生成模型（Generative Model）和判别模型（Discriminative Model）的互相博弈学习产生相当好的输出。

生成网络G是一个生成图片的网络，它接收一个随机的噪声z，通过这个噪声生成图片，记做G(z)。

判别网络D是一个判别网络，判别一张图片是不是“真实的”。它的输入参数是x，x代表一张图片，输出D（x）代表x为真实图片的概率，如果为1，就代表100%是真实的图片，而输出为0，就代表不可能是真实的图片。

在训练过程中，生成网络G的目标就是尽量生成真实的图片去欺骗判别网络D。而D的目标就是尽量把G生成的图片和真实的图片分别开来。这样，G和D构成了一个动态的“博弈过程”。

最后博弈的结果在最理想的状态下，对生成网络G来说，它可以生成足以“以假乱真”的图片G(z)。对于判别网络D来说，它难以判定G生成的图片究竟是不是真实的，因此判断真实图片的概率D(G(z)) = 0.5。

这样，我们就得到了一个可以用来生成图片的生成式模型G。

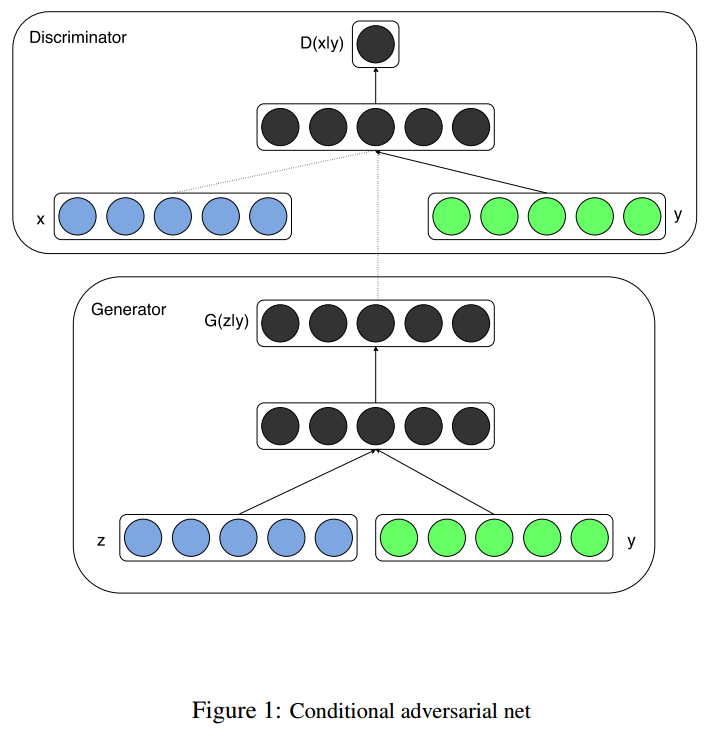
然而，这种不需要预先建模的方法缺点是太过自由了，对于较大的图片，较多的 pixel的情形，基于简单 GAN 的方式就不太可控了。为了解决GAN太过自由这个问题，一个很自然的想法是给GAN加一些约束，于是便有了Conditional Generative Adversarial Nets（CGAN）。

这项工作提出了一种带条件约束的GAN，在生成模型（D）和判别模型（G）的建模中均引入条件变量y（conditional variable y），使用额外信息y对模型增加条件，可以指导数据生成过程。这些条件变量y可以基于多种信息，例如类别标签，用于图像修复的部分数据，来自不同模态（modality）的数据。如果条件变量y是类别标签，可以看做CGAN 是把纯无监督的 GAN 变成有监督的模型的一种改进。

条件生成式对抗网络（CGAN）是对原始GAN的一个扩展，生成器和判别器都增加额外信息 yy为条件, yy 可以使任意信息,例如类别信息,或者其他模态的数据。如 Figure 1 所示，通过将额外信息 yy 输送给判别模型和生成模型,作为输入层的一部分,从而实现条件GAN。在生成模型中,先验输入噪声 p(z)p(z) 和条件信息 yy 联合组成了联合隐层表征。对抗训练框架在隐层表征的组成方式方面相当地灵活。类似地，条件 GAN 的目标函数是带有条件概率的二人极小极大值博弈（two-player minimax game ）：

公式

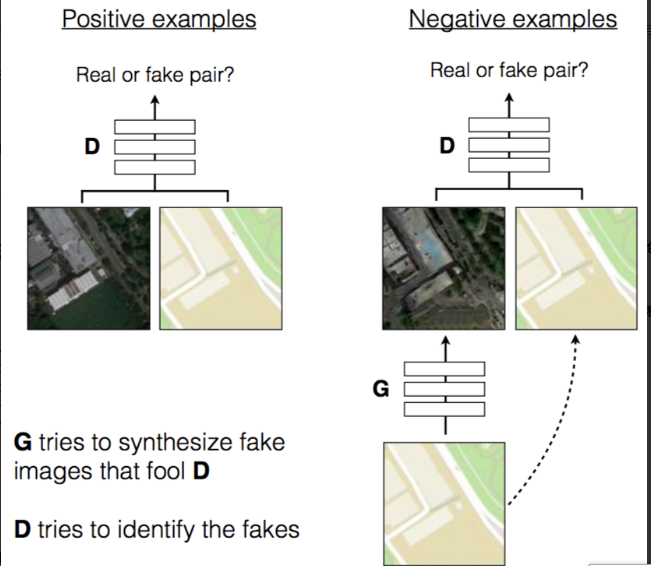
CGAN的网络结构如下图示意：



基于CGAN的pix2pix模型，可用于实现多种配对图像翻译任务。配对图像翻译包括很多应用场景，输入和输出都是图片且尺寸相同。

pix2pix原理如下，它具有典型的CGAN结构，但G只接受一个固定的输入X，可以理解为一个条件C，即不需要随机噪音，然后输出翻译后的版本Y。

D接受一个X（CGAN中的C）和一个Y（真假样本），并判断X和Y是否为配对的翻译。



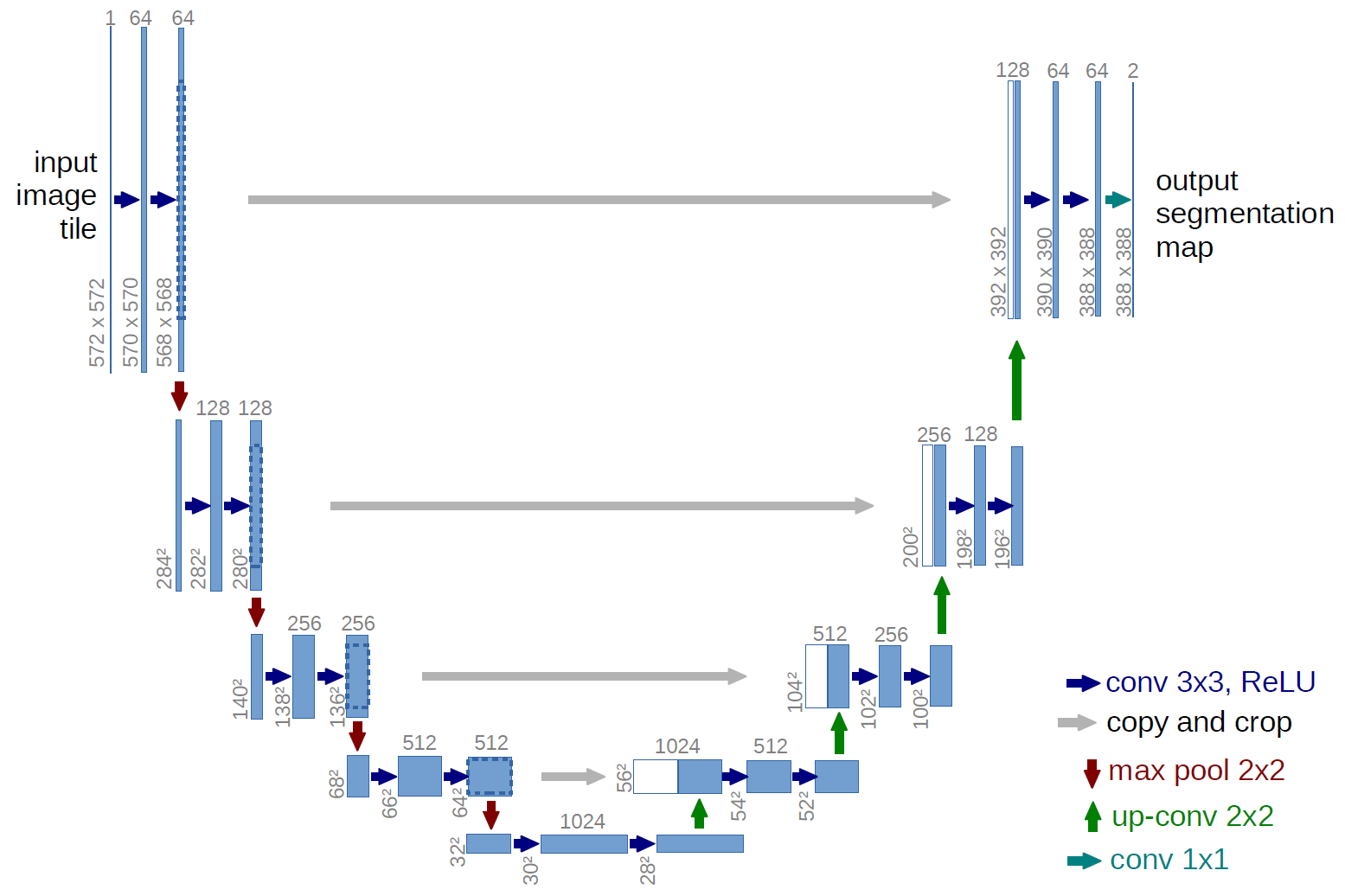
除了标准的GAN损失函数之外，pix2pix还考虑了生成样本和真实样本之间的L1距离作为损失：

IMG_256

GAN损失负责捕捉图像高频特征，L1损失负责捕捉低频特征，使得生成结果既真实且清晰。

生成器G使用Unet实现，主要用到Skip-Connection来学习配对图像之间的映射。

判别器D使用了PatchGAN的思想，之前是对整张图片给出一个分数，PatchGAN则是将一张图片分为很多块，对每一块都给出一个分数。



4.实验细节

实际训练过程中的参数调整：lamda

Loss函数：

标准的GAN损失函数：

公式

以及生成样本和真实样本之间的距离L1

IMG_256

可以明显看出，在本地训练一千次的效果与在bitahub上训练100000次的效果有巨大的不同。（此处我们有存图吗？有的话贴一个）

在尚未修改参数时，效果有些奇怪，实物车的车体显得十分虚浮，猜测是（）导致。修改参数后，（）。

经过多次调整后，获得了较好的实物图。

实验过程中，其实也存在熟悉bitahub的使用方法用了较长时间。在制作数据集时，500张图中有一张非标准格式的，导致输入不正确。算是一些无用功吧。

Epoch：100000

指标情况：（再贴个图？）

1. 心得体会

目前网络上缺乏对深度学习初学者比较友好的平台。

1. 分工合作

张浩恩、胡子谦：修改代码，训练网络。

曾益茗：寻找数据集。