计算机科学与探索
Journal of Frontiers of Computer Science and Technology doi: 10.3778/j.issn.1673-9418.1909058

生成对抗网络的研究进展综述

吴少乾，李西明+
华南农业大学 数学与信息学院，广州 510642
+ 通讯作者 E-mail: liximing@scau.edu.cn

摘　要：自生成对抗网络诞生以来，对其的研究已经成为机器学习领域的一个热点。它利用对抗学习的机制训练模型，解决了当年生成算法无法解决的问题。由于 GANs 的优势，研究者们对其进行深入地研究，产生了许多 GANs 的衍生模型，这使得 GANs 得到了快速的发展，形成了所谓的 GAN-Zoo。GANs 被广泛应用于视觉领域、音频领域、自然语言领域及其他各种领域中，如图像生成、图像翻译、文本生成、音频转换和自然语言翻译等。从传统 GANs 出发，对近几年内 GANs 的研究中较为突出的方面进行介绍总结，首先介绍了传统 GANs 的基本理论，然后对近年来 GANs 的主要衍生模型进行总结，最后总结了 GANs 在图像领域和信息安全领域中的主要应用成果。
关键词：生成对抗网络；散度函数；神经网络；生成模型
文献标志码：A **中图分类号：**TP30

吴少乾，李西明. 生成对抗网络的研究进展综述[J].计算机科学与探索
WU S Q, LI X M. A Survey on the Research Progress of Generating Adversarial Networks[J]. Journal of Frontiers of Computer Science and Technology

A Survey on the Research Progress of Generating Adversarial Networks

WU Shaoqian, LI Ximing+
College of Mathematics and Informatics, South China Agricultural University, Guangzhou 510642, China

Abstract: Since the birth of Generative Adversarial Networks, the research on it has become a hot spotin the field of machine learning. It uses the mechanism training model of Adversarial Learning to solve the problem that the generation algorithm cannot solve. Due to the advantages of GANs, researchers have conducted in-depth research on it and produced a large number of derivative models of GANs, which empower the rapid development of GANs and the formation of so-called GAN-Zoo. GANs is widely used in visual field, audio field, natural language field and other fields, such as image generation, image translation, text generation, audio conversion, natural language translation and so on. Based on the traditional GANs, we introduce and summarizes the prominent aspects of GANs research in recent years. First, we introduce the basic theory of GANs, the summarize the main derivative models of GANsin recent years, and finally summarize the main application results of GANs in image field and information security field.
Key words: GANs; Divergence function; Neural networks; Generative model

* The National Natural Science Foundation of China under Grant No. 61872152 (国家自然科学基金); the National Natural Science Foundation of China under Grant No. 61872409 (国家自然科学基金); 2018 Provincial Rural Revitalization Strategy Special Project of Guangdong Provincial Department of Agriculture under Grant No.54 (2018 年广东省农业厅省级乡村振兴战略专项项目(粤农计〔2018〕54 号)); Guangdong Program for Special Support of Top-notch Young Professionals under Grant No. 2015TQ01X79 (广东省特支计划).

1. 背景

近年来，伴随着 Alpha Go 成为第一个击败人类职业围棋世界冠军的人工智能机器人开始，机器智能得到了迅速发展。而机器智能需要我们应用大量的数据去训练机器，使其更加智能。其中，生成算法通常被视为机器对训练数据“理解”程度的衡量标准，生成算法通过学习一组信息之后生成信息分布，通过深度学习领域中的基于反向传播和 Dropout 算法的判别式模型，将高维丰富的感知器输入映射到类别标签上[1,2]，以此来判别生成算法生成的分布的正确性，若生成的分布正确，那么机器也一定正确地理解了这些信息。因此，生成算法被看作是对系统理解的一种度量，常应用于生成文本[3,4]、图像[5,6]、语音[7,8]及生成系统行为和系统状态[9]等领域。

然而，由于在最大似然估计和相关策略方面通常存在许多难以解决的概率计算问题，且在生成上下文时无法充分利用分段线性单元的优势，导致深度生成模型没有得到广泛的应用。但是，在统计信号处理和机器学习中，以产生或再现与真实样本难以区分的样本为目标的研究依旧是一大热点。特别是，获取高维数据分布的生成模型是一项具有挑战性但又非常重要的工作，因为它们对于各种应用程序（如文本到演讲合成、语音转换、图像到图像的翻译和照片编辑等)都具有重要的意义。

2014 年，Goodfellow 等人[10]提出了一种通过对抗训练来评估生成模型的新框架——生成对抗网络(Generative Adversarial Networks, GANs)。它的出现为解决工程和数学领域中高维度概率密度分布中采样和训练的问题提供了很大的帮助，迅速成为人工智能学界一个热门的研究方向。GANs 的基本思想源自双人博弈论，由一个生成器和一个判别器构成，通过对抗学习的方式进行训练，目的是估测数据样本的潜在分布并生成新的数据样本，这使得 GANs 在各个领域得到迅速的发展，被广泛应用于各个算法、模型和神经网络中，从朴素贝叶斯到深度信念网络，再到高斯混合模型、隐马尔可夫模型、潜在狄利克雷分配和受限玻尔兹曼机等。

近年来，对于 GANs 的基础研究和实际应用研究，都取得了相当不错的成果。本文将对 GANs 近几年来的研究成果进行总结，介绍 GANs 的基础理论及其实际应用成果。第二节，介绍传统 GANs 的基本理论。第三节，介绍由传统 GANs 衍生而来的比较出名的衍生模型。第四节，介绍 GANs 在图像和信息安全领域中的应用。第五节，总结全文，展望 GANs 未来的发展方向。

2. GANs 的基本原理

2.1. GANs 原理

GANs 是一种深度生成模型，它包含两个相互竞争的神经网络模型：生成模型 G 和判别模型 D，其中 G 的目的是学习真实样本的分布，生成相似度逼近真实样本的生成样本，而 D 的目的则是判断训练样本来自真实样本还是生成样本。简单来说，两个模型通过不断地对抗训练，G 最大化生成以假乱真的生成样本，D 最小化自己判别错误的概率。

模型 G 和 D 通过不断地对抗训练，使 D 正确判别训练样本来源的概率最大化，同时，使 G 生成的生成数据与真实数据的相似度最大化。在训练优化 D 时，当 D 的输入为真实样本 x 时，希望 D(x)趋于 1；当输入为生成样本 G(z)时，则希望$D\big(G(z)\big)$趋于 0，即希望$1-D\big(G(z)\big)$趋于 1，故极大化模型 D。在训练优化 G 时，输入只有噪声z，此时希望生成样本$G(z)$通过 D 后的概率值为 1，即希望$D\big(G(z)\big)$趋于 1，也就是说希望$1-D\big(G(z)\big)$趋于 0，故极小化模型 G。因此，D 和 G 的训练可以表示为关于值函数$V(G,D)$的极小化极大的双方博弈问题：

$$\min_G \max_D V(D,G) = \mathbb{E}_{x\sim p_{data}(x)}[\lg D(x)] + \mathbb{E}_{x\sim p_z(z)}\left[\lg\left(1-D\big(G(z)\big)\right)\right]$$

训练过程中，两个模型交替迭代，先固定 G，训练 D，更新 D 的参数，然后固定 D，更新迭代 G 的参数，最终达到模型稳定。Goodfellow 等人在[14]中证明了当且仅当$p_z = p_{data}$时，极大化极小的双方博弈问题存在全局最优解，即达到纳什均衡，此时生成模型 G 学会了真实样本p_{data}的分布，使得判别模型 D 的准确率稳定停留于$1/2$上，即使得 D 只能对训练样本在 0 或 1 之间进行随机猜测。

2.2. GANs 优缺点

相比较于以前的建模框架，GANs 框架具有以下的一些优势。相比较于蒙特卡洛估计、玻尔兹曼机和生成随机网络等生成网络，GANs 的计算过程不需要使用马尔可夫链，学习过程也不需要近似推理，这使

得它能够更好地利用分段线性单元的优势，仅通过反向传播来计算梯度，从而回避了近似计算困难的概率难题。与完全明显的信念网络相比，GANs 因为不需要在采样序列生成不同的数据，使其能够更快的产生样本。相比于变分自编码器，GANs 没有变分下界，不需要引入任何决定性偏置，是渐进一致的，而变分方法引入的决定性偏置是有偏差的，由于变分自编码器优化的是对数似然的下界，而不是似然度本身，从而导致变分自编码器的生成效果比 GANs 差。相比较于非线性独立分量分析，GANs 不需要对潜在变量（生成器的输入值）的大小进行限制，即不要求生成器输入的潜在变量有任何特定的维度或要求生成器是可逆的。

GANs 虽然解决了生成式模型中的一些问题，对其他生成算法的发展具有一定的启发意义，但 GANs 也并非完美，在解决了已有问题的同时也产生了一些新的问题。由于 GANs 采用对抗学习的方法，导致模型收敛性的不稳定，虽然 GANs 在纳什均衡时达到最优，但是只有当梯度下降在凸函数的情况下才能保证实现纳什均衡。而训练过程需要保证两个对抗网络的平衡和同步，否则难以得到很好的训练效果，但在实际过程中，两个对抗网络的同步不易把控，因此训练过程可能不稳定，这就导致了模型训练很难收敛的问题。此外，GANs 框架使用极大化极小的概念，这使得在训练过程中难以对模型训练情况进行评价。GANs 在训练过程中，生成模型可能出现退化现象，总是生成同样的样本点，无法继续学习，导致生成模型奔溃，而当生成模型崩溃时，判别模型也会对相似的样本点指向相似的方向，使得训练无法继续，从而导致了 GANs 的模型崩溃问题。相比其它生成式模型，GANs 无需预先建模，直接对真实样本进行采样训练，从而真正达到理论上可以完全逼近真实样本，这是 GANs 的最大优势，但同时也这会使得模型过于自由。而对于尺寸大、像素高的图片，简单的 GANs 又会导致模型不太可控，这就导致了 GANs 模型过于自由不可控的问题。

Table 1 Classification of GANs derived models

表 1 GANs 的衍生模型分类

基于损失函数改进的 GANs	f-GANs、Least-Square GANs、Loss-sensitive GAN、WGAN、WGAN-GP、WGAN-LP、DRAGAN、BEGAN、Fisher GAN、EBGANs 等	
基于模型应用改进的 GANs	基于网络架构改进的 GANs	CGAN、DCGAN、InfoGAN、StackGAN、AL-GAN、ASGAN、CycleGAN、DualGAN 等
	基于编码器改进的 GANs	BEGAN、VAE-GAN、BiGAN、tDCGAN、Adversarial Autoencoders、[32]、[33]、[40]等
	GANs 的其他改进	LAPGAN、MGAN、3D-GAN、SRGAN、ESRGAN 等

3. GANs 的衍生模型

自 2014 年 Goodfellow 等人提出生成对抗网络以来，GANs 已经成为最流行的深度生成模型之一，其研究也取得了快速的进展，衍生出了上百种相关的模型，在短短的几年里已在图像处理、自然语言、计算机视觉等领域中得到了广泛的应用，最终形成了所谓的 GANs-zoo。本节将从 GANs 模型中的损失函数、网络体系结构及其模型应用改进等方面的研究进展进行整理分析，并介绍其中具有代表性的模型，并将一些比较常见的 GANs 衍生模型进行分类，如表 1 所示。

3.1 基于损失函数改进的 GANs

Nowozin 等人[11]对 Nguyen 等人[12]提出的变散度估计框架进行扩展，提出f − GANs模型，将散度估计扩展到模型估计，并称这种新方法为变分发散最小化（VDM），并证明了生成对抗训练是 VDM 框架的一个特例。他们使用生成模型 Q 和变分函数 T 作为生成对抗网络的两个模型。通过实验表明，GANs 架构中的 JS 散度可以被任意的f-散度（不同散度的总称）代替，证明了 GANs 的广泛应用性。虽然使用交叉熵的定义可以保证模型收敛的速度，但是在决策边界为真的一边的样本均会被分类为真，即使是假样本，这样就会导致即使使用与真实样本相差甚远的假样本更新生成器时，判别器仍会判断为真，这就让假样本成功

欺骗了判别器。从而导致了更新生成器的时候出现梯度弥散的问题，进而使得传统 GANs 生成的图片质量不高以及训练过程不稳定的系列问题。针对以上问题，Xudong Mao 等人在[13]中提出了一种最小二乘生成对抗网络（LSGANs），该网络将传统 GANs 的目标函数中的交叉熵损失函数替换成最小二乘损失函数，并将最小二乘损失函数作为判别器，根据样本到决策边界的距离进行惩罚，从而将假样本移向决策边界，这样就会生成更多的梯度来更新生成器，同时保证了模型的收敛性和稳定性，以此解决了 GANs 中梯度消失的问题，但 LSGANs 无法解决这个核心问题——如何更好地测量生成数据和真实数据之间的散度距离。

针对散度距离测量的问题，Arjovsky 等人[14]提出的 WGAN 抛弃了传统 GANs 的 JS 散度定义，采用 Earth Mover 距离（EM 距离）来计算两个分布的距离，利用 EM 距离来监控模型的好坏，解决 GAN 训练不稳定、模型崩塌和生成模型的评价问题。WGAN 在处理 Lipschitz 限制条件时采用了权重裁剪方法，将参数的梯度值限制在一定的范围内，以便通过倒数限制 D(x)的增长速度，但是该方法中的权重裁剪值却不好确定，这使得有时仍然只能生成较差的样本或无法收敛。针对 WGAN 存在的缺点，Gulrajani 等人在[15]中提出了 WGAN 的改进版 WGAN-GP，WGAN-GP 提出了一种新的 Lipschitz 连续性限制手法，即通过给判别函数添加梯度惩罚项，将参数与限制联系起来，达到 Lipschitz 限制条件，以此解决 WGAN 的权重裁剪导致参数集中化、梯度爆炸和梯度消失的问题，但是对于梯度的模大于 1 的区域的值，WGAN-GP 虽然也做出了惩罚，但是却无法保证每一个值的梯度的模都小于等于 1，并且该方法计算成本很高。

以上所述均存在一个较强的假设，即鉴别器在每一步中（在函数空间中）都处于最优状态，相比较 WGAN 和 WANG-GP，Petzka 等人[16]提出一种较弱的正则化项来执行 Lipschitz 限制。而 Kodali 等人[17]提出的 DRAGAN 则将交替梯度更新过程看作 regretminimization 进行训练，以达到纳什均衡，并且证明了在无参数限制和无需要求判别器在每一步中都处于最优状态的条件下，模型训练能够渐近收敛，在此基础上提出了一种新的梯度惩罚算法，支持更快的训练，实现更好的稳定性和建模性能，但本质上只是采用了不同的梯度惩罚函数。同样，Berthelot 等人在[18]中提出了一种新的平衡强化方法——BEGAN，该方法结合 EM 距离的损失来训练基于自动编码器的 GANs。BEGAN 抛弃了通过估计真实分布和生成分布之间的差距，反而通过估计分布的分布误差之间的相似度，以达到判别目的，但是 BEGAN 主要应用于图像领域中，对于高分辨率的图像的效果一般。为此，在 BEGAN 的基础上，Yanchun Li 等人[19]在判别器中新增去噪损失函数，以获取更多与真实分布相关的信息，以此提高训练效果和收敛稳定性，但其效果并未超过 WGAN-GP。

在[20]中，Mroueh 等人利用奇异值分解概念，将分布嵌入到有限维特征空间中，并根据其均值和协方差特征统计进行匹配，同时匹配均值特征和二阶矩阵特征，使得真实数据分布和生成数据分布之间的嵌入协方差差异最大化，以此提升训练效果。在[21]中，Mroueh 等人建立了一种积分概率矩阵框架——FisherGAN 来训练 GANs，它在 WGAN-GP 的基础上继续改进，当 WGAN-GP 对判别器函数的梯度做出惩罚时，Fisher GAN 会对判别器的二阶矩阵进行约束，使得模型可以稳定训练，但是 FisherGAN 对于积分概率度量(the Integral Probability Metrics,IPM)的约束较强，缺乏一定的灵活性。在神经网络中，将梯度的模限制在一个范围内，抽象地来说就是让产生的函数更平滑一些，最常见的做法便是正则化，而频谱归一化则可以使正则化产生更明确的限制，因此，Miyato 等人在[22]中提出了频谱归一化 GANs，利用频谱范数[23]标准化神经网络的参数矩阵 W，从而让神经网络的梯度被限制在一个范围内，以便让正则化产生更明确地限制，但这同时也牺牲了模型的收敛速度。

3.2 基于模型应用改进的 GANs

上一小节主要介绍了从损失函数方面进行改进的各种 GANs 模型，通过各种改进，不但稳定了模型的收敛性，同时也解决了模型崩溃的问题。而在这一小节中，将介绍从模型的内部网络体系结构和结合模型应用等方面进行改进的各种 GANs 衍生模型。

在无监督学习中，从大量无标记数据集中学习到可重复使用的特征是十分有意义，特别是在在计算机视觉领域中，如果能从大批量无标记图像和视频中学习到良好的中间特征，就可以将它用于诸如图像分类这样的监督学习任务。要建立图像的良好特征，可以通过训练无监督学习的 GANs 模型，并把判别模型和

生成模型作为特征提取器，然后再应用到有监督学习上。因此，在基于 GANs 的思想上，2015 年 Radford 等人[24]提出了一种将有监督学习中的深度卷积神经网络 CNNs 和无监督学习的 GANs 结合在一起的架构——深度卷积对抗生成网络（Deep Convolutional GANs, DCGANs）。通过在不同训练集上训练表明，不论是判别模型还是生成模型，也不论是单个对象还是图像全局场景，DCGANs 都能学习到一系列特征，同时 GANs 训练的稳定性以及生成结果质量都有了极大的提升，DCGANs 也因此建立起了 CNNs 在有监督学习和无监督学习之间的桥梁。

训练传统 GANs 需要找到具有连续高维参数的非凸博弈的纳什均衡，而传统 GANs 通常使用梯度下降法进行训练，这种算法虽然可以找到损失函数的最小值，但是无法找到博弈的纳什均衡。并且当该算法用于寻找纳什均衡时，该算法可能无法收敛，即使做了结构细化的 DCGANs，其训练过程仍难以收敛。因此，Salimans 和 Goodfellow 在[25]中提出 Improved-DCGANs，针对 DCGANs 的训练过程，利用特征匹配、小批量判别、历史平均、单侧标签平滑和虚拟批次正态化五种不同的增强方法，使训练朝着收敛方向进行。

针对 GANs 无需预先建模，导致模型过于自由不可控的问题，Mirza 和 Osindero[26]提出了一种给 GANs 加上约束条件的模型，简称 CGANs。CGANs 就是一种带条件约束的 GANs，在 G 和 D 的模型中均引入条件变量 y，通过将 y 作为 G 和 D 输入层的一部分来进行调节，以此提高对模型的控制。这个改进虽然简单，但是被证明非常有效，并广泛用于后续的相关工作中[27,28,29]。已有的文本图像生成方法可以生成比较粗糙的模型，但是却无法生成必要的细节和生动的物体，虽然 CGANs 的出现提高了生成图像的质量，但是 CGANs 仍无法产生高质量的图片。因此，Zhang 等人[6]在 CGANs 的基础上提出了 StackGAN 来生成高质量的图片。StackGAN 通过堆叠的方法来实现目的，即将合成过程分解为两个较易处理的阶段。第一阶段的 GANs 利用文本描述粗略勾画物体主要的形状和颜色，生成低分辨率的图片；第二阶段的 GANs 修正了第一阶段的结果，生成细节丰富的高分辨率图片，但是 StackGAN 无法处理复杂文本。为了进一步提高生成样本的质量，稳定 GANs 的训练，Zhang 等人[30]提出了对 StackGAN 进一步改进后的 StackGAN++，相比于 StackGAN，StackGAN++由多个生成器和多个鉴频器组成，它们以树状结构排列，让同一场景对应的多个尺度的图像由树的不同分支生成，以此提升模型对复杂文本的处理能力。但是 StackGAN++仍无法很好地处理很复杂的文本，因此，Johnson 等人[31]提出了一种从场景图生成图像的端到端的方法，以处理更长更复杂的文本，与从文本描述生成图像的方法相比，该方法能够从结构化场景图中明显地推理出对象和关系，并生成具有多个可识别对象的复杂图像。

Donahue 等人[32]在 GANs 的基础上，结合编码器的定义提出一种双向生成对抗网络，不但能将潜在样本映射到生成的数据，而且能够将数据逆映射到潜在表示上。在 CGANs 的基础上，Perarnau 等人[33]则将编码器和 CGANs 相结合，提出了用于图像编辑的可逆的 CGANs，将真实图像映射到高特征空间和条件表示中，这就允许根据任意的属性重构和修改真实图像，即能够重新生成具有确定性的复杂修改的真实图像。在[2]中，Karacan 等人利用反卷积神经网络和卷积神经网络构造了新的条件 GANs——属性-布局条件生成对抗性网络 (Attribute-Layout Conditioned GAN，AL-CGAN)。AL-CGAN 模型被拆解成两部分研究，即单属性条件的 A-CGAN 模型和单空间布局条件的 L-CGAN 模型。同时还会对缺失的空间布局进行补充，对每个场景进行粗语义标注，对缺失的属性进行属性预测。

上述表明了深度卷积网络能够提升 GANs 生成高分辨率图片的细节，但是由于卷积网络的局部感受的限制，如果要生成大范围相关的区域，卷积网络就会出现问题。Qian 等人[34]提出将 AttentiveNet 引入传统的 GANs 中，以此生成注意力图(attentive map)，并将注意力图应用于生成网络和判别网络中，使网络能够快速准确地定位到图像中的重点关注区域。但由于 GANs 很难捕捉几何结构特征，使得 GANs 在某些图像类别上很难建模，因此依靠卷积来建立不同图像区域之间的依赖关系的模型，其依赖关系的传递只能通过大范围的多个卷积层来实现。随着卷积大小的增加，网络的真实容量也在增加，但却损失了计算效率。而 self-attentive，却能够做到依赖性和计算效率的平衡。因此，Zhang 等人[35]提出将 self-attention 机制加入 GANs 中，让生成器和判别器可以自动学习图像中的重要目标，形成了模型 SAGAN。SAGAN 克服了传统

GANs 模型均在低分辨率特征图的空间局部点上来生成高分辨率的细节的缺陷——SAGAN 可以从所有的特征处生成细节，并且 SAGAN 的判别器可以判别两幅具有明显差异的图像是否具有一致的高度精细特征，但仍有很大的提升空间。此外，Brock 等人[36]提出的 BigGAN 以 SAGAN 架构为基础，将正交正则化应用于生成器，使其适用于简单的“截断技巧”，并使用铰链损失作为 GANs 目标函数，同时使用类条件 BatchNorm 和含投影的 D 向 G 提供类信息，该方法极大地提高了生成图像的逼真性和精细度。

上诉介绍的各种改进只是众多改进中的一小部分，除此之外，研究者们利用各种相关的方法继续改进传统 GANs，以改善传统 GANs 的各种不足，例如：Ghosh 等人[37]利用多主体的 GANs 来改进传统 GANs 模型崩塌的问题；Zhao 等人[38]引入能量定义，提出了基于能量地生成对抗网络（EBGAN），该模型将判别器看作一个能量函数，将低能量赋给数据流行附近地区域，将高能量赋给其他区域，结果表明，EBGAN 比传统的 GANs 更加稳定，生成的图像分辨率更高；Guo−Jun Qi 等人[39]提出的损失敏感型生成对抗网络，进一步利用真实数据密度的 Lipschitz 正则条件对其损失函数进行正则化处理，得到一个正则化模型，该模型提升了 GANs 生成新数据的泛化能力；Yujia Li 等人[40]将最大均值差异作为损失函数，并结合自动编码器，结果表明它能够比传统 GANs 更有效地生成更好的生成模型；Tolstikhin 等人[41]使用自提高的训练机制，通过融合对各生成模型，达到更好的拟合数据分布的目的，从而克服了 GANs 训练难及模型崩溃问题。

4. GANs 的应用

GANs 作为一种生成模型，它并不局限于特定的数据类型，可以应用于各种数据，如图像、音频、文本等。它也不局限于特定的任务，可以应用于各种任务，如图像的处理、视频生成、恶意检查、通信保护和密码破译等。

4.1. GANs 在图像领域中的应用

针对 GANs 模型过于自由问题，除了给其加上约束条件外，我们也可以通过让 GANs 分次完成任务，一次生成一部分，分多次生成完整的目标，来避免这个问题。因此，L.Denton 等人[42]采用这个思想，在基于 CGANs 的基础上，提出了改进模型 LAPGANs。LAPGANs 是一种生成式参数模型，通过带有 Laplacian 金字塔框架的级联卷积网络，以逐步求精的迭代方式生成高质量的自然图像样本，LAPGANs 训练出来图像比 GANs 训练出来的图像更加自然，边缘也更加明确，但由于 LAPGANs 的网络深度较深，使得逐步求精的方法增加了网络简单记忆输入样本的难度。Chuan Li 等人[43]将对抗性生成式网络应用于 Markovian 环境中，学习相同内容的不同描述之间的映射，通过生成网络 G 将输入的图像直接解码为合成图像的像素，然后利用判别器 D 去学习区分实际的特征块和不合适的合成特征块，通过对抗性训练反卷积神经网络来合成纹理，但是该方法在处理非纹理的数据上的表现不足。

GANs 模型不仅适用于二维空间，而且适用于三维空间。但不同于 GANs 在二维空间的应用，在三维空间中，GANs 的生成器需要建立从二维空间到三维空间的映射关系，并依据该映射关系将 2D 图像生成 3D 形状的对象，这个映射关系既是重点也是难点，直接影响 3D 对象的生成质量。Wu 等人[44]在 DCGANs 的基础上进行改进，提出了 3D 生成对抗网络(3D-GAN)，它利用生成对抗网络和体卷积网络[45]，建立从概率空间到 3D 对象空间的映射，再利用判别器为三维对象的描述提供有用的特征信息，但是合成的 3D 图像的锐度和边缘信息还有待提高。Gadelha 等人[46]提出的投影生成对抗网络（PrGANs），利用其训练一个可以准确表示 3D 形状的深度生成模型，将 2D 图像生成 3D 对象，并在传递给判别器之前将其转换为 2D 图像，通过迭代训练周期，生成器通过改进生成的 3D 体素形状来完善 2D 图像到 3D 对象的投影结果。将 GANs 应用于三维空间上的探索还有很多，例如：Jie Cao 等人[47]利用一种 3D 辅助二重生成对抗网络(AD-GAN)来精确地将人脸图像旋转到任意指定的角度；Weiyue Wang 等人[48]结合 3D 解码器 GANs 和长期循环卷积网络，以低分辨率填充缺失的三维数据，使三维模型具有语义合理性和上下文细节。

随着神经网络在无监督学习和半监督学习领域上的发展，我们可以通过对输入样本的重构来对数据分布进行显式建模，但是基于重构的学习方法往往会学习并保存全部输入样本的特征。在 GANs 的基础上，Xin Yi 等人[49]提出了无监督和半监督学习下的分类生

成式对抗网络框架，即将神经网络分类器与对抗生成模型相结合，对抗生成模型对训练有素的分类器进行正则化，以此使得与分类器一起学习的生成器能够生成高视觉保真度的图像。但由于传统 GANs 的输入是一组没有任何限制且完全随机的噪声信号 z，这使得我们无法将 z 的具体维度和数据的语义特征对应起来，导致输入向量对输出产生不明确性的影响。因此，Xi Chen 等人在[50]中提出了无监督学习的 InfoGANs，他们将原本的输入向量 z 拆成子向量 c（可解释的隐变量，表示对输出产生影响的因素）和子向量z'（不可压缩的噪声）。InfoGANs 通过约束隐变量 c 与生成数据之间的关系，使 c 能够直接代表数据某个方面的语义信息，进而使得 c 与生成数据具有较高的互信息，以此解决隐变量可解释性的问题。

为了提高图像的分辨率，Ledig 等人[51]提出了一种超分辨率生成对抗网络（SRGAN），实现了低分辨图像合成 4 倍放大的高分辨率图像，但生成的图像的纹理信息并不够真实，且常常伴有噪声。因此，Xintao Wang 等人[52]提出了一种增强分辨率的生成对抗网络（ESRGAN），对网络的结构、对抗性损失和感知性损失进行改进，并在此基础上，引入残差-剩余密集块作为基本的网络构建单元，得益于这些改进，为亮度一致性和纹理恢复提供更强的监控，产生的纹理比 SRGAN 更真实和自然，进一步提高了视觉的质量。GANs 应用于图像超分辨率的例子还有很多，例如：Huang 等人[53]提出的双通道生成对抗网络通过感知全局结构和局部细节，实现了从一张单侧照片合成高分辨率的正面人脸图像；Kupyn 等人[54]提出了一种基于 CGANs 和内容损失的端到端学习法 DeblurGAN，能够去除由于运动而产生的模糊；Kupyn 等人对 DeblurGAN 进行改进得到 DeblurGAN-v2[55]，大大提高了去模糊效率、质量和灵活性。

图像到图像的转换是图像领域中一大热门研究方向，GANs 在该方向的研究也取得了一定的进展。Phillip Isola 等人[28]在 CGANs 的基础上进行图像到图像的转换的研究，利用成对的数据集进行实验，表明了 CGANs 是一种很有前途的图像到图像的转换方法，特别是那些涉及高度结构化图形输出的图像。图像到图像的转换通常使用一组成对的图像对来学习输入图像和输出图像之间的映射，但是，对于许多任务，成对的训练数据很难获取，因此，Zhu 等人[56]提出了一种在没有成对数据的情况下，设计双生成器与双鉴别器，形成双向风格迁移的 CycleGAN 模型，学习将图像从域 X 转换为目标域 Y 的方法，解决了 pix2pix 必须使用成对数据的问题。Yu-Sheng Chen 等人[57]提出一种利用非成对数据来实现图像增强的方法，其基于一个结构类似于 CycleGAN 的双向 GANs，并进行改进，以此来实现图像到增强图像的转换。

除了上诉应用外，GANs 在各个领域的应用也十分广泛，Xiaolong Wang 等人[58]对图像生成过程进行了分解，提出了一种基于风格和结构的生成反求网络；Creswell 等人[59]将 GANs 应用于检索；Jun-Yan Zhu 等人[60]提出利用生成式对偶神经网络直接从数据中学习自然图像流形；包仁达等人[61]提出了一种掩模控制的自动上妆 GANs，能够重点编辑上妆区域，约束人脸妆容编辑中无需编辑的区域不变，保持主体信息；Reed 等人[9]综合了图像和描述在哪个位置绘制什么内容的指令；Vondrick 等人[62]利用大量未标记的视频来学习视频识别任务(例如动作分类)和视频生成任务(例如未来预测)的场景动力学模型；Taigman 等人[63]研究了将一个域上的样本转移到另一个域上的模拟样本的问题，可以应用于包括数字和人脸图像在内的视觉领域；Denton 等人[64]介绍了一种简单的基于对抗性损失的画中学习的半监督学习方法；Mogren 等人[65]提出了一种基于连续序列数据的生成式对抗模型，并将其应用于古典音乐的集合；武随烁等人[66]针对 GANs 无法有效提取图像局部与全局特征间依赖关系、以及各类别间的依赖关系，提出一种用于 GANs 的孪生注意力模型；桑亮等人[67]基于 GANs 的深度卷积网络，采用端对端的方式，复原了由于抖动或运动所导致的模糊图像的细节信息；时澄等人[68]利用 GANs 对残缺图像进行补全。除此之外，GANs 的应用数不胜数，但仍需要我们继续研究。

4.2 GANs 在信息安全领域中的应用

在数据隐私保护中，如何保证数据集的可利用性和隐私性的平衡是极为重要的。随着 GANs 应用领域的不断扩大，其在隐私保护中也有所应用。GANs 利用自身的优势，将噪声添加到潜在空间而不是直接添加到数据中，减少了整体的信息损失，同时保证了隐私。Triastcyn 等人[69]提出了一种生成人工数据集的方法，在生成对抗网络的判别器中加入高斯噪声层，使输出和梯度相对于训练数据具有不同的私密性，然后

利用生成器组件合成具有保密性的人工数据集，不但保留了真实数据的统计特性，同时为这些数据提供了差分隐私保护。Beaulieu-Jones 等人[70]结合 GANs 和差分隐私提出了差分隐私辅助分类生成对抗网络，用于生成医疗临床数据。针对当 GANs 应用于私人或敏感数据（如病人的医学病历），并且分布的集中可能泄露关键的病人信息的问题，LiyangXie 等人[71]提出一种差分私有 GANs 模型——DPGAN 模型，通过在学习过程中向梯度中添加精心设计的噪声来实现 GANs 的差分隐私。同样，Chong Huang 等人[72]结合 GANs 提出一种上下文感知隐私模型，通过巧妙添加噪声来实现私有数据的发布。Frigerio 等人[73]通过差异隐私定义提出了一个保护隐私的数据发布框架，从时间序列到连续数据和离散数据的生成，均可以很容易地适应不同的用例，以保证在发布新的开放数据的同时保护用户的个性。Nicolas 等人[74]提出了一种教师——学生模式的深度网络隐私保护方法，利用教师深度模型和学生 GANs 模型，通过训练从而达到保护训练数据集的目的，但是这种将模型训练过程视为黑盒，仅从模型外部添加隐私保护机制，其保护效果的可控性是不够的。

GANs 除了应用于隐私保护外，还被应用于恶意检测中。为了有效地检测包括零日攻击在内的恶意软件，Kim 等人[75]提出了一种转移深度卷积生成对抗网络(tDCGAN)，基于深度自编码技术，利用实际数据和 tDCGAN 生成的修改数据学习各种恶意软件的特征，提取有意义的特征进行恶意软件检测。GANs 同样适用于信用卡欺骗检测，Fiore 等人[76]训练一个 GANs 模型来输出模拟的少数类的欺骗例子，然后将这些例子与训练数据合并成一个增强训练集，从而提高分类器对少数类欺骗例子的分类效率。为了有效识别受害者向诈骗者发送大额转账，Yujun Zheng 等人[77]提出了一种基于 GANs 的模型来计算银行每笔大额转账的欺诈概率，让银行采取适当的措施，以此防止潜在的骗子在概率超过阈值时取钱。除了这些恶意检测，GANs 还可以用于检测僵尸网络，Chuanlong Yin 等人[78]提出了一种基于 GANs 的僵尸网络检测模型框架——Bot-GAN，生成模型不断生成伪样本，以辅助原检测模型提高性能。

GANs 的广泛适用性，使其在信息安全领域中最为古老的密码学中同样也适用。2016 年，Abadi 等人[79]利用 GANs 的对抗学习机制，将传统的对称加密体系中的通信双方及敌手用神经网络进行代替，以此实现加解密的过程，达到保护通信过程的目的。在[80]中李西明等人改进 Abadi 等人的模型，进行了抗泄漏加密通信测试，发现了利用生成对抗网络实现抗泄漏加密通信的可能性。Coutinho 等人[81]利用选择明文攻击的概念改进了 Abadi 等人的模型，证明了神经网络在适当的环境下可以学习一次性密码本。除了保护通信之外，GANs 同样也适用于密码破译。Gomez 等人[82]基于 GANs 提出 CipherGAN，用于破译古典密码学中移位密码和维吉尼亚密码，Hitaj 等人[83]提出了一种利用 GANs 来增强密码破译的新方法——PassGAN，通过在泄露密码列表中训练 GANs 来实现密码破译，为 GANs 在密码学上的应用前景提供了更广阔的道路。

5. 讨论

自 2014 年 GANs 的诞生以来，它作为一种新型的生成模型得到了广泛的应用。作为一个生成模型，GAN 模型避免了一些传统生成模型在实际应用中的一些困难，巧妙地通过对抗学习来近似一些不可解的损失函数，这是其最大的创新之处也是其产生问题的根源。针对其产生的各种问题，研究者们近几年不断地进行研究分析，将 GANs 与各种相关的方法进行改进融合，得到了许多有效的成果。但目前的 GANs 仍存在不足及缺陷，有赖于研究者们进一步的解决。在应用上，GANs 被广泛应用于图像、文本生成、自然语言和信息安全等领域中，其中应用最为热门的领域当属图像领域。

虽然 GANs 已经在许多方面得到应用，但其仍存在广大的应用前景，特别是在信息安全领域。在未来的工作中，我们可以从以下几点深入研究 GANs 在信息安全领域中的应用。

1) 利用 GANs 的特性，结合图像水印技术，将信息附着在图像上，进行传输，达到消息的安全传递的效果。
2) 利用 GANs 的对抗学习机制，分析病毒样本，预测新型病毒，形成自动防御体系。
3) 将 GANs 与经典密码学相结合，分析破译率，并利用 GANs 的对抗学习机制，保证公钥密码体制下的安全通信，实现消息的可认证性和保密性。

致谢：

本文受国家自然科学基金（61872152）、国家自然科学基金（61872409）、2018 年广东省农业厅省级乡村振兴战略专项项目（粤农计〔2018〕54 号）和广东省特支计划科技创新青年拔尖人才项目（2015TQ01X79）资助。

参考文献：

[1] Harm de Vries, Florian Strub,Jérémie Mary, et al. Modulating early visual processing by language[C]//In Advances in Neural Information Processing Systems 30: Annual Conference on Neural Information Processing Systems 2017, 4-9 December 2017, Long Beach, CA, USA, 2017, 6597–6607.

[2] Levent Karacan, Zeynep Akata, Aykut Erdem, et al. Learning to generate images of outdoor scenes from attributes and semantic layouts[DB/OL]. (2016-12-01). CoRR, abs/1612.00215.

[3] Vincent Dumoulin, Jonathon Shlens, Manjunath Kudlur. A learned representation for artistic style[DB/OL]. (2016-10-24). CoRR, abs/1610.07629.

[4] Dustin Tran, Rajesh Ranganath, David M. Blei. Hierarchical implicit models and likelihood-free variational inference[C]// In Advances in Neural Information Processing Systems 30: Annual Conference on Neural Information Processing Systems 2017, 4-9 December 2017, Long Beach, CA, USA, 2017: 5529-5539.

[5] Jonas Gehring, Michael Auli, David Grangier, et al. Convolutional sequence to sequence learning[C]//In Proceedings of the 34th International Conference on Machine Learning, ICML 2017, Sydney, NSW, Australia, 6-11 August 2017, 2017: 1243-1252.

[6] Han Zhang, Tao Xu, Hongsheng Li. Stackgan: Text to photo-realistic image synthesis with stacked generative adversarial networks[C]//In IEEE International Conference on Computer Vision, ICCV 2017, Venice, Italy, October 22-29 2017, 2017: 5908-5916.

[7] Oliver Bendel. The synthetization of human voices[J]. AI Soc. 2019, 34(1): 83-89.

[8] Deep neural networks for acoustic modeling in speech recognition: The shared views of four research groups[J]. IEEE Signal Process. Mag., 2012, 29(6): 82-97.

[9] Scott E. Reed, Zeynep Akata, Santosh Mohan, et al. Learning what and where to draw[C]//In Advances in Neural Information Processing Systems 29: Annual Conference on Neural Information Processing Systems 2016, December 5-10 2016, Barcelona, Spain, 2016: 217-225.

[10] Ian J. Goodfellow, Jean Pouget-Abadie, Mehdi Mirza, et al. Generative Adversarial Networks[DB/OL]. (2014-06-10). CoRR, abs/1406.2661.

[11] Sebastian Nowozin, Botond Cseke, Ryota Tomioka. f-gan: Training generative neural samplers using variational divergence minimization[C]//In Advances in Neural Information Processing Systems 29: Annual Conference on Neural Information Processing Systems 2016, December 5-10 2016, Barcelona, Spain, 2016: 271-279.

[12] X. Nguyen, M. J. Wainwright, M. I. Jordan. Estimating divergence functionals and the likelihood ratio by convex risk minimization. Information Theory[J]. IEEE, 2010, 56(11): 5847-5861.

[13] Xudong Mao, Qing Li, Haoran Xie, et al. Least Squares Generative Adversarial Networks[J]. ICCV 2017: 2813-2821

[14] Martín Arjovsky, Soumith Chintala, Léon Bottou. Wasserstein GAN.[DB/OL].(2017-01-26). CoRR, abs/1701.07875.

[15] Ishaan Gulrajani, Faruk Ahmed, Martín Arjovsky, et al. Improved Training of Wasserstein GANs[C]//In Advances in Neural Information Processing Systems 30: Annual Conference on Neural Information Processing Systems 2017, 4-9 December 2017, Long Beach, CA, USA, 2017: 5769-5779.

[16] Henning Petzka, Asja Fischer, Denis Lukovnicov. On the regularization of Wasserstein GANs.[DB/OL].(2017-09-26).CoRR, abs/1709.08894.

[17] N Kodali, J Abernethy, J Hays, et al. On Convergence and Stability of GANs[DB/OL]. (2017-05-19). CoRR, abs/1705.07215.

[18] David Berthelot, Tom Schumm, Luke Metz. BEGAN: boundary equilibrium generative adversarial networks [DB/OL]. (2017-03-31). CoRR, abs/1703.10717.

[19] Yanchun Li, Nanfeng Xiao, Wanli Ouyang. Improved boundary equilibrium generative adversarial networks[J]. IEEE Access, 2018, 6:11342-11348.

[20] Youssef Mroueh, Tom Sercu,Vaibhava Goel. Mcgan: Mean and covariance feature matching GAN[C]//In Proceedings of the 34th International Conference on Machine Learning, ICML 2017, Sydney, NSW, Australia, 6-11 August 2017, 2017: 2527-2535.

[21] Youssef Mroueh, Tom Sercu. Fisher GAN[C]//In Advances in Neural Information Processing Systems30: Annual Conference on Neural Information Processing Systems 2017, 4-9 December 2017, Long Beach, CA, USA, 2017, 2510-2520.

[22] Takeru Miyato,Toshiki Kataoka, Masanori Koyama, et al. Spectral normalization for generative adversarial networks. [DB/OL]. (2018-02-16). CoRR, abs/1802.05957.

[23] Takeru Miyato, Shin-ichi Maeda, Masanori Koyama, et al. Virtual adversarial training: a regu-larization method for supervised and semi-supervised learning[DB/OL]. (2018-04-13). CoRR, abs/1704.03976.

[24] Alec Radford, Luke Metz, Soumith Chintala. Unsupervised representation learning with deep convolutional generative adversarial networks[DB/OL]. (2015-11-19). CoRR, abs/1511.06434.

[25] Tim Salimans, Ian J. Goodfellow, Wojciech Zaremba, et al. Improved Techniques for Training GANs[C]//In Advances in Neural Information Processing Systems 29: Annual Conference on Neural Information Processing Systems 2016, December 5-10, 2016, Barcelona, Spain, 2016: 2226-2234.

[26] Mehdi Mirza, Simon Osindero. Conditional generative adversarial nets[DB/OL].(2014-11-06). CoRR, abs/1411.1784.

[27] Augustus Odena, Christopher Olah, Jonathon Shlens. Conditional image synthesis with auxiliary classifier gans[C]//In Proceedings of the 34th International Conference on Machine

Learning, ICML 2017, Sydney, NSW, Australia, 6-11 August 2017, 2017: 2642-2651.

[28] Phillip Isola, Jun-Yan Zhu, Tinghui Zhou, et al. Image-to-image Translation with Conditional Adversarial Networks[C]// In 2017 IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2017, Honolulu, HI, USA, July 21-26, 2017: 5967-5976

[29] Vladimir A. Knyaz, Vladimir V. Kniaz, Fabio Remondino. Image-to-voxel model translation with conditional adversarial networks[C]//In Computer Vision - ECCV 2018 Workshops-Munich, Germany, September 8-14, 2018, Proceedings, Part I, 2018: 601-618.

[30] Han Zhang, Tao Xu, Hongsheng Li, et al. StackGAN++: Realistic Image Synthesis with Stacked Generative Adversarial Networks[J]. IEEE Trans. Pattern Anal. Mach. Intell, 2019, 41(8): 1947-1962.

[31] Justin Johnson, Agrim Gupta, Li Fei-Fei. Image Generation from Scene Graphs[C]//In 2018 IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2018, Salt Lake City, UT, USA, June 18-22, 2018, 2018: 1219-1228.

[32] Jeff Donahue, Philipp Krähenbühl, Trevor Darrell. Adversarial feature learning[C]//In 5th International Conference on Learning Representations, ICLR 2017, Toulon, France, April 24-26, 2017, Conference Track Proceedings, 2017.

[33] Guim Perarnau, Joost van de Weijer, Bogdan Raducanu, et al. Invertible conditional gans for image editing[DB/OL]. (2016-11-19). CoRR, abs/1611.06355.

[34] Rui Qian, Robby T. Tan, Wenhan Yang, et al. Attentive generative adversarial network for raindrop removal from a single image[DB/OL]. (2017-11-28). CoRR, abs/1711.10098.

[35] Han Zhang, Ian J. Goodfellow, Dimitris N. Metaxas et al. Self-attention generative adversarial networks[DB/OL] (2018-05-28). CoRR, abs/1805.08318.

[36] Andrew Brock, Jeff Donahue, Karen Simonyan. Large scale GAN training for high fidelity natural image synthesis[DB/OL]. (2018-09-28). CoRR, abs/1809.11096.

[37] Arnab Ghosh, Viveka Kulharia, Vinay P. Namboodiri, et al. Multi-agent Diverse Generative Adversarial Networks[C]//In 2018 IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2018, Salt Lake City, UT, USA, June 18-22, 2018, 2018: 8513-8521.

[38] Junbo Jake Zhao, Michaël Mathieu, Yann LeCun. Energy-based Generative Adversarial Networks[C]//In 5th International Conference on Learning Representations, ICLR 2017, Toulon, France, 24-26April, 2017, Conference Track Proceedings, 2017.

[39] Guo-Jun Qi. Loss-Sensitive Generative Adversarial Networks on Lipschitz Densities[DB/OL]. (2017-01-23). CoRR abs/1701.06264

[40] Yujia Li, Kevin Swersky, Richard S. Zemel. Generative Moment Matching Networks[C]//Proceedings of the 32nd International Conference on Machine Learning, ICML 2015, Lille, France, 6-11 July 2015.JMLR Workshop and Conference Proceedings 37, 2015: 1718-1727.

[41] Ilya O. Tolstikhin, Sylvain Gelly, Olivier Bousquet, et al. AdaGAN: Boosting Generative Models[C]//Annual Conference on Neural Information Processing Systems 2017, 4-9 December 2017, Long Beach, CA, USA. Neural Information Processing Systems Conference, 2017: 5424-5433.

[42] Emily L. Denton, Soumith Chintala, Arthur Szlam, et al. Deep generative image models using a laplacian pyramid of adversarial networks[DB/OL]. (2015-06-18). CoRR, abs/1506.05751.

[43] Chuan Li, Michael Wand. Precomputed real-time texture synthesis with markovian generative adversarial networks[C]// In Computer Vision-ECCV 2016-14th European Conference, Amsterdam, The Netherlands, October 11-14, 2016, Proceedings, Part III, 2016: 702-716.

[44] Jiajun Wu, Chengkai Zhang, Tianfan Xue, et al. Learning a Probabilistic Latent Space of Object Shapes via 3D Generative-Adversarial Modeling[C]//In Advances in Neural Information Processing Systems Annual Conference on Neural Information Processing Systems 2016, December 5-10, 2016, Barcelona, Spain, 2016: 82-90.

[45] Daniel Maturana, Sebastian Scherer. Voxnet: A 3d convolutional neural network for real-time object recognition[C]//In 2015 IEEE/RSJ International Conference on Intelligent Robots and Systems, IROS 2015, Hamburg, Germany, September 28 - October 2, 2015, 2015: 922-928.

[46] Matheus Gadelha, Subhransu Maji, Rui Wang. 3D Shape Induction from 2D Views of Multiple Objects[DB/OL]. (2016-12-08). CoRR abs/1612.05872.

[47] Jie Cao, Yibo Hu, Bing Yu, et al. 3D Aided Duet GANs for Multi-View Face Image Synthesis[J]. IEEE Trans. Information Forensics and Security, 2019, 14(8): 2028-2042.

[48] Weiyue Wang, Qiangui Huang, Suya You, et al. Shape Inpainting Using 3D Generative Adversarial Network and Recurrent Convolutional Networks[C]//IEEE International Conference on Computer Vision, ICCV 2017, Venice, Italy, October 22-29, 2017. IEEE Computer Society 2017, 2017: 2317-2325.

[49] Xin Yi, Ekta Walia, Paul Babyn. Unsupervised and Semi-supervised Learning with Categorical Generative Adversarial Networks assisted by wasserstein distance for dermoscopy image classification[DB/OL]. (2018-04-10). CoRR, abs/1804.03700.

[50] Xi Chen, Yan Duan, Rein Houthooft, et al. Infogan: Interpretable representation learning by information maximizing generative adversarial nets[C]//In Advances in Neural Information Processing Systems 29: Annual Conference on Neural Information Processing Systems 2016, December 5-10, 2016, Barcelona, Spain, 2016: 2172-2180.

[51] Christian Ledig, Lucas Theis, Ferenc Huszar, et al. Photo-realistic single image super-resolution using a generative adversarial network[C]//In 2017 IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2017, Honolulu, HI, USA, July 21-26, 2017, 2017: 105-114.

[52] Xintao Wang, Ke Yu, Shixiang Wu, et al. ESRGAN: Enhanced Super-Resolution Generative Adversarial Networks[C]// Computer Vision-ECCV 2018 Workshops-Munich, Germany, September 8-14, 2018, Proceedings, Part V. Lecture Notes in Computer Science 11133, Springer, 2018: 63-79.

[53] Rui Huang, Shu Zhang, Tianyu Li, 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, ICCV 2017, Venice, Italy, October 22-29, 2017. IEEE Computer Society 2017, 2017: 2458-2467.

[54] Orest Kupyn, Volodymyr Budzan, Mykola Mykhailych, et al. DeblurGAN: Blind Motion Deblurring Using Conditional Adversarial Networks[C]//2018 IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2018, Salt Lake City, UT, USA, June 18-22, 2018. IEEE Computer Society 2018, 2018: 8183-8192.

[55] Orest Kupyn, Tetiana Martyniuk, Junru Wu, et al. DeblurGAN-v2: Deblurring (Orders-of-Magnitude) Faster and Better[DB/OL]. (2019-08-10). CoRR abs/1908.03826.

[56] Jun-Yan Zhu, Taesung Park, Phillip Isola, et al. Unpaired Image-to-Image Translation Using Cycle-Consistent Adversarial Networks[C]//IEEE International Conference on Computer Vision, ICCV 2017, Venice, Italy, October 22-29, 2017. IEEE Computer Society 2017, 2017: 2242-2251

[57] Yu-Sheng Chen, Yu-Ching Wang, Man-Hsin Kao, et al. Deep Photo Enhancer: Unpaired Learning for Image Enhancement From Photographs With GANs[C]//2018 IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2018, Salt Lake City, UT, USA, June 18-22, 2018. IEEE Computer Society 2018, 2018: 6306-6314.

[58] Xiaolong Wang, Abhinav Gupta. Generative image modeling using style and structure adver-sarial networks[C]//In Computer Vision-ECCV 2016-14th European Conference, Amsterdam, The Netherlands, October 11-14, 2016, Proceedings, Part IV, 2016: 318-335.

[59] Antonia Creswell, Anil Anthony Bharath. Adversarial training for sketch retrieval[C]//In Computer Vision-ECCV 2016 Workshops-Amsterdam, The Netherlands, October 8-10 and 15-16, 2016, Proceedings, Part I, 2016: 798-809.

[60] Jun-Yan Zhu, Philipp Kr ähenb ühl, Eli Shechtman, et al. Generative visual manipulation on the natural image manifold[C]// In Computer Vision-ECCV 2016-14th European Conference, Amsterdam, The Netherlands, October 11-14, 2016, Proceedings, Part V, 2016, 597–613.

[61] Bao Renda, Geng Han, Zhu Defa, et al. Automatic Makeup with Region Sensitive Generative Adversarial Networks[J]. Journal of Software, 2019, 30(4): 896-913.

[62] Carl Vondrick, Hamed Pirsiavash, Antonio Torralba. Generating videos with scene dynamics[C]//In Advances in Neural Information Processing Systems 29: Annual Conference on Neural Information Processing Systems 2016, December 5-10, 2016, Barcelona, Spain, 2016: 613-621.

[63] Yaniv Taigman, Adam Polyak, Lior Wolf. Unsupervised cross-domain image generation[DB/OL]. (2016-11-07). CoRR, abs/1611.02200.

[64] Emily L. Denton, Sam Gross, Rob Fergus. Semi-supervised learning with context-conditional generative adversarial networks[DB/OL].(2016-11-19).CoRR, abs/1611.06430.

[65] Olof Mogren. C-RNN-GAN: continuous recurrent neural networks with adversarial training[DB/OL]. (2016-11-29). CoRR, abs/1611.09904.

[66] CaoWenwen, Kang Bing, Yan Jun, et al. Sparse Representation Target Tracking via Multi-Source Data Fusion[J]. Computer Engineering and Application, 2019, 55(6): 1-7.

[67] Sang Liang, Gao Shuang, Yi Zengshan. Motion Deblurring Based on Generative Adversarial Networks[J]. Computer Engineering and Application, 2019, 55(6): 173-177.

[68] Shi Cheng, Pan Bin, Guo Xiaoming, et al. Application of Generative Adversarial Networks in Image Completion[J]. Journal of Frontiers of Computer Science & Technology, 2019, 13(8): 1402-1410.

[69] Aleksei Triastcyn, Boi Faltings. Generating Differentially Private Datasets Using GANs[DB/OL]. (2018-03-08). CoRR abs/1803.03148.

[70] Brett K. Beaulieu-Jones, Zhiwei Steven Wu, Chris Williams, et al. Privacy-preserving generative deep neural networks support clinical data sharing[J]. Circulation: Cardiovascular Quality and Outcomes, 2019, 12(7).

[71] Liyang Xie, Kaixiang Lin, Shu Wang, et al. Differentially Private Generative Adversarial Network[DB/OL]. (2018-02-08). CoRR abs/1802.06739.

[72] Chong Huang, Peter Kairouz, Xiao Chen, et al. Generative Adversarial Privacy[DB/OL]. (2018-07-13). CoRR abs/1807.05306

[73] Lorenzo Frigerio, Anderson Santana de Oliveira, Laurent Gomez, et al. Differentially Private Generative Adversarial Networks for Time Series, Continuous, and Discrete Open Data[DB/OL]. (2019-01-08). CoRR abs/1901.02477.

[74] Nicolas Papernot, Mart ín Abadi, ÚlfarErlingsson, et al. Semi-supervised Knowledge Transfer for Deep Learning from Private Training Data[DB/OL]. (2016-10-18). CoRR abs/1610.05755

[75] Jin-Young Kim, Seok-Jun Bu, Sung-Bae Cho. Zero-day Malware Detection Using Transferred Generative Adversarial Networks based on Deep Autoencoders[J]. Inf. Sci. 2018,460-461: 83-102

[76] Ugo Fiore, Alfredo De Santis, Francesca Perla, et al. Using generative adversarial networks for improving classification effectiveness in credit card fraud detection[J]. Inf. Sci. 2019,479: 448-455

[77] Yujun Zheng, Xiao-Han Zhou, Weiguo Sheng, et al. Generative Adversarial Network based Telecom Fraud Detection at the Receiving Bank[J]. Neural Networks, 2018, 102(6): 78-86

[78] Chuanlong Yin, Yuefei Zhu, Shengli Liu, et al. An Enhancing Framework for Botnet Detection Using Generative Adversarial Networks[C]//2018 International Conference on Artificial Intelligence. May 26-28, Chengdu, China. NJ: IEEE, 2018: 288-234

[79] Mart ín Abadi, David G. Andersen. Learning to Protect Communications with Adversarial Neural Cryptography[DB/OL]. (2016-10-21). CoRR abs/1610.06918

[80] Li Ximing, WuJiarun, Wu Shaoqian, et al. Study on adversarial encryption based on generative adversarial networks[J/OL]. Computer Engineering and Applications:1-6[2019-09-10].http://kns.cnki.net/kcms/detail/11.2127.tp.20190703.1803.008.html.

[81] Murilo Coutinho, Robson de Oliveira Albuquerque, F ábio Borges, et al. Learning Perfectly Secure Cryptography to Protect Communications with Adversarial Neural Cryptography[J]. Sensors, 2018,18(5): 1306.

[82] Aidan N. Gomez, Sicong Huang, Ivan Zhang, et al. Unsupervised Cipher Cracking Using Discrete GANs[DB/OL]. (2018-01-15). CoRR abs/1801.04883.
[83] Briland Hitaj, Paolo Gasti, Giuseppe Ateniese, et al. PassGAN: A Deep Learning Approach for Password Guessing[DB/OL]. (2017-09-01). CoRR abs/1709.00440.

附中文参考文献：

[61] 包仁达, 庾涵, 朱德发, 等. 基于区域敏感生成对抗网络的自动上妆算法[J]. 软件学报,2019,30(4):896-913.
[66] 曹雯雯, 康彬, 颜俊, 等. 面向多源数据融合的稀疏表示目标跟踪[J].计算机工程与应用,2019,55(6):1-7.
[67] 桑亮, 高爽, 尹增山. 基于生成对抗网络的运动模糊图像复原[J]. 计算机工程与应用, 2019,55(6):173-177.
[68] 时澄, 潘斌, 郭小明, 等. 生成式对抗网络在图像补全中的应用[J]. 计算机科学与探索,2019,13(8):1402-1410.
[80] 李西明, 吴嘉润, 吴少乾, 等. 基于生成对抗网络的抗泄露加密算法研究[J/OL]. 计算机工程与应用:1-6[2019-09-10]. http://kns.cnki.net/kcms/detail/11.2127.tp.20190703.1803.008.html.

WU Shaoqian was born in 1994. He is an M.S. candidate at South China Agricultural University. His research interest is Machine learning and Information security.
吴少乾(1994-),男（汉）,广东揭阳人,华南农业大学在读硕士研究生，主演研究领域为机器学习和信息安全。

LI Ximing was born in 1974. He received Ph.D. degree in cryptography from South China Agricultural Universityin 2011. His research interests include Information security, Intelligent image processing and Machine learning.
李西明(1974-),男（汉）,山东临清人, 2011 年获华南农业大学密码学博士学位。高级工程师,博士,主要研究方向为信息安全、智能图像处理和机器学习。