

**人工智能概论课程报告**

**题目： 基于生成式对抗神经网络实现车道线检测**

**姓 名**

**学 号**

**年 级**

**专 业**

**教 师**

**学 院 信息工程学院**

2021年 12 月 1 日

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **小组成员姓名** |  |  |  |  |
| **分工** | 代码编写 | 前期准备 | 资料整理 | 代码测试 |

目录

[一、 背景介绍 1](#_Toc27854)

[二、 设计思路 1](#_Toc16526)

[三、 实施步骤 2](#_Toc10262)

[四、 运行结果 4](#_Toc13529)

[五、发现的问题、排除方法和改进措施总结 6](#_Toc16752)

# 背景介绍

近年来，自动驾驶在学术界和工业界都受到了广泛的关注。其中最基本和最具有挑战性的任务是在真实场景中进行车道线检测从而辅助驾驶。但是由于恶劣场景的存在，如遮挡、雾霾、黑暗、强光反射等，准确地进行车道线检测极具挑战性。

车道检测是自动驾驶汽车以及一般计算机视觉的关键组件。这个概念用于描述自动驾驶汽车的路径并避免进入另一条车道的风险。现在国内主要的自动驾驶技术提供商，车厂自研代表有特斯拉、小鹏，Tier 1阵营中，有老玩家百度Apollo、小马智行、以及新入场的华为、大疆、滴滴等。

个人觉得车道线检测其实是个很有意思、具有难度且有需求的研究方向。我们既可以了解传统图像如何处理解决这个问题，也可以通过深度学习来解决。国内做自动驾驶方向的公司很多，而且这个方向与公司业务也是强关联的，认为很有助于找工作。

# 设计思路

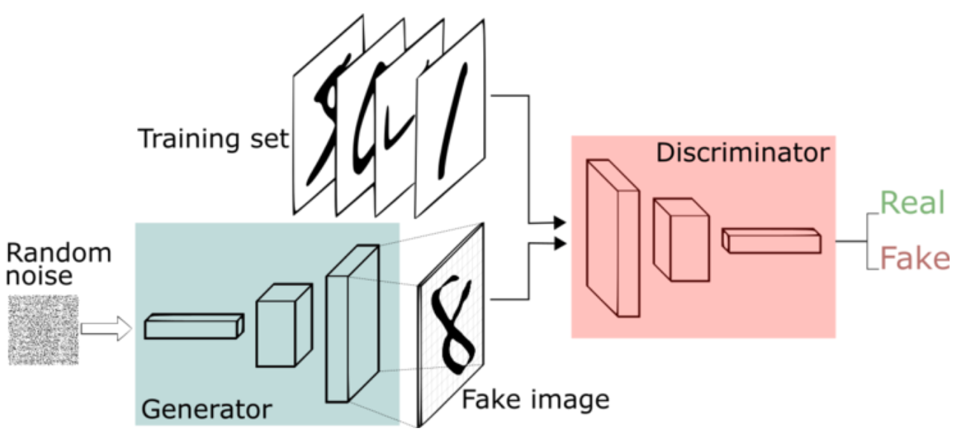


图2-1 原理图

GAN的基本原理其实非常简单，这里以生成图片为例进行说明。假设我们有两个网络，G（Generator）和D（Discriminator）。正如它的名字所暗示的那样，它们的功能分别是：

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)。

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。

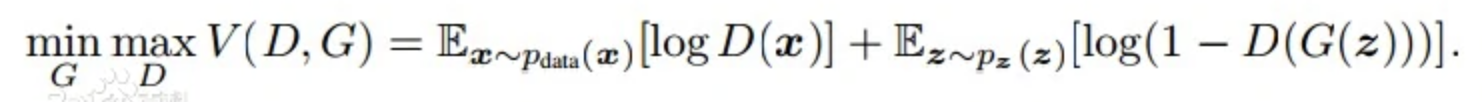
在训练过程中，生成网络G的目标就是尽量生成真实的图片去欺骗判别网络D。而D的目标就是尽量把G生成的图片和真实的图片分别开来。这样，G和D构成了一个动态的“博弈过程”。

最后博弈的结果是什么？在最理想的状态下，G可以生成足以“以假乱真”的图片G(z)。对于D来说，它难以判定G生成的图片究竟是不是真实的，因此D(G(z)) = 0.5。

这 样我们的目的就达成了：我们得到了一个生成式的模型G，它可以用来生成图片。

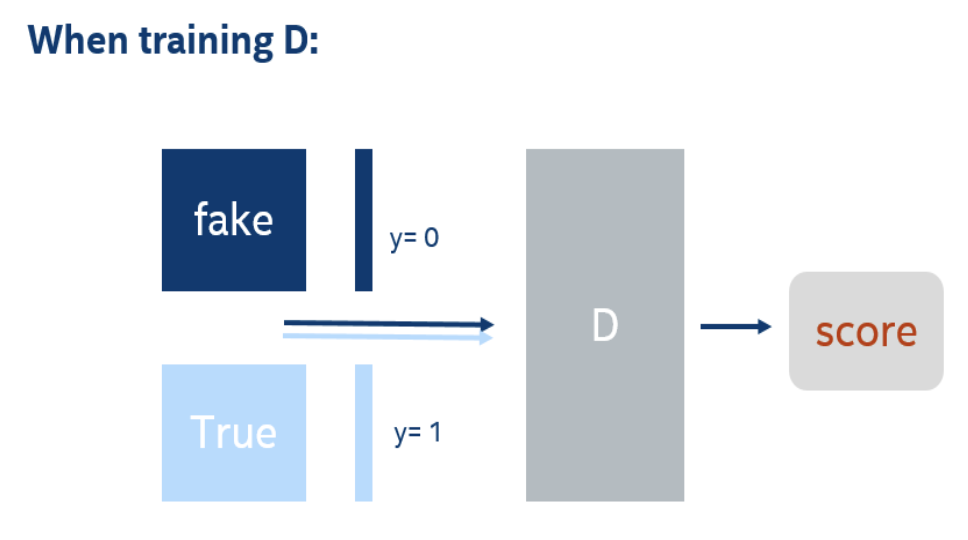
Goodfellow从理论上证明了该算法的收敛性 [1] ，以及在模型收敛时，生成数据具有和真实数据相同的分布（保证了模型效果）。

GAN公式：



公式中x表示真实图片，z表示输入G网络的噪声，G(z)表示G网络生成的图片，D(·)表示D网络判断图片是否真实的概率。

根据GAN的算法：



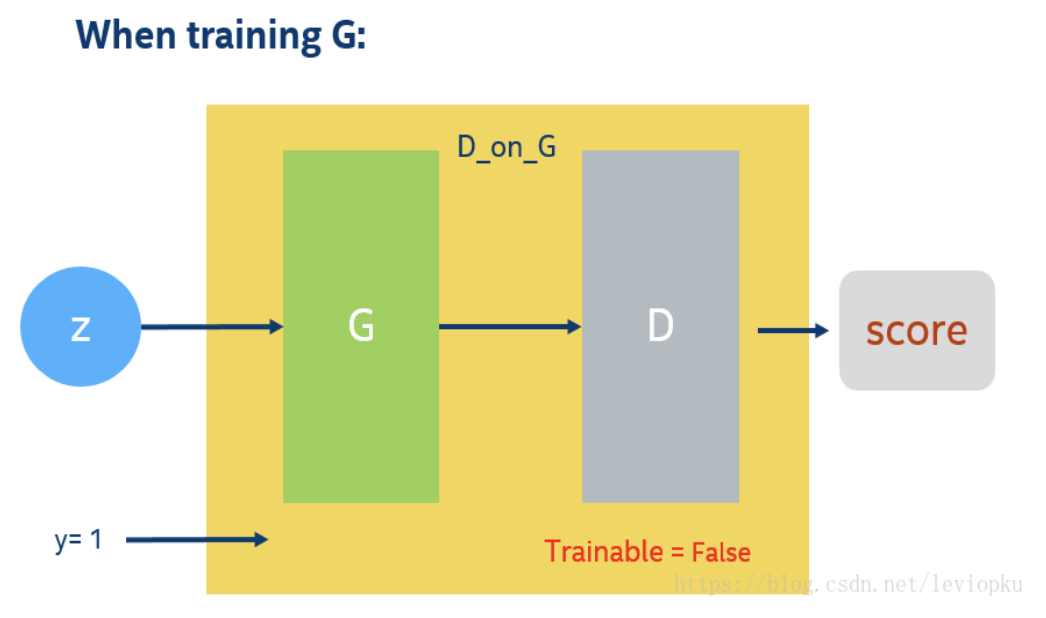
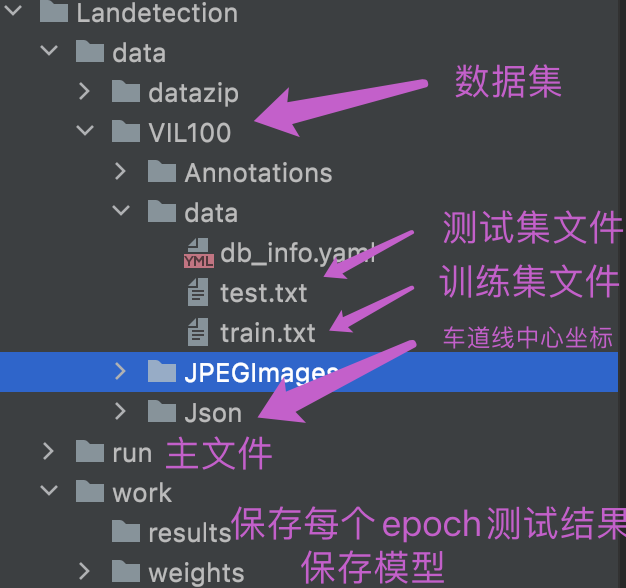


图2-2 GAN算法图

GAN的训练在同一轮梯度反传的过程中可以细分为2步，先训练D在训练G；注意不是等所有的D训练好以后，才开始训练G，因为D的训练也需要上一轮梯度反传中G的输出值作为输入。

图2-3 文件图示

在该算法中，使用了8000个训练集和2000个测试集，其中训练集有10个做验证集。使用了1000次的epoch迭代。

# 实施步骤

1.数据预处理

1. 整理数据集，划分训练集和测试集

图3-1 测试集和训练集划分结果

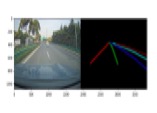
1. 从训练集中随机抽取一张图像进行可视化

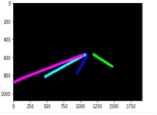
图3-2 随机抽取图像可视化

1. 再次从训练集中随机抽取一张图像进行可视化



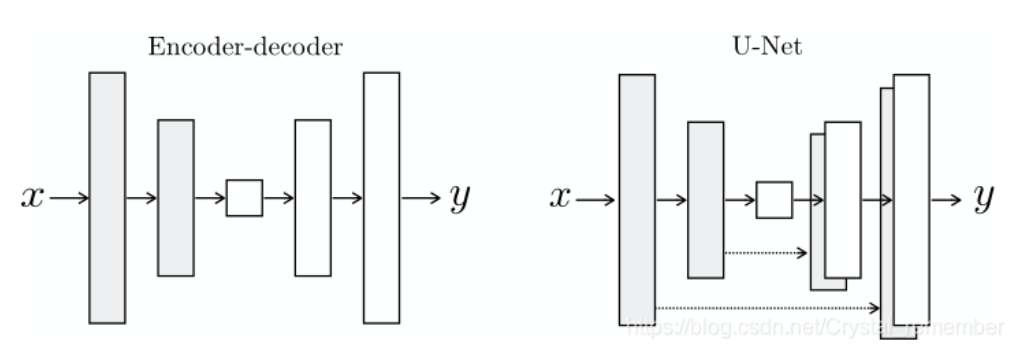
图3-3 再次随机抽取图像可视化

1. 可视化图像对应的标签

图3-4 可视化图像对应的标签

2.生成器模型搭建

（1）结构：

图3-5 生成器模型结构图

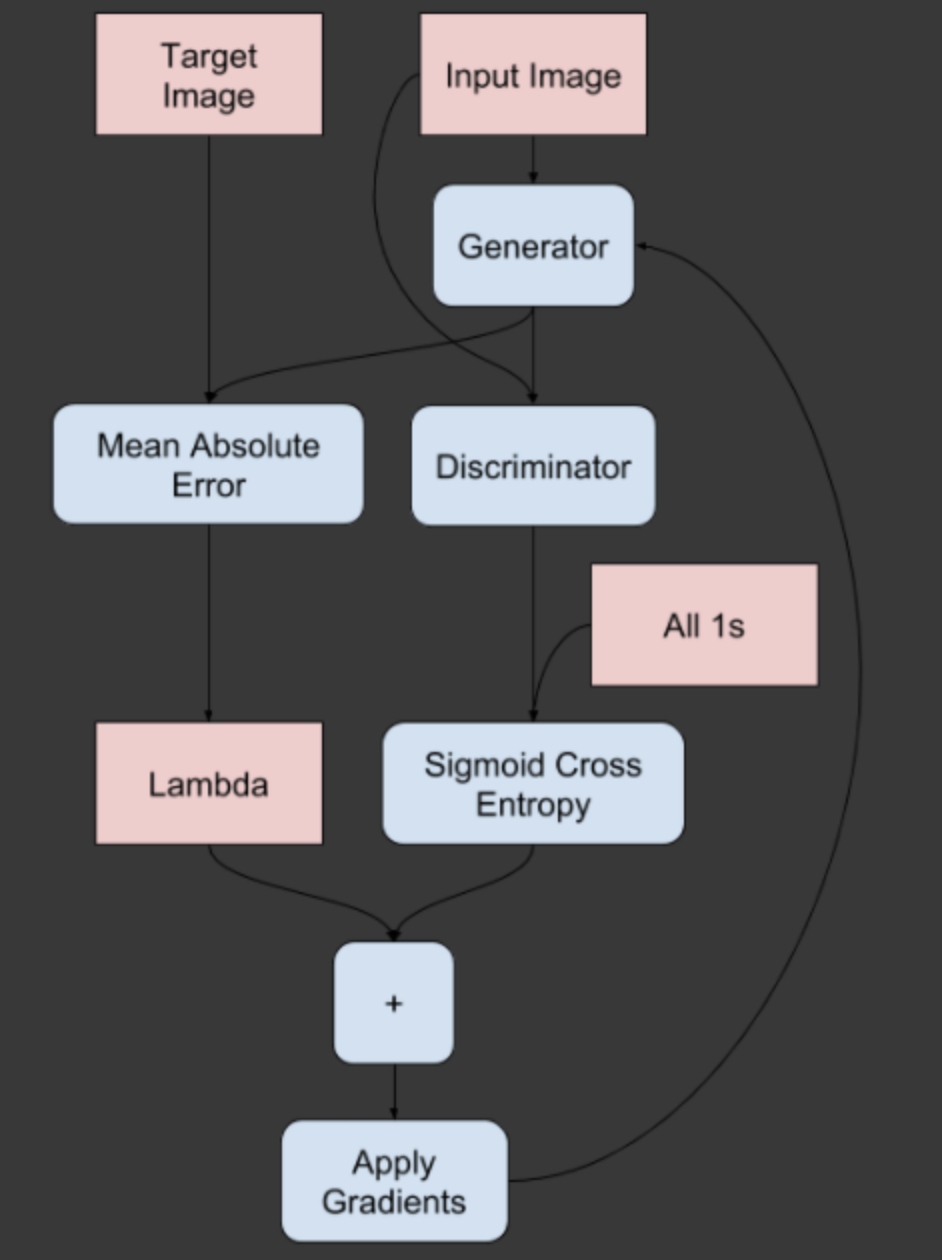
即一个修改过的U-Net网络，跳连连接编码器和解码器。

编码器（下采样）的每个块由卷积层，batchnorm层和leakyrelu层构成；

解码器（上采样）的每个块由反卷积层，batchnorm层和Dropout层（仅应用于前三个块）构成。

（2）生成器损失：

1. 生成的图像和元素为1的张量间的叉熵损失；
2. 生成的图像和真实图像的L1损失，这使得生成图像结构上和真实图像相似。

图3-6

生成器G的结构采用的是U-Net。在车道线检测任务中，需要从一个图像到另一个图像，输入输出虽然在表面细节不同，但是底层大致的结构是相同的，所以输入与输出需要粗略的对齐。对于图像的任务而言，需要输入输出之间除了共享高层语义信息之外，还需要能够共享底层的语义信息。这样，U-Net就可以被运用起来。特别的，为了更好的传递信息，Pix2Pix的作者加了跳连, 直接从i层传输到n-i层，n是总得网络层数，i是与n-i层之间有相应的通道。

3.判别器模型搭建

在Pix2Pix中，判别器输入两张图像，一张是输入给生成器G的图像，另一张是生成器G的输出的图像。也就是说，对于判别器而言，只有高质量的输出图像是不能骗过判别器的，有两张图像的对应关系能使判别器更好地工作。

判别器使用的所谓的PatchGAN结构（可看做卷积）。生成的矩阵中的每个点代表原图的一小块区域（patch）。通过矩阵中的各个值来判断原图中对应每个Patch的真假。原来用一个值衡量图像，现在用生成的矩阵来衡量。

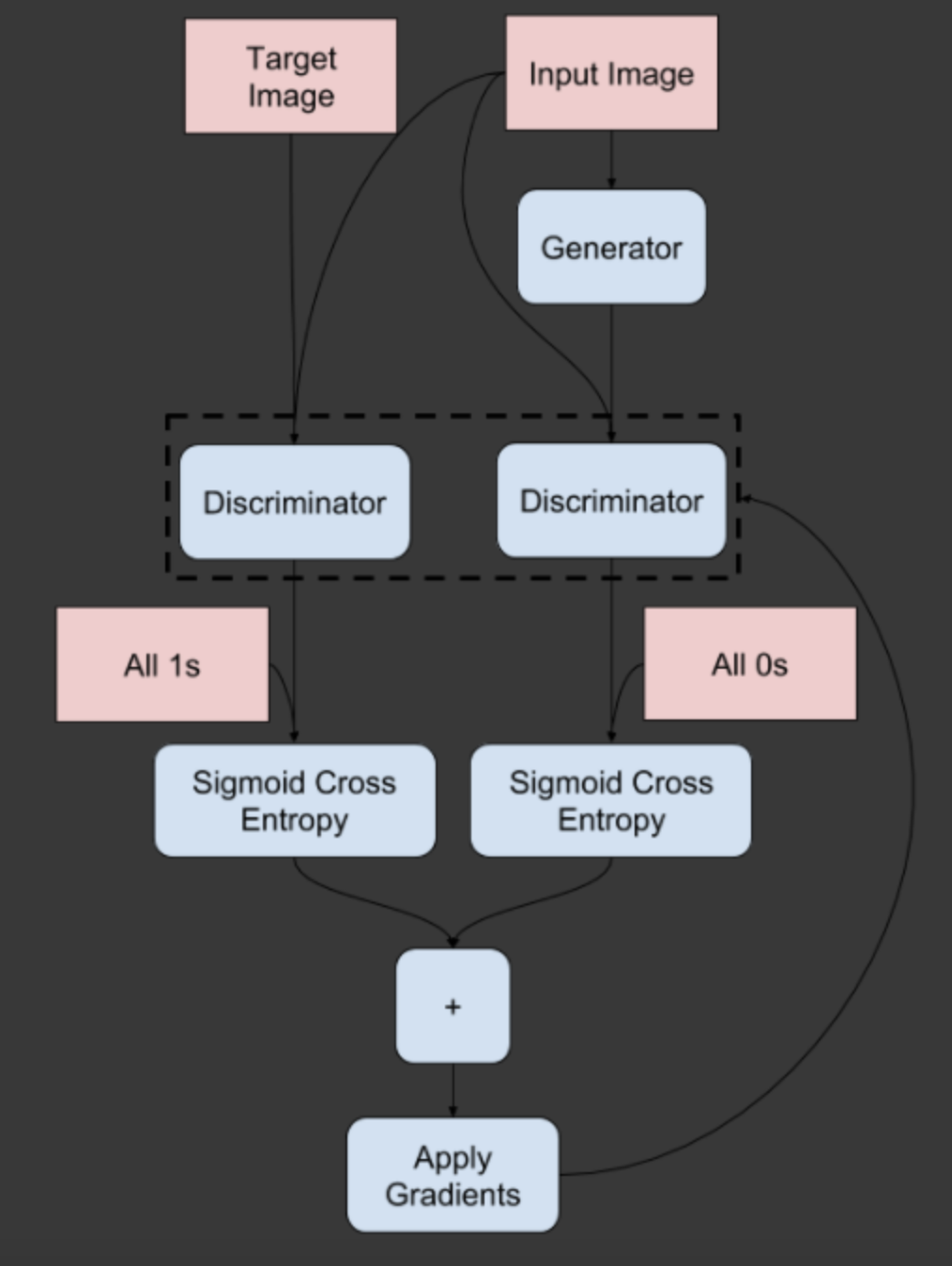


图 3-7

4.生成器、判别器模型测试

完成生成器和判别器的组网后，测试一下两者的输出是否符合我们的预期

测试结果：

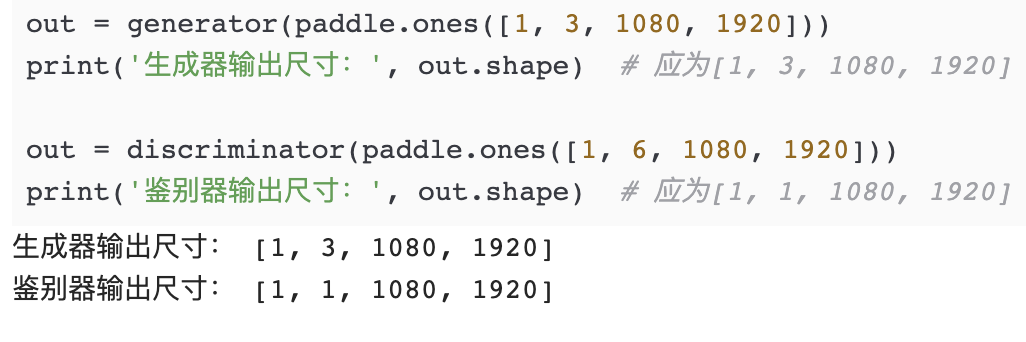


图3-8 测试结果

5.模型训练

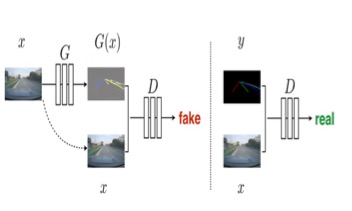


图3-9 pix2pix基于GAN实现图像翻译

更准地讲是基于cGAN（conditional GAN，也叫条件GAN），因为cGAN可以通过添加条件信息来指导图像生成，因此在图像翻译中就可以将输入图像作为条件，学习从输入图像到输出图像之间的映射，从而得到指定的输出图像。

Pix2Pix，通过随机向量z和图像x生成需要图像y，即{z,x} -> y生成器G用于生成尽可能愚弄判别器D的图像，判别器D尽可能分辨出生成器G生成的假图以及真实图像。

在Pix2Pix的优化目标包含2个部分。一部分是cGAN的优化目标。cGAN的优化目标如下所示：

LcGAN​(G,D)=​Ex,y​[logD(x,y)]+Ex,z​[log(1−D(x,G(x,z))],​

z表示随机噪声，判别器D的优化目标是使得LcGAN(G,D)的值越大越好，而生成器G的优化目标是使得log(1−D(x,G(x,z))) 越小越好。

这里需要注意的是正如GAN论文提到的，这样的计算方法在训练时容易出现饱和现象，也就是判别器D很强大，但是生成器G很弱小，导致生成器G基本上训练不起来，因此可以将生成器G的优化目标从最小化log(1−D(x,G(x,z)))修改为最大化log(D(x,G(x,z))) ，Pix2Pix算法采用修改后的优化目标。

另一部分是L1距离，用来约束生成图像和真实图像之间的差异，这部分借鉴了其他基于GAN做图像翻译的思想，只不过这里用L1而不是L2，目的是减少生成图像的模糊。

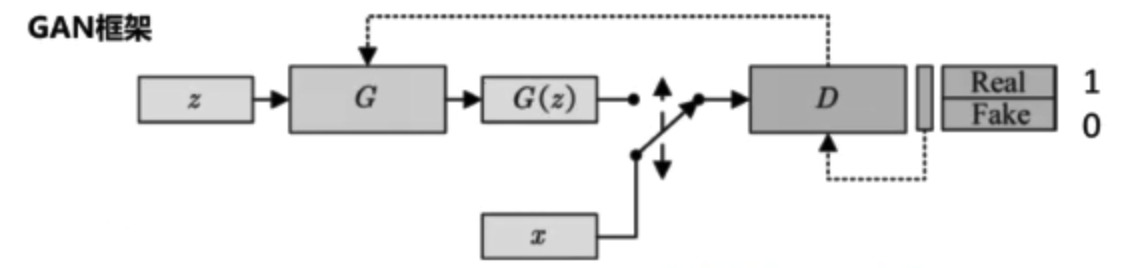


图3-10 GAN框架

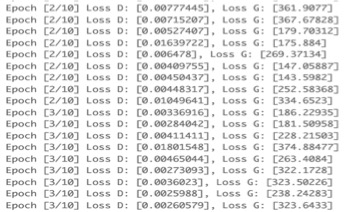


图3-11 模型训练展示

# 运行结果

为生成器加载权重，再次读取数据后，输出运行结果

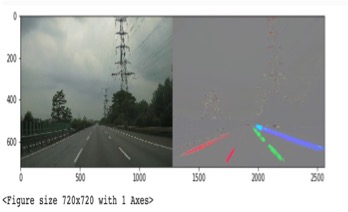


图4-1 部分输出结果展示

# 五、发现的问题、排除方法和改进措施总结

1. 由于选题的项目是基于PaddlePaddle深度学习框架的，对于Paddle及编程环境搭建并不是特别熟悉，可能看着页面上凭空给出的代码及操作陷入困境。通过查阅官网的详细教程，查看其API和函数文档等完成项目搭建。
2. 在环境搭建过程中，发现python解释器中的包和版本容易混乱冲突。通过搭建搭建虚拟环境，目前配置是Anaconda+PyCharm的组合，以此来保持全局解释器的干净整洁。
3. 在代码编写过程中，遇到很多包多次尝试仍无法导入，通过网络查找各种解决方法等，顺利的完成包的导入。
4. 选取的 VIL-100数据集，体量庞大，经过简短的训练，发现如果完全训练完成整个数据集，用时太长，我们通过整理相应数据，等比例减少数据，以此保证在适当时间内完成项目运行。
5. 在学习过程中发现，资料显示，GAN存在训练不稳定、梯度消失、模式崩溃的问题会导致训练结果比较差，导致生成结果的一些信息缺失，特征不全。

本项目基于经典的生成式对抗神经网络Pix2Pix实现了车道线检测任务，一种基于生成式对抗网络的无人驾驶车道线检测方法，其特征在于：包括车道线图片的生成网络、判别网络和检测算法三个部分。与以往基于分割的车道线检测不同，基于GAN方法的车道线检测不仅能识别出表观上的车道线，还能识别出语义上的车道线。通过这个项目我们可以比较清晰的学会了三个概念：GAN是什么、如何实现一个简单的GAN、GAN可以做什么。从神经网络到深度学习，去感受深度学习的变化和突破，去了解人工智能的发展。