

生成对抗网络研究综述

刘鹤丹, 叶汉平, 徐梦真, 赵旭磊

(厦门大学嘉庚学院 信息科学与技术学院, 福建 漳州 363105)

摘要: 生成对抗网络 GAN (Generative Adversarial Networks) 于 2014 年被 Lan 等人提出, 该生成式模型受到二人零和博弈的影响, 通过生成器和判别器的相互比较而形成。凭借生成数据质量较好的优势, GAN 已被广泛应用于各科研领域, 如图像生成与修复、人脸识别等多个领域中, 均有较好的表现。本文介绍了 GAN 的基本概念、代表性模型、主要应用领域等基础知识, 将 GAN 与传统类型的算法进行了优劣比较, 最后阐述了对 GAN 未来前景的看法。

关键词: GAN; 神经对抗网络; 二人博弈; 人工智能; 深度学习; 生成式模型

中图分类号: TP391

文献标识码: A

文章编号: 2095-1302 (2022) 11-0093-05

0 引言

随着人工智能和深度学习的发展, GAN 的出现成为了深度学习领域学者们研究的热点。GAN 在多个领域均表现得极其出色, 如图像的生成^[1]、修复^[2]、识别^[3-5]等。机器学习算法可分为监督学习和无监督学习两类, 监督学习譬如 YOLO 系列^[6]需要依赖已知的带标记数据进行训练, 这类算法虽有较好的效果但消耗成本较高; 而无监督学习因其无须提前标记好数据, 受到越来越多学者们的青睐, 并且 GAN 具有无须监督和做数据标记的优势, 解决了数据集问题^[7]。GAN 不仅可生成高质量图像, 还可进行图像增强、图像迁移。

1 GAN 的理论基础

GAN (生成对抗网络)^[8]由两大部分组成: 生成器网络 (Generative Model)、判别器网络 (Discriminative Model)。生成器根据要求会不断生成与实际标签数据相近的数据, 并传递给判别器; 判别器会区分生成器生成数据的结果与实际标签之间的区别, 并且判别器会根据判断后产生的误差反向传递至生成器中进行更新, 生成器收到反馈后会生成更加接近的数据再传递给判别器, 迭代数次后, 直至生成器生成的数据使判别器无法区分数据的真实性才停止训练。

如图 1 所示, 生成器 G 接收到随机变量 z 的数据后, 会生成假样本数据 $G(z)$; 将其传入到判别器 D 中, 判别器会对

接收的真实样本数据 x 进行判别处理并输出生成结果。生成结果为判别器 D 输入的是否为真实分布的概率, 若为 1 则数据为真实数据, 为 0 则为假数据。同时, 判别器 D 会把输出结果返回给生成器 G 用于其训练。当 D 的输出概率值为 0.5 左右, 表明无法区别数据来源, 即模型已达到最优状况, 则停止训练。而该停止条件如果导致判别器 D 判别没调好, 会造成不收敛、模式崩溃、梯度弱化, 甚至消失等问题, 当梯度消失时, 训练的生成器 G 相当于没有训练时的状态。这也是 GAN 模型难以训练的原因^[9]。

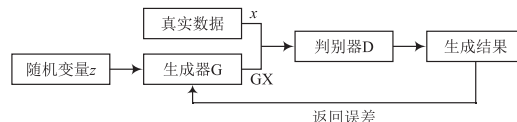


图1 GAN网络模型的基本结构示意图

2 GAN 变体模型

GAN 的训练较为困难, 因此在研究者们的不断探索中出现了很多不同的变体模型, 本章对具有代表性的变体模型进行简述与分析。

2.1 DCGAN

DCGAN (Deep Convolutional GAN)^[10], 由 Alec Radford 于 2015 年提出, 是最基本的 GAN 版本, 通过改变卷积神经网络的结构提高了样本质量和收敛速度, 许多 GAN 模型都基于 DCGAN 进行了改进。它可以很好地适应于卷积神经网络, 能够有效实现高质量图片的生成和相关模型生成。其中包含了一种突破性的关键技术: 批归一化。DCGAN 的生成效果如图 2 所示。虽然 DCGAN 的批归一化效果被证明是有效的, 但经过实验发现如果改变某些层的 BN 或者改变激活函数, 都可能导致网络生成的图像为噪声。

收稿日期: 2022-03-15 修回日期: 2022-04-13

基金项目: 国家级大学生创新创业训练计划项目 (202213469054);

福建省中青年教育科研项目 (JAT200922); 漳州市
自然科学基金 (ZZ2021J10); 厦门大学嘉庚学院预研项目
(YY2019L02)



图2 DCGAN 的生成效果

2.2 BigGAN

BigGAN (Large Scale GAN)^[11] 被称为史上最强的 GAN 图像生成器, 首次生成具有高保真度和低品种差距的图像。它与传统的 GAN 的区别之一在于训练中采用了很大的 Batch 值, 同时也增强了卷积的通道数和网络参数, 还包含了“截断技巧”和模型稳定性的控制等。随着相关研究的不断进步, 还衍生出了 BiGAN、BigBiGAN。BigGAN 的生成效果如图 3 所示。



图3 BigGAN 的生成效果

2.3 StyleGAN

StyleGAN (A Style-based Generator Architecture for GAN)^[12] 从 ProGAN 中演变而来, 具有可基于样式的生成器, 可生成更高质量的高分辨率图像。StyleGAN 将生成的过程可控化, 可通过数据集中的属性转换图像中的风格, 譬如可以实现无监督地修改人脸姿势、身份、发型等, 生成相对应的图像, 还可以生成汽车、卧室等高质量图像。

StyleGAN 参考了 ProGAN, 发现渐进层的视觉特征会受层和分辨率的影响, 越高的分辨率越可进行更细微和精确地控制。根据不同的分辨率范围, 在 StyleGAN 中分为粗糙、中等、高质三种类型。StyleGAN 在 GAN 模型基础上删除了传统输入, 添加了噪声 noise, 使用了自适应实例归一化 (AdaIN)。

StyleGAN 极大地提高了研究者们对 GAN 合成的理解和可控性。随着 StyleGAN 的发展, 其版本已经延续到 StyleGAN3。StyleGAN 的生成效果如图 4 所示。



图4 StyleGAN 的生成效果

2.4 StackGAN

StackGAN (Text to Photo-Realistic Image Synthesis With Stacked GAN)^[13] 是首个可根据文本描述来生成图像分辨率达到 256×256 的网络模型。StackGAN 的训练方式分为两个阶段。其中 Stage-I 会通过给定的文字生成低分辨率 (64×64) 图片; Stage-II 在 Stage-I 基础上生成高分辨率 (256×256) 图片并捕获被 Stage-I 忽视的文字信息, 修正 Stage-I 结果的缺陷、添加细节。这种分段式模型可能由于每个任务找不到重点而导致生成失败。StackGAN 的生成效果如图 5 所示。

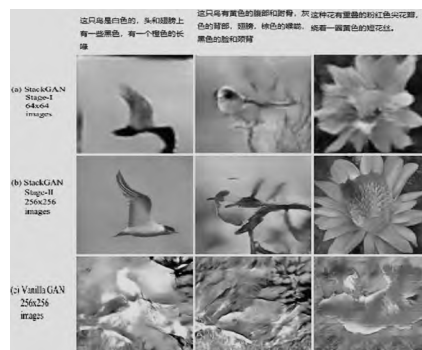


图5 StackGAN 的生成效果

2.5 CycleGAN

CycleGAN (Unpaired Image-To-Image Translation Using Cycle-Consistent GAN) 的核心思想是: 设有 X 域和 Y 域两个域, 从 X 中映射 Y , 再从 Y 中映射 X , 循环往复^[14], 如图 6 所示。

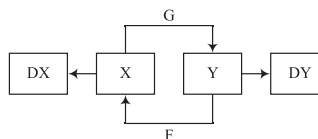


图6 CycleGAN 的核心原理

CycleGAN 主要应用于域迁移 (Domain Adaptation) 领域。域迁移是将数据从一个域移动到另一个域的过程, 譬如将照片中的马转换成斑马。在传统算法中, 一般对于两个域之间的某一物体相互转换需要该两个域之间具有相同的成对图片

作为数据集进行训练,如 pix2pix^[15],而 CycleGAN 并不需要成对的图片作为训练数据,只需要有充足的图片数据集就

可实现物品之间的相互转换。CycleGAN 的生成效果如图 7 所示。

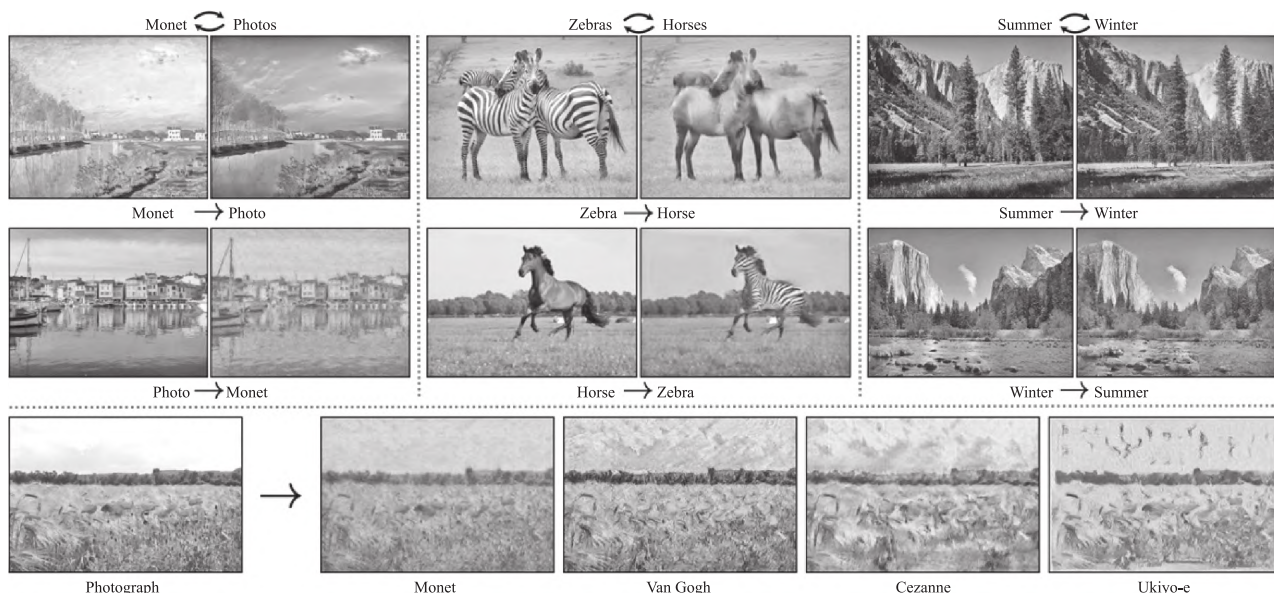


图 7 CycleGAN 的生成效果

3 GAN 系列的主要应用领域

3.1 图像生成

GAN 作为一个生成模型,能从大量的无标签数据中无监督地学习到一个具备生成各种形态(图像、语音、语言等)数据能力的函数(生成器),因此可以达到数据增广的目的。但 GAN 并不是单纯地对真实数据的复现,而是具备一定的数据内插和外插作用。例如:路面生成^[16]、机器翻译^[17-18]、二次元风格转换^[19]、人脸生成^[20]等。

3.2 图像超分辨率

图像超分辨率相关的研究中一个比较重要的课题是对天文图像和卫星图像做超分辨率,该课题的成果在各领域得到广泛应用。例如 SRGAN (Super-Resolution Generative Adversarial Network)^[21]可应用于图像超分辨率方面。它基于相似性感知方法^[22]提出了损失函数,可有效解决恢复后的图像丢失高频细节问题。

3.3 图像合成

图像合成是通过某种形式的图像描述创建新图像的过程。CycleGAN^[14]和 pix2pix^[15]均属于图像合成领域,具体应用例如:应用于场景合成的 pix2pix^[15],应用于人脸合成^[23]的 TpGAN^[24],应用于文本到图像合成的 StackGAN^[13],应用于风格迁移的 CycleGAN^[14]。

3.4 其他

GAN 除了被应用于以上三个领域外,在其他领域也有

着很好的表现,如 3D 打印^[25]、图形修复^[26]、图片编辑^[27]、人体相关的姿态估计^[28]、自动驾驶^[29]、目标跟踪检测^[30-32]等。

4 GAN 与传统算法对比

4.1 GAN 的优势

GAN 是一种生成式模型,相比于其他生成模型只用到了反向传播,而无需复杂的马尔科夫链,可产生更为清晰和真实的样本;其采用无监督训练,可广泛用在无监督、半监督学习领域。相比于其他机器学习模型,不受样本脆弱性的影响,可应用于图片风格迁移、超分辨率、图像补全、去噪等场景,避免了损失函数设计的困难;还可以通过 GAN 生成以假乱真的样本,缓解了小样本机器学习的困难。

4.2 GAN 的劣势

训练 GAN 需要达到纳什均衡,采用梯度下降法不一定都可以实现。目前并没有达到纳什均衡的更好方法,因此训练 GAN 并不稳定。

(1) 梯度消失:随着训练次数增加,G、D 的生成和判别能力逐渐增强且互相影响,当 D 训练到最好的情况下,原始 GAN 中 G 的损失函数存在自相矛盾的缺陷,从而导致 GAN 很难训练,微调敏感。

(2) 模式崩溃:模型只学习到真实样本分布的一部分,导致模型生成的样本单一,样本差异较小。

(3) GAN 不适合处理文本等离散形式的数据。

5 前景预测

如何根据简单的随机输入生成多样的、能够更好地与人类进行交互的数据,更好地配合深度学习,是 GAN 的近期发展方向。如何让 GAN 与特征学习、模仿学习、强化学习等技术更好地融合,利用 GAN 的特性或博弈论思想来推动人工智能的发展与应用,提高其对世界的理解甚至激发其创造能力是值得研究的问题。GAN 是生成式模型,采用模型学习方式来估测其分布并生成同分布的新样本,因此在图像视觉计算、语音语言处理、信息安全等领域有巨大的应用价值。但 GAN 解决了生成式模型问题的同时也引入了新的问题^[33],譬如有限的注释、多样性有限、受限数据、训练波动大等。以下四个方面是今后针对 GAN 进行研究的新方向:

(1) 资源问题的解决受数据集的限制,使得 GAN 没有普适性和安全性。

(2) 可结合其他机器学习算法改善 GAN 的目标函数和框架,对结构进行拓展。

(3) 实现统一的度量标准。由于 GAN 的速度有快有慢,各自都有自己的优势,所以无法判断模型的好坏。

(4) 需要更加完善的体系来减少梯度损失、模型不稳定等因素带来的干扰。

6 结 语

本文简要介绍了 GAN 的基本概念、代表模型及主要应用领域,从多个角度对 GAN 和传统算法进行深入的对比与分析,阐述了 GAN 在当前社会的研究意义以及未来前景预测。

注:本文通讯作者为叶汉平。

参 考 文 献

- [1] 杨曼婷. 基于 GAN 的图像生成算法研究[D]. 淮南: 安徽理工大学, 2021.
- [2] 高杰. 基于生成对抗网络的图像修复算法研究[D]. 南京: 南京邮电大学, 2021.
- [3] 陈旭. 基于生成对抗网络的姿态归一化人脸识别方法研究[D]. 重庆: 重庆邮电大学, 2021.
- [4] KRIZHEVSKY A, SUTSKEVER I, HINTON G E. ImageNet classification with deep convolutional neural networks[C]// Proceedings of the 25th Annual Conference on Neural Information Processing Systems. Lake Tahoe: MIT Press, 2012: 1106-1114.
- [5] WU J, QIAN X Z. Application of compact deep convolutional neural network in image recognition[J]. Journal of frontiers of computer science and technology, 2019, 13 (2): 275-284.
- [6] 林健巍. YOLO 图像检测技术综述[J]. 福建电脑, 2019, 35 (9): 80-83.
- [7] 梁俊杰, 韦舰晶, 蒋正锋. 生成对抗网络 GAN 综述[J]. 计算机科学与探索, 2020, 14 (1): 1-17.
- [8] CRESWELL A, WHITE T, DUMOULIN V, et al. Generative adversarial networks: an overview [J]. IEEE signal processing magazine, 2018, 35 (1): 53-65.
- [9] 汪美琴, 袁伟伟, 张继业. 生成对抗网络 GAN 的研究综述[J]. 计算机工程与设计, 2021, 42 (12): 3389-3395.
- [10] RADFORD A, METZ L, CHINTALA S. Unsupervised representation learning with deep convolutional generative adversarial networks [C]// ICLR 2016. San Juan, Puerto Rico: ICLR, 2016.
- [11] BROCK A, DONAHUE J, SIMONYAN K. Large scale GAN training for high fidelity natural image synthesis [C]// ICLR 2019. NEW Orleans, LA, United States: ICLR, 2019.
- [12] KARRAS T, LAINE S, AILA T. A style-based generator architecture for generative adversarial networks[C]// 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Long Beach, CA, USA: IEEE, 2019.
- [13] ZHANG H, XU T, LI H S, et al. StackGAN: text to photo-realistic image synthesis with stacked generative adversarial networks[C]// 2017 IEEE International Conference on Computer Vision (ICCV). Venice, Italy: IEEE, 2017: 5908-5916.
- [14] ZHU J Y, PARK T, ISOLA P, et al. Unpaired image-to-image translation using cycle-consistent adversarial networks[C]// 2017 IEEE International Conference on Computer Vision (ICCV). Venice, Italy: IEEE, 2017: 2223-2232.
- [15] ISOLA P, ZHU J Y, ZHOU T H, et al. Image-to-image translation with conditional adversarial networks[C]// Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Los Alamitos: IEEE Computer Society Press, 2017: 5967-5976.
- [16] 裴莉莉, 孙朝云, 孙静, 等. 基于 DCGAN 的路面裂缝图像生成方法[J]. 中南大学学报(自然科学版), 2021, 52 (11): 3899-3906.
- [17] 夏珺, 周湘贞, 隋栋. 基于循环生成对抗网络的机器翻译方法研究[J]. 南京师大学报(自然科学版), 2022, 45 (1): 104-109.
- [18] 苏依拉, 王昊, 贺玉玺, 等. 基于对抗学习的蒙汉神经机器翻译[J]. 计算机系统应用, 2022, 31 (1): 249-258.
- [19] 王一凡, 赵乐义, 李毅. 基于生成对抗网络的图像动漫风格化[J]. 计算机工程与应用, 2022, 58 (18): 104-110.
- [20] 赵志舜. 基于生成对抗网络的人脸转正和多姿态人脸生成方法研究[D]. 广州: 广东工业大学, 2021.
- [21] LEDIG C, THEIS L, HUSZÁR F, et al. Photo-realistic single image super-resolution using a generative adversarial network[C]// IEEE Conference on Computer Vision and Pattern Recognition. Piscataway, NJ, USA: IEEE, 2017: 4681-4690.
- [22] DOSOVITSKIY A, BROX T. Generating images with perceptual similarity metrics based on deep networks [C]// Advances in neural information processing systems 29. [S.l.]: NIPS, 2016: 658-666.
- [23] KARRAS T, AILA T, LAINE S, et al. Progressive growing of gans for improved quality, stability, and variation [C]// ICLR 2018. Vancouver, Canada: IEEE, 2019: 3177-3181.

- [24] HUANG R, ZHANG S, LI T, et al. Beyond face rotation: global and local perception gan for photorealistic and identity preserving frontal view synthesis [C]// IEEE International Conference on Computer Vision. New York, NY: IEEE, 2017: 2458-2467.
- [25] WU J, ZHANG C, XUE T, et al. Learning a probabilistic latent space of object shapes via 3d generative-adversarial modeling [C]// 30th Annual Conference on Neural Information Processing Systems. Barcelona, Spain: Neural Information Processing Systems Foundation, 2016: 82-90.
- [26] YE H R A, CHEN C, YIAN L T, et al. Semantic image inpainting with deep generative models[C]// 2017 IEEE Conference Computer Vision and Pattern Recognition (CVPR) . Honolulu, HI, United States: IEEE, 2017: 6882-6890.
- [27] 周作为. 结合图像与文本的条件图像合成研究 [D]. 北京: 北京交通大学, 2020.
- [28] SI C, WANG W, WANG L, et al. Multistage adversarial losses for pose-based human image synthesis [C]// Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. New York, USA: IEEE, 2018: 118-126.
- [29] 刘仕焜. 基于 C-GAN 模型的自动驾驶车辆轨迹预测 [J]. 综合运输, 2021, 43 (9): 85-90.
- [30] SHEN Y, JI R, ZHANG S, et al. Generative adversarial learning towards fast weakly supervised detection[C]// Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. New York, USA: IEEE, 2018: 5764-5773.
- [31] SONG Y, MA C, WU X, et al. Vital: visual tracking via adversarial learning[C]// Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. New York, USA: IEEE, 2018: 8990-8999.
- [32] WANG X, LI C, LUO B, et al. Sint++: robust visual tracking via adversarial positive instance generation[C]// The IEEE Conference on Computer Vision and Pattern Recognition (CVPR) . New York, USA: IEEE, 2018: 4864-4873.
- [33] 魏富强, 古兰拜尔·吐尔洪, 买日旦·吾守尔. 生成对抗网络及其应用研究综述 [J]. 计算机工程与应用, 2021, 57 (19): 18-31.

作者简介: 刘鹤丹 (1983—), 女, 锡伯族, 辽宁人, 博士, 副教授, 主要研究方向为计算机视觉、机器学习。

叶汉平 (2001—), 男, 福建人, 本科生, 主要研究方向为计算机视觉、机器学习。

徐梦真 (2002—), 女, 福建人, 本科生, 主要研究方向为计算机视觉、机器学习。

赵旭磊 (2000—), 男, 浙江人, 本科生, 主要研究方向为计算机视觉、机器学习。

《物联网技术》杂志栏目简介

特别专访: 紧跟行业热点, 邀请行业专家、企业领袖针对我国物联网产业核心技术、标准体系、产业培育、应用示范、信息安全、公共服务等方面进行权威分析、解读, 为读者了解物联网产业现状、发展趋势提供参考。

行业资讯: 跟踪报道物联网产业重要信息, 涵盖国家、地方物联网扶持政策, 重大项目实施。物联网行业标准制定进展; 新产品及其特点介绍、新技术研发动态。

专题介绍: 物联网促进工业、农业、流通、生态环境、安全生产、交通管理、公共安全、城市基础设施管理、智能家居等传统行业信息化、智能化提升方面的顶层设计、系统化论述; 全球各国物联网产业相关, 如传感网、ITC、云计算、大数据专题解析。

方案案例: 精选物联网技术在智慧政务、智能安防、智能交通、智能电网、智能物流、智慧工业、智能医疗、智能家居、智慧社区等方向的典型解决方案以及具有示范意义的成功应用案例。

iCAN 论坛: 与国际大学生物联网创新创业大赛合作, 每期优选大赛获奖作品, 写实报道大学生令人惊喜的创新成果, 带您一览高校学子“传递 iCAN 理念、激发创新热情、点燃创业梦想”的中国梦。

学术研究: 在物联网技术全面感知、可靠传输、智能处理与应用三个层面, 刊载物联网技术学术研究成果, 为各高校、科研机构、企业研发部门、行业用户搭建交流平台, 促进物联网技术的创新、交流和产业繁荣。