

GAN 在自动驾驶数据集生成方面的应用

田凯茜

(西北工业大学 自动化学院, 陕西 西安 710072)

摘要: 随着传感器技术、人工智能技术等技术的突破,无人驾驶技术得到了空前的发展。自动驾驶汽车期望通过多传感器的目标检测达到全方位的环境感知。其中视觉传感器作为最贴近真实人类感知的传感器,更加成为了自动驾驶汽车需要克服的技术重难点。目前,由于安全问题,处于试验阶段的自动驾驶汽车无法在真实道路上采集大量复杂路况照片,而最新的神经网络算法需要大量的数据集才能得到可靠性较高的参数。除此以外,自动驾驶汽车在视觉方面还需要大量特殊复杂视觉死区的图像数据,这类数据难以获得且数量极少。提出一种基于生成对抗网络来生成大量真实道路图片,丰富自动驾驶汽车图像数据库,并且增加道路图像的情景多样性和全面性。采用 KITTI 数据集进行实验验证,生成对抗网络能够很好地生成与真实道路相似的图像。

关键词: 自动驾驶; 生成对抗网络; DCGAN; 图像处理

doi: 10.3969/j.issn.1006-8554.2019.11.010

0 引言

自动驾驶汽车产业近些年飞速的发展,各大互联网公司都在自动驾驶方面开展了研发和产业羡慕,如百度“Apollo”,Tesla“Autopilot”,奥迪 A8,Google“Waymo”等。现市面上大部分的自动驾驶汽车都是 L2 级别的,即在大部分实现自动化,但仍需要人为注意力集中的监管。随着人工智能方面的发展,通过大规模神经网络建模,能够很好地推动自动驾驶从 L2 ~ L3 级别的跨越。GAN(生成对抗网络)作为最前沿的研究热点,提供给我们丰富数据集的新方法。生成对抗网络由 Goodfellow 提出,是一个博弈模型包含生成器和判别器。生成器用于提取数据分布,生成符合样本分布的仿真数据。判别器是一个分类器,用于给仿真数据评分,越像真实数据评分越低,否则越高。DCGAN 将传统生成对抗网络的全连接层用卷积层替代,去掉了池化层并将生成器使用了判别器的学习结果。DCGAN 作为 GAN 的衍生,提供了强大的图片生成能力。基于现有小样本数据来生成高质量的仿真数据,为特征提取和训练提供海量数据。

1 DCGAN 算法原理简述

1.1 DCGAN 的数学原理

GAN 的原理为判别器 $D(x)$ 从判别的角度尽可能地区分真实样本和仿真样本,因此判别器 $D(x)$ 的目标函数是最大化 $D(x)$ 的值。

而生成器 $G(z)$ 从生成样本数据的角度尽可能地生成和样本数据同分布的数据,因此最小化 $D(G(z))$ 的值被确定为生成器的目标函数。最终得到的最佳判别器如下式:

$$D_{G(z)} = \frac{p_{data(x)}}{p_{data(x)} + p_{G(z)}}$$

通过两个模型的不断博弈,最终达到平衡状态。GAN 网络对应的 loss 函数如下式:

$$\min_G \max_D V(D, G) = E_{x \sim p_{data(x)}} [\log D(x)] + E_{z \sim p_{z(0)}} [\log(1 - D(G(z)))]$$

式中 $p_{z(0)}$ 是生成器输入的先验噪声信号, $p_{data(x)}$ 是样本数据的

分布。

DCGAN 则是在 GAN 的基础上,将判别模型和生成模型的网络优化为全卷积网络。数学原理和上述原理一致。经典的 CNN 网络卷积核如下:

$$o = \left\lfloor \frac{i + 2p - k}{s} \right\rfloor + 1$$

式中 i 为输入大小,步长为 s ,padding 为 p ,输出大小为 o 。DCGAN 中的分布卷积则为

$$o' = s(i' - 1) + k - 2p$$

1.2 DCGAN 的结构实现

依据算法目的为生成仿真的自动驾驶图像传感器获取的图像数据,按照上述数学原理设计 DCGAN 的实现结构。

首先设计生成器,生成器的架构由一层全连接层和四层反卷积网络构成。全连接层由输入层,变换层和激活层组成。输入层的神经元个数为 $64 \times 8 \times \text{img_width} \times \text{img_length}$,激活函数在本系统中采用 ReLu 激活函数

$$y = \max(0, x)$$

全连接层后接的四层,该部分由反卷积实现,称作反卷积层。卷积核的核大小均为 5×5 ,步长为 2,使用高斯分布。第二层的反卷积层输出通道为 64×4 ,第三层为 64×2 ,第四层为 64×1 ,激活函数同样采用 ReLu 激活函数。最后一层的输出通道数为 3,即对应彩色图像中的红、绿、蓝三色通道。

设计判别器,判别器由四层卷积层和一层全连接层构成。前四层是有卷积形式出现,其中的卷积层的卷积核大小为 5×5 ,同样使用 Gauss 分布。前四层卷积层的激活函数均使用 LeakyReLu 函数

$$y_i = \begin{cases} x_i & \text{if } x_i \geq 0 \\ \frac{x_i}{a_i} & \text{if } x_i < 0 \end{cases}$$

其中 a_i 是 $(1, +\infty)$ 区间内的固定参数。对于 ReLu 函数来说,Leaky ReLu 激活函数是优化形式,原理不是将小于零的值均置为零,而是按照比例赋值。最后一层的全卷积层则属输

出一个神经元。其卷积核大小也为 5×5 。

2 基于自动驾驶数据集实验验证

2.1 实验数据集的选取

自动驾驶的数据集相较于普通的图像识别数据集少很多。但是随着研究需求的不断增长,目前,也存在一些很不错的数据集如 RobotCar (牛津大学对多种道路情境的数据采集)、Cityscape(城市道路街景)、BDDV(自动驾驶视频数据集)、TORCS、CARLA 等。我们最终选择了目前国际认可度较高,规模最大的数据集 KITTI(见图1)。

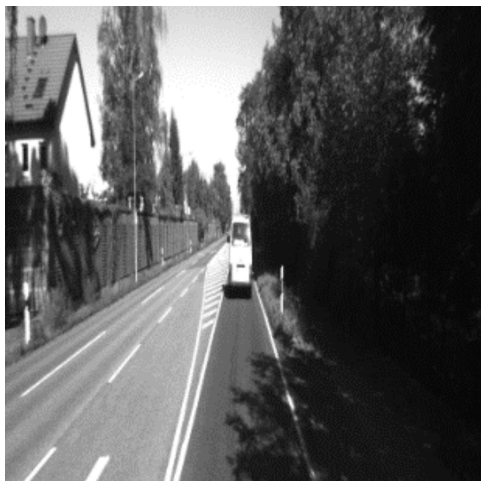


图1 KITTI 数据集

其中包含了多种不同情境下汽车视角的路况照片。是一个很好的训练样本。

2.2 实验平台的搭建与训练

本次实验基于 Windows 10 系统搭建的 tensorflow-gpu 的运行平台。

首先是环境的搭建,在 Windows 10 平台上依次安装 Anaconda3、Pycharm、Python 和 pip。在 Windows 环境安装 Python 3.5 版本,安装 CUDA 驱动程序,安装 TensorFlow 的 Nightly Build 版,最后安装 CuDNN 库。

使用 KITTI 数据集中道路图像作为训练集,使用 DCGAN 在 tensorflow 下训练。Epoch 230 时得到如下训练结果(见图2)。



图2 DCGAN 生成图像

3 结果分析

3.1 实验结果分析

我们可以看出在训练次数达到足够大的时候,DCGAN 能够很好地生成逼真现实道路的图片。并且,我们可以使用上述生成的图片来扩充自动驾驶情境下的数据集,同时也能生成一些常规道路难以采集的图片,对于自动驾驶视觉图像分析有着促进作用。

DCGAN 的原理简单,运用灵活。能够通过多次训练,学习数据的分布并生成相似度极高的仿真图片。在传统的生成模型中是一个很大的突破,在自动驾驶领域数据集拓展方面有着很强大的应用效果,并且在其他领域也有很好的应用前景。但是 DCGAN 存在训练稳定性不够高的问题,比如生成器和判别器的训练平衡难以掌控,训练时长较长,耗费的计算资源较多等神经网络共同的缺点,还有很大的优化空间。

3.2 自动驾驶 DCGAN 生成图像模糊及解决方法

由上述实验结果图,我们会发现图像存在模糊的问题。是由于在生成器中,采用了四层反卷积网络。多层反卷积网络通过不断地迭代会产生不均匀重叠的问题,反映在最终生成的图像上就是模糊的问题。

为了避免上述现象的产生,有两种解决办法。一种是改变内核大小,通过步幅来划分内核大小。另外一种方法是通过插值法调整图像大小和分辨率再执行卷积操作。

参考文献:

- [1] Bolin Zhou, Jihu Zheng, Pei Yin et al. Research on Autonomous Driving Simulator Control and Decision Algorithms based on Computer Vision Methods [P]. Proceedings of the 3rd International Conference on Computer Engineering, Information Science & Application Technology (ICCIA 2019) 2019.
- [2] 曹仰杰,贾丽丽,陈永霞,等.生成式对抗网络及其计算机视觉应用研究综述[J].中国图象图形学报,2018,23(10):1433-1449.
- [3] 樊雷.一种基于 TensorFlow 的 DCGAN 模型实现[J].电脑知识与技术,2017,13(36):219-221.
- [4] 笪陈宇,唐明,雷鑑铭.基于自动驾驶的机器视觉原理及应用[J].无线互联科技,2018,15(10):124-125.