

生成对抗网络模型综述

张营营^{1,2,3}

(1.中国科学院上海微系统与信息技术研究所 上海 200050; 2.上海科技大学 上海 201210;
3.中国科学院大学 北京 100049)

摘要: 针对生成对抗网络(GAN)这一热点模型,介绍其发展和应用的趋势。本文主要对比了现有几种典型的生成对抗网络模型及其变体:生成对抗网络(GAN)、条件生成对抗网络(CGAN)、深度卷积生成对抗网络(DCGAN)、半监督生成对抗网络(SGAN)信息生成对抗网络(InfoGAN)。同时本文系统地总结了生成对抗网络各种变体在计算机视觉领域的主要应用及性能优劣。文章最后分析了生成对抗网络存在的问题,以及对生成对抗网络研究趋势做了总结和展望。

关键词: 生成对抗网络; 无监督学习; 卷积神经网络; 图像生成

中图分类号: TP391.4

文献标识码: A

文章编号: 1674-6236(2018)05-0034-04

A survey on generative adversarial networks

ZHANG Ying-ying^{1,2,3}

(1.Shanghai Institute of Microsystem and Information Technology, Chinese Academy of Sciences, Shanghai 200050, China; 2.School of Information Science and Technology, ShanghaiTech University, Shanghai 201210, China; 3.University of Chinese Academy of Sciences, Beijing 100049, China)

Abstract: Towards generative adversarial networks(GAN) model, this paper firstly analyzed the trend of its development and application, then introduced and compared existing representative GAN models and its variants: GAN, Conditional GAN, Deep Convolutional GAN, Semi-Supervised GAN, Info GAN and etc. Also, this paper summarized the application of GAN and its variants in computer vision community. At last this paper analyzed the existing problems of GAN and made the summary and outlook on future work of GAN models.

Key words: GAN; unsupervised learning; CNN; image generation

近年来,深度学习在很多领域的都取得了突破性进展,但大家似乎发现了这样的一个现实,即深度学习取得突破性进展的工作基本都是判别模型相关的。2014年 Goodfellow等人启发自博弈论中的二人零和博弈,开创性地提出了生成对抗网络^[5](GAN)。生成对抗网络包含一个生成模型和一个判别模型。其中,生成模型负责捕捉样本数据的分布,而判别模型一般情况下是一个二分类器,判别输入是真实数据还是生成的样本。这个模型的优化过程是一个“二元极小极大博弈”^[1]问题,训练时固定其中一方(判别网络或生成网络),更新另一个模型的参数,交替迭代,最终,生成模型能够估测出样本数据的分布。生成对抗网络的出现对无监督学习^[15],图

片生成^[16]的研究起到极大的促进作用。生成对抗网络已经从最初的图片生成,被拓展到计算机视觉的各个领域,如图像分割^[2]、视频预测^[3]、风格迁移^[4]等。

1 生成对抗网络

生成对抗网络(GAN)包含一个生成模型^[6] G 和一个判别模型^[7] D ,其结构如图1所示。生成对抗网络的目的是学习到训练数据的分布 p_g ,为了学习该分布,首先定义一个输入噪声变量 $p_z(z)$,接下来将其映射到数据空间 $G(z; \theta_g)$,这里的 G 就是一个以 θ_g 作为参数的多层感知网络构成的生成模型。此外,定义一个判别模型 $D(x; \theta_d)$ 用来判断输入的数据是来自生成模型还是训练数据, D 的输出为 x 是训练数据的概

收稿日期:2017-03-08 稿件编号:201703095

作者简介:张营营(1992—),男,甘肃陇南人,硕士研究生。研究方向:计算机视觉。

率。最后训练 D 使其尽可能准确地判断数据来源, 训练 G 使其生成的数据尽可能符合训练数据的分布。

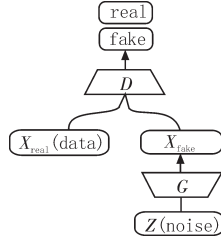


图1 生成对抗网络模型结构图

算法1 生成对抗网络的批随机梯度下降方法。

for number of training iterations **do**

for k steps **do**

 • Sample minibatch of m noise samples $\{z^{(1)}, \dots, z^{(m)}\}$ from noise prior $p_g(z)$.

 • Sample minibatch of m examples $\{x^{(1)}, \dots, x^{(m)}\}$

from data generating distribution $p_{data}(x)$.

 • Update the discriminator by ascending its stochastic gradient:

$$\nabla_{\theta_d} \frac{1}{m} \sum_{i=1}^m \left[\log D(x^{(i)}) + \log(1 - D(G(z^{(i)}))) \right]$$

end for

 • Sample minibatch of m noise samples $\{z^{(1)}, \dots, z^{(m)}\}$ from noise prior $p_g(z)$.

 • Update the generator by descending its stochastic gradient:

$$\nabla_{\theta_g} \frac{1}{m} \sum_{i=1}^m \log(1 - D(G(z^{(i)})))$$

end for

换句话说, D 和 G 的优化过程是一个二元极大问题, 其目标函数可以表示为:

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

值得说明的是, D 和 G 的优化必须交替进行, 因为在有限的训练数据情况下, 如果先将 D 优化完成会导致过度拟合, 从而模型不能收敛。训练生成对抗网络的算法伪代码参考算法1。

2 生成对抗网络的各种变体

2.1 条件生成对抗网络

条件生成对抗网络 (CGAN) 是生成对抗网络的一个扩展, 其结构图如图2所示, 它的生成模型和判

别模型都基于一定的条件信息 y 。这里的 y 可以是任何的额外信息, 例如, 类别标签或者数据属性等。这样生成模型 G 就有两个输入: $p_z(z)$ 和 y 。一般情况下, $p_z(z)$ 和 y 以隐藏节点连接的方式结合在一起。因此, 该二元极大问题的目标函数就变为:

$$\min_G \max_D V(D, G) = E_{x \sim p_{data}} [\log D(x|y)] + E_{z \sim p_g} [\log(1 - D(G(z|y)))] \quad (2)$$

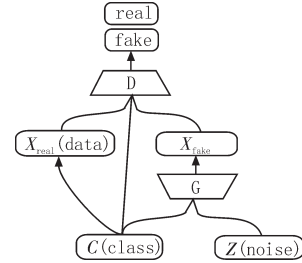


图2 条件生成对抗网络模型结构图

条件生成对抗网络可以根据输入条件不同生成相应类型的图片。

2.2 深度卷积生成对抗网络

深度卷积生成网络 (DCGAN) 将卷积神经网络 (CNN)^[12] 引入到了生成模型和判别模型当中, 使得生成性能有了质的提升, 以至于后来很多工作都在该网络的基础上进行改进。该网络结构的几个设计要点为: 1) 将卷积网络中的池化层^[14] 用相应步长的卷积层代替; 2) 在生成模型和判别模型中都使用了批归一化层^[13]; 3) 去掉了网络中的全连接层^[11]; 4) 在生成模型中采用 ReLU^[7] 激活函数; 5) 在判别模型中采用 LeakyReLU^[9] 激活函数。其生成模型结构图如图3所示。深度卷积生成网络相对于一般的生成对抗网络而言具有更强大的生成能力, 同时训练起来更加稳定、容易, 生成样本更加多样化等优点。深度卷积生成网络生成的图片足以以假乱真, 但缺点是生成图像分辨率比较低 (64 * 64), 这也是目前各种生成对抗网络及其变体所具有的共同问题, 因此, 如何生成高质量、高分辨率图像将会是一个热门研究方向。

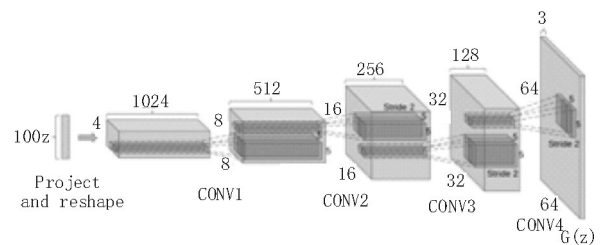


图3 深度卷积生成网络中生成模型结构图

2.3 半监督生成对抗网络

半监督生成对抗网络(SGAN)的判别模型不仅判断图像的来源(由 G 生成或者来自训练数据)同时判断图像类别,这样使得判别模型具有更强的判别能力,其网络结构图如图4所示。另一方面,输入网络的类别信息也在一定程度上提高了生成模型生成图片的质量,因此半监督抗网络的性能要比普通的生成对抗网络的性能略好一些。

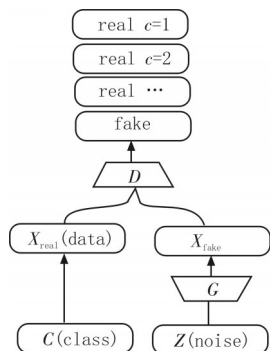


图4 半监督生成对抗网络生成结果生成

2.4 信息生成对抗网络

信息生成对抗网络^[10](InfoGAN),其结构图如图5所示,它在一般的生成对抗网络的基础上增加了一个潜在编码 c ,其中 c 可以包含多种变量,比如在MNIST中, c 可以一个值来表示类别,一个高斯分布的值来表示手写体的粗细。网络中生成模型的输入 $p_z(z)$ 和 y ,输出为 $G(z, c)$ 。为了避免网络没有监督使用 c ,信息生成对抗网络在优化的目标函数中增加了一项: $I(c; G(z, c))$ 用来表示共同信息的程度。因此该网络的目标函数为:

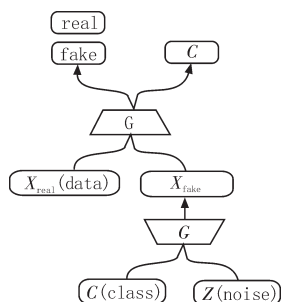


图5 信息生成对抗网络模型结构图

$$\min_G \max_D V(D, G) = E_{x \sim p_{data}(x)} [\log D(x)] + E_{z \sim p_z(z)} [\log (1 - D(G(z, c)))] - \lambda I(c; G(z, c)) \quad (3)$$

信息生成对抗网络可以通过改变潜在编码 c 的

值来控制生成图片的属性,例如生成不同倾斜度或者粗细的数字。

3 生成对抗网络及其变体性能的比较

目前对于生成图像的质量好坏的评价没有一个量化的标准,主要依靠个人感官评价。最早Goodfellow等人提出的生成对抗网络其实生成图像的能力非常有限,生成模型和判别模型都采用多层感知器,而且是首次提出该理论,实验调参经验缺乏一定的积累。但对后续发展的贡献是巨大的,因为这是一个理论从无到有的过程,无论后续的工作取得什么样的成就,都是最初生成对抗网络理论的基础上做的改进。半监督生成对抗网络和条件生成对抗网络是紧接着出现的改进工作,它们的性能比原始的生成对抗网络略好。深度卷积生成对抗网络的出现是一个里程碑意义的工作,它使得生成对抗网络的性能有了质变,它首次将深度卷积网络引入生成对抗网络,而且为稳定训练网络提供一套比较完整的参数配置。在后续工作中绝大多数模型参数配置都采用它的配置。信息生成对抗网络的主要贡献是引入了一个能够控制生成图像属性的潜在编码,从而让网络生成人为规定特殊属性的图像。

4 生成对抗网络在计算机视觉领域的几个应用

4.1 图像分割

图像分割是指将图像中的内容根据不同的语义分割开来。绝大多数分割的方法是将图像的每个像素进行分类,这样就忽略了空间一致性这个重要信息。生成对抗网络做图像分割时,生成模型被图像像素分类的网络替换,判别模型用于判断是网络的分类产生的分割图(Class predictions)还是标准的分割图(Ground truth)。其流程图如图6所示。

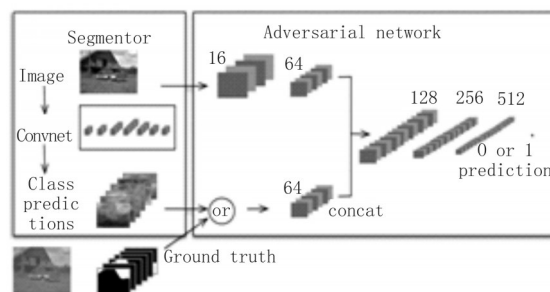


图6 生成对抗网络应用于图像分割的流程图

4.2 视频预测

视频预测是根据当前的几帧视频预测接下来的一帧或者多帧视频。一般常用的做法是用最小二乘回归视频帧的逐个像素值,这样做的一个问题就是生成视频帧存在动作模糊(Motion blur)。一种采用生成对抗网络的做法是,将现有视频帧改变不同尺寸输入生成模型 G , 让其输出接下来的真值帧(Ground truth frame)。图7给出了生成对抗网络应用于视频预测时生成网络部分示意图。判别模型和一般生成对抗网络相同,即判断视频帧是生成的还是训练数据原有的。

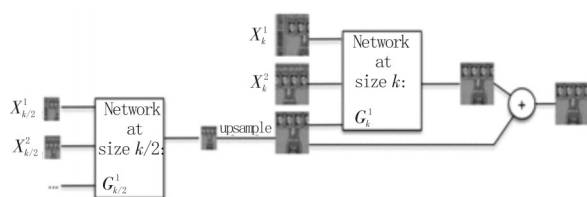


图7 视频预测生成对抗网络生成模型示意图

4.3 风格迁移

风格迁移是指将目标图像的风格迁移到源图像当中,使得源图像保留内容的同时具有目标图像的风格,比如将卡通人物的风格迁移到真实人脸图像使其成为卡通头像。生成对抗网络用于图像风格迁移时其特点有3个:1)将源图像输入生成网络后,生成网络输出目标图像风格的图片;2)生成图像和源图像特征相同(或者说保持内容);3)当把目标图像输入生成网络后,生成网络仍然输出目标图像(相当于单位映射)。风格迁移生成对抗网络的结构如图8所示。

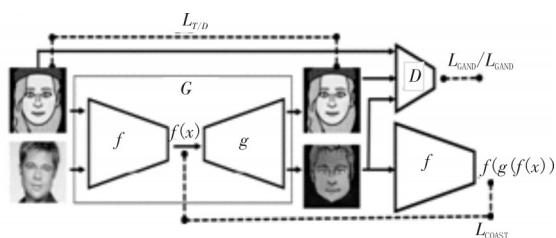


图8 风格迁移生成对抗网络结构图

5 结论

生成对抗网络在计算机视觉的各个领域具有广泛的应用前景,虽然经过几年的发展其性能有了很大程度的提升,但仍然存在以下几个问题:1)缺乏一个量化的评价结果好坏的标准;2)生成图片的分辨

率有限,很难生成高清图像;3)网络训练对参数配置敏感,不稳定。因此,如何构建一个训练容易,能够生成高质量高分辨率图像的网络是接下的研究热点之一。鉴于目前学术界对于生成对抗网络综述的文章非常少,本文分析比较了生成对抗网络及其变体,对其做了一个总结和概述,同时分析了目前存在的一些问题,展望了该领域未来的研究趋势。本文的目的是方便研究者对生成对抗网络有一个总体的认识和把握,希望对大家在生成对抗网络相关领域的科研有所帮助或启示。

参考文献:

- [1] 王海军. 集值型两个函数的极小极大定理[J]. 数学的实践与认识, 2015(8):293-298.
- [2] 宋熙煜,周利莉,李中国,等. 图像分割中的超像素方法研究综述[J]. 中国图象图形学报, 2015, 20(5):599-608.
- [3] 齐美彬,陈秀丽,杨艳芳,等. 高效率视频编码帧内预测编码单元划分快速算法[J]. 电子与信息学报, 2014(7):1699-1705.
- [4] 蔡兴泉,郭天航. 面向手机应用的图像色彩风格迁移系统设计与实现[J]. 信息通信, 2016(6): 139-140.
- [5] Ian J. G, Jean P, Shah, Mehdi M, Bing X, Yoshua B. Generative adversarial nets. [C]// NIPS. 2014: 2672-2680.
- [6] 王军,夏利民,夏胜平. 利用生成模型的人体行为识别[J]. 国防科技大学学报, 2016,38(2):68-74.
- [7] 黄利文. 一种判别模型的择优方法[J]. 江西理工大学学报, 2013,34(1):96-99.
- [8] Zeiler M D, Ranzato M, Monga R. On rectified linear units for speech processing[C]// ICASSP 2013:3517-3521.
- [9] Maas A L, Hannun A Y, Ng A Y. Rectifier nonlinearities improve neural network acoustic models.[C]//ICML. 2013: 1152-1160.
- [10]Chen X, Duan Y, Houthoof R. Infogan: Interpretable representation learning by information maximizing generative adversarial nets.[C]// NIPS. 2016: 2172-2180.
- [11]李玉鑑,张婷,胡海鹤. 基于多层感知器的深度核映射支持向量机[J]. 北京工业大学学报, 2016,

(下转第43页)

量的实验测试其性能具体测试内容如下:当无人机定点装置,处于目标点位置飞行较为稳定时,通过将遥控器拨到手动通道,人为的将无人机拉到远离目标点的位置,再将遥控器迅速拨回自动通道,观察无人机定点装置的运动,并利用Vicon室内目标定位系统记录下无人机XY通道的运动数据,如图7所示。

通过实验数据可以看到,在开始位置将无人机拉到距离目标为(210,-200)位置处,通过控制律的调节,在误差允许的范围内,可以视为将无人机拉回原点,即目标位置,经测试能较好的实现定点功能。

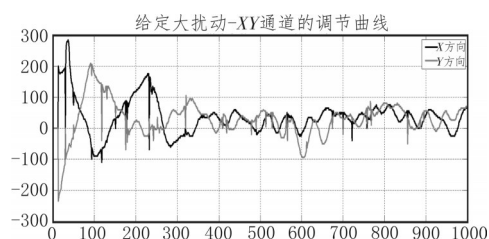


图7 XY通道跟踪误差曲线(纵轴单位:cm,横轴单位:ms)

4 结束语

本文通过在simulink平台上搭建四旋翼无人机模型,和实际中搭建无人机硬件平台和地面站,分别从仿真与实验的角度证明了基于视觉的无人机在定点飞行时的优良性能。

参考文献:

- [1] 李雷,陈桂琴,田园. 无人机的发展历程与展望[J]. 科技论坛, 2011(32):6.
- [2] 吴瑶瑶,白云飞,曹科才. 扑翼飞行器硬件平台设计与实现[J]. 信息化研究, 2016,42(3):38-42.
- [3] 吴显亮,石宗英,钟宜生. 无人机视觉导航研究综述[J]. 系统仿真学报, 2010,22(1):62-65.
- [4] 郭佳琪,吴杰,蒋一琦. 高压线巡线飞行器的研究与设计[J]. 南京工程学院学报:自然科学版,

2012,10(1):39-42.

- [5] 徐东升. 惯性导航中加速度计和陀螺仪性能研究[J]. 佳木斯大学学报:自然科学版, 2012,30(3):334-337.
- [6] 刘畅,杨良军,刘峰. 一种基于惯性/视觉信息融合的无人机自主着陆导航算法[J]. 导航定位与授时, 2016,3(6):7-11.
- [7] 盛哲. 基于视觉导航的无人机位姿控制与自动返航技术[D]. 上海:上海交通大学, 2015.
- [8] Zhang X, Xian B. Autonomous flight control of a nano quadrotor helicopter in a GPS-Denied environment using on-Board vision[J]. IEEE Transactions on Industrial Electronics, 2015:1-11.
- [9] 王翔鹏. GPS失效情况下农用无人机视觉导航关键算法研究[D]. 青岛:青岛科技大学, 2014.
- [10] 刘国宏,郭文明. 改进的中值滤波去噪算法应用分析[J]. 计算机工程与应用, 2010,46(10):187-189.
- [11] 赵高长,张磊,武风波. 改进的中值滤波算法在图像去噪中的应用[J]. 应用光学, 2011,32(4):678-682.
- [12] 黄忠主. 面向监视视频的运动轨迹提取方法研究[D]. 长沙:国防科技大,2011.
- [13] 闫钧华,陈少华,艾淑芳. 基于Kalman预测器的改进的CAMShift目标跟踪[J]. 中国惯性技术学报, 2014,22(4):536-542.
- [14] 韩崇昭,朱洪艳,段战胜. 多源信息融合[M]. 北京:清华大学出版社,2006.
- [15] 李俊,李运堂. 四旋翼飞行器的动力学建模及PID控制[J]. 辽宁工程技术大学学报:自然科学版, 2012,31(1):114-117.
- [16] 付瑞玲,乐丽琴. 基于MATLAB/Simulink的PID参数整定[J]. 工业控制计算机, 2013,26(8):75-76.

(上接第37页)

- 42(11):1652-1661.
- [12] Krizhevsky A, Sutskever I, Hinton G E. ImageNet classification with deep convolutional neural networks [C]// NIPS. 2012: 1097-1105.
- [13] Ioffe S, Szegedy C. Batch Normalization: Accelerating Deep Network Training by Reducing Internal Covariate Shift[C]//ICML. 2015: 448-456.

- [14] 李丹,沈夏炯,张海香,等. 基于Lenet-5的卷积神经网络改进算法[J]. 计算机时代, 2016(8):4-6.
- [15] 王建新,钱宇华. 符号数据的无监督学习:一种空间变换方法[J]. 计算机科学, 2016,43(1):89-93.
- [16] 陈中贵,欧阳永昇,曹娟. 特征保持的马赛克图像生成方法[J]. 计算机辅助设计与图形学学报, 2014,26(4):520-527.