

# The Applications of Generative Adversarial Networks in the Field of Computer Vision

Chunyu Jiang, Lin Cheng\*, Xiaomingliang Li

School of Information Science and Engineering, Lanzhou University, Lanzhou Gansu

Email: \*chenglin@lzu.edu.cn

Received: Oct. 31<sup>st</sup>, 2018; accepted: Nov. 13<sup>th</sup>, 2018; published: Nov. 20<sup>th</sup>, 2018

---

## Abstract

The emergence of generative adversarial networks is another milestone in the field of computer vision. It provides a new tool for solving various image prediction problems. For this purpose, from the investigation of relevant literatures and latest research results, the characteristics of unsupervised learning algorithms are highlighted. At present, the topology of the generative adversarial networks is continuously optimized to its basic performance. It has become one of the hot topics of current research to play their respective strengths in improving the difficulties such as non-convergence, mode collapse and so on.

## Keywords

Generative Adversarial Networks, Computer Vision, Image Prediction, Unsupervised Learning

---

## 生成对抗网络在计算机视觉领域的应用

江春雨, 程琳\*, 黎晓明亮

兰州大学信息科学与工程学院, 甘肃 兰州

Email: \*chenglin@lzu.edu.cn

收稿日期: 2018年10月31日; 录用日期: 2018年11月13日; 发布日期: 2018年11月20日

---

## 摘要

生成对抗网络(GAN, Generative Adversarial Networks)的出现是计算机视觉领域又一里程碑式的发展, 它为了解决各种图像预测问题提供了新型工具。以此为目的, 本文通过相关文献的调研以及结合最新成果

---

\*通讯作者。

的研究, 突出了生成对抗网络的无监督学习算法特征。目前生成对抗网络拓扑结构的研究在不断优化着基础生成对抗网络的性能; 在改进难收敛、模式崩溃等缺点方面发挥着各自的长处, 已然成为当下研究的热点方向之一。

## 关键词

生成对抗网络, 计算机视觉, 图像预测, 无监督学习

Copyright © 2018 by authors and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

## 1. 引言

生成对抗网络最早是在 2014 年由 Open AI 的 Ian Good fellow 等人提出的一种创新型模型[1]。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帮助。随后经过两年对生成对抗网络的研究与总结, Ian Good fellow 等人在 2016 年神经信息处理系统进展大会上对生成对抗网络的相关知识体系进行了汇总[2]。生成对抗网络的理论在此基础上被证明了是一种具有非监督学习性质的收敛模型, 这就为其在后续的应用中提供了可行性和实践性。它能够在不需要假设数据分布的前提下, 来对一种随机分布进行直接采样, 进而得到全局最优解。生成对抗网络的优点之一是它回避了一些比较复杂的概率运算, 在训练过程中不涉及到最大似然估计和马尔科夫链等复杂算法[3], 又因为生成对抗网络善于捕捉模型的分布方式等一些独特的优点, 其概念和模型被学术界广泛接受。至此以后以生成对抗网络作为研究基础的拓扑结构以指数爆炸的速度发展着, 从 GAN ZOO 中统计的自 2014 年以来以生成对抗网络作为研究基础的论文数量的数据中可窥见一斑。

为此, 针对生成对抗网络非监督学习这一特点, 本文着重分析了非监督学习相较于监督学习的优势, 对比了一些主流的无监督生成模型; 主要总结了相对于标准生成对抗模型扩展模型——条件生成对抗网络(CGAN, Conditional Generative Adversarial Nets)、深度卷积生成对抗网络(DCGAN, Deep Convolution Generative Adversarial Nets)、循环生成对抗网络(Cycle GAN, Cycle Conditional Generative Adversarial Nets)做出概括的应用。生成对抗网络及其各种衍生模型的进一步研究必将有效解决计算机视觉问题, 而且利用生成对抗网络来提升机器理解世界的的能力, 也必将成为人工智能中的重要研究方向。

## 2. 监督学习与无监督学习

### 2.1. 深度学习的研究实质

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, 机器学习又走向了人们的视野。至此, 机器学习在经历了数十年的跌宕起伏发展之后, 又迎来了它的第二次浪潮——深度学习。如今, 深度学习已然成为了 21 世纪计算机学科的热门研究方向之一。深度学习最重要也最广泛的应用是数据挖掘(Data Mining), 它是利用大数据的方法来学习样本特征, 并对其多特征的特点进行分类或者预测[4], 从而找到一些解决复杂问题的途径并提高执行效率的方法。由此看来, 深度学习的实质是通过构建具有很多隐层的深度学习模型和海量的训练数据, 来习得更有用的特征, 从而最终提升机器分类或预测的准确性。

## 2.2.两种学习算法的区别和适用场景

从训练神经网络的方式上区分,大部分深度学习算法可以分为监督机器学习算法、无监督机器学习算法。

在通常情况下,深度学习所要训练的数据集中会包含很多特征,例如对于每一朵鸢尾花都有一个其所属品种类别。而描述的特征既可以是像鸢尾花品种这样可分类的,也可以是连续或者是二进制的[5]。以此根据预测器的特征来建立类标签(Class Label)分布的简明模型,然后对未来的实例进行预测。如果在这些实例中应用到的是已知标签,例如:Iris 数据集中注明了每个鸢尾花的样本属于何种品种,通过监督学习算法对 Iris 数据集进行研究,来学习如何根据测量结果将样本划分为三个不同品种[6]。那么这种学习被称之为监督学习。相反,如果这些实例没有标签,那么称之为非监督学习[7]。

针对于训练集中数据的特征可以将监督学习的算法分成回归和分类两种,其中分类的问题占据了监督学习中的大部分。回归的目的是为了对连续数据进行预测,通过样本数据的每一个样本特征拟合接近于真实值的预测值。图 1 中红色部分表示了利用回归算法对追加了噪声的二元函数趋势的非线性拟合曲线图。而分类是建立在回归之上,根据已知训练的样本和特征向量标签,通过计算特征参数来建立判别函数以对特征样本进行分类,其输出是离散数据。而在无监督学习中,因其训练数据集中只有特征而没有标签,只能通过数据之间的相同的某种或多种属性将他们分成相似样本的集合,这种方法称之为聚类(Clustering) [8]。在聚类分析中,希望能够找到一种算法能自动将相同元素分为紧密关系的子集或簇。

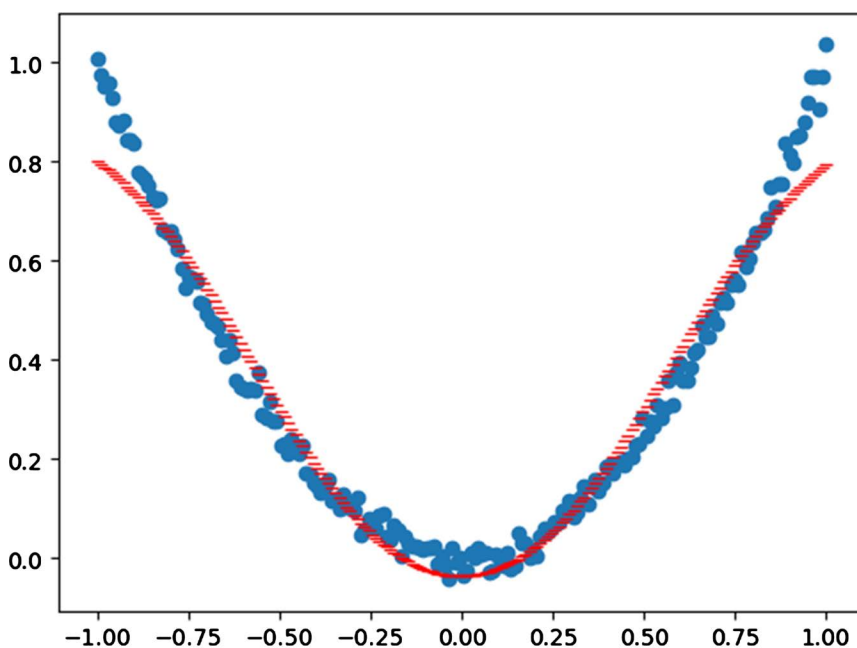


Figure 1. Nonlinear fitting curve (red line) for scatter plot of binary functions with added noise  
图 1. 对追加了噪声的二元函数散点图的非线性拟合曲线图

虽然以具有监督学习算法性质的卷积神经网络(CNN, Convolutional Neural Network)能够以很高的准确率识别物体,使得它在图像分类、物体检测、姿态估计、图像分割、人脸识别等领域崭露头角[9]。但是随着需要训练的模型大型化和复杂化,深度学习模型在训练监督型数据上往往需要花费大量的精力和时间,而且实际上需要训练的未标签数据要远远多于已标签数据,由此催生了基于无监督学习领域的发展,以期望建立一种可兼容小数据集进行训练的通用系统。

### 2.3.几种基于无监督学习生成模型算法的对比

生成模型在概率统计理论和深度学习领域是一种重要的模型，它是指利用随机的数据来生成一系列可以观测的数据模型[10]。主要的无监督学习生成模型对比如表 1 所示：

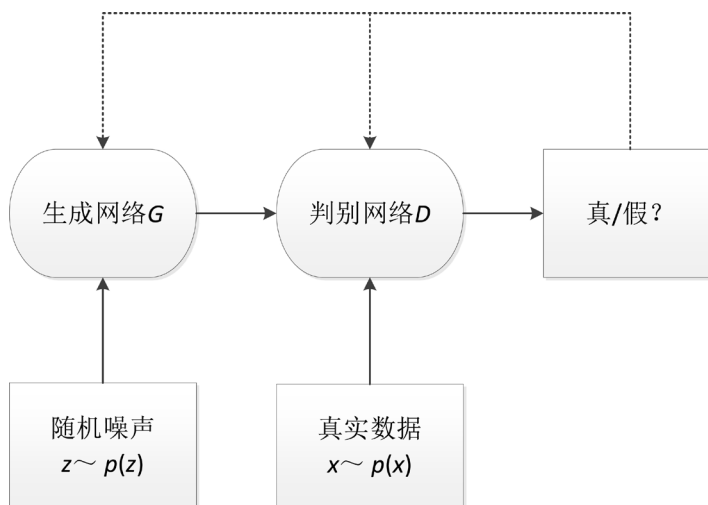
**Table 1.** Comparison of main generative models

**表 1.** 主要的生成模型的对比

名称	英文表达	所属类型	主要特征
自动编码器	Autoencoder	无监督学习	自编码器针对原始数据训练样本，利用数据的压缩算法提取特征，生成和原始数据集尽可能相同的数据。
改进自动编码器	Variational Autoencoder	无监督学习	取代标准自动编码器所产生的实数值向量，产生均值向量和标准差向量来优化 KL 散度。
生成对抗网络模型	Generative Adversarial Networks	无监督学习	包含一个生成网络和一个判别网络。两个网络在博弈的过程中，生成网络最终实现经判别网络“认证”的不同于原数据集的“自然数据集”。

### 3. 生成对抗网络模型原理与框图

生成对抗网络的思想来源于极大极小对策和纳什均衡原理，其本质上是一个对抗博弈的过程[11]。它主要实现两个基本功能：根据模型分布来提供密度估计和生成样本数据。一个完整的生成对抗网络结构包含了一个生成网络(G, Generator)和一个判别网络(D, Discriminator)，其基本结构如图 2 所示。



**Figure 2.** The basic framework of GAN

**图 2.** 生成对抗网络的结构示意图

G 的输入是一个随机分布的噪声  $z$ ，其分布用  $p(z)$  表示。它捕获真实数据  $x$  的分布  $p(x)$ ，并把随机噪声映射到新的数据空间，输出结果为  $G(z)$ ，并使  $G(z)$  尽量接近真实数据的分布  $p(x)$ 。

D 有两个输入，一个来源于真实数据中的样本  $x$ ，另一个来自生成器所产生的伪造数据  $G(z)$ 。D 按照其输出的表现形式可以归类于一个二分类器。它通过比对两个样本的数据分布，输出一个  $[0, 1]$  之间的标量，用于衡量生成器生成样本的好坏。按照设定，该值越接近于 0 表示生成的效果越差；相反，越接近于 1 表示生成效果越好。

由于最开始的 G 是从无标签的数据集中学习，输出随机分布的噪音，此时 D 的性能要远远好于 G，

所以往往需要在起初加快生成模型的收敛。随着时间的进行，在 G 提高生成能力的同时，D 也不断地提高着对生成的样本判别能力，以期望得到更加“真实”的样本。这样两个网络始终维持在一个相互竞争的状态，就同时提高了各自的能力。理论上在两个模型达到纳什均衡状态后，可以认为 G 已经学到真实的数据分布，模型达到最优。两个模型都可以通过非线性映射函数的设计来实现，例如多层感知器 (Multi-Layer Perception)，但是在实际的操作中多数情况下采用深度神经网络模型来确保达到良好的训练效果。

## 4. 生成对抗网络的应用

### 4.1. 条件生成对抗网络

GAN 的缺点之一是无法控制判别网络 G 所生成样本的形式，这是由于样本是随机产生的缘故。解决的方式之一是在 G 和 D 的输入层中附加一些类标签等辅助信息  $y$  [12]，这时 GAN 就演变成为了 CGAN。CGAN 相较于 GAN 最主要区别在于：在 CGAN 中，G 和 D 均需要对样本分布进行观察，例如在需要指定生成一张大树的图片时，D 也需要判断这是否是一张包含大树的伪造图片 [13]，这样每个网络就增加了能够描述生成或判别样本的能力。

利用 CGAN 可以实现一种图像——图像之间的转换，即图像翻译。图像翻译是计算机视觉领域图像处理不可或缺的任务之一。简言之，它能够将图像的一个域的属性迁移到另一个域的属性之中，来实现图像内容的改变。利用 CGAN 模型不仅能够学习从输入图像到输出图像之间的映射，还可以学习只有一个损失函数来训练这种映射，这就解决了传统方法中不同模型需要不同的损失函数公式的问题，因而这种模型可以作为一种通用的方法来产生高清图片。图 3 和图 4 分别是应用 CGAN 将轮廓图和像素图转化成为了对应的实物图片 [14]，以此提升了视觉效果。



Figure 3. Transform a contour map to physical map of a rucksack with CGAN  
图 3. 借助 CGAN 将包的轮廓图转换成为实物图

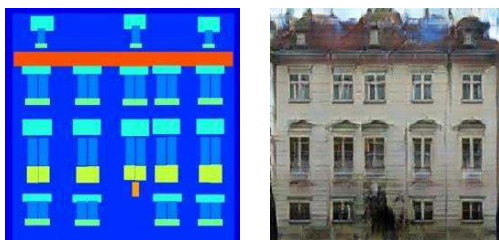


Figure 4. Transform a contour map to physical map of a building with CGAN  
图 4. 借助 CGAN 将建筑的轮廓图转换成为实物图

### 4.2. 深度卷积生成对抗网络

GAN 的另一个缺点在于生成网络 G 的输出并不稳定，经常会导致无意义的输出。同时为了进一步优化生成样本图片质量并加快模型的收敛速度，文献 [15] 将 GAN 中的生成网络 G 和判别网络 D 均设计成

为了有约束的卷积神经网络拓扑结构，其生成网络 G 的模型如图 5 所示。

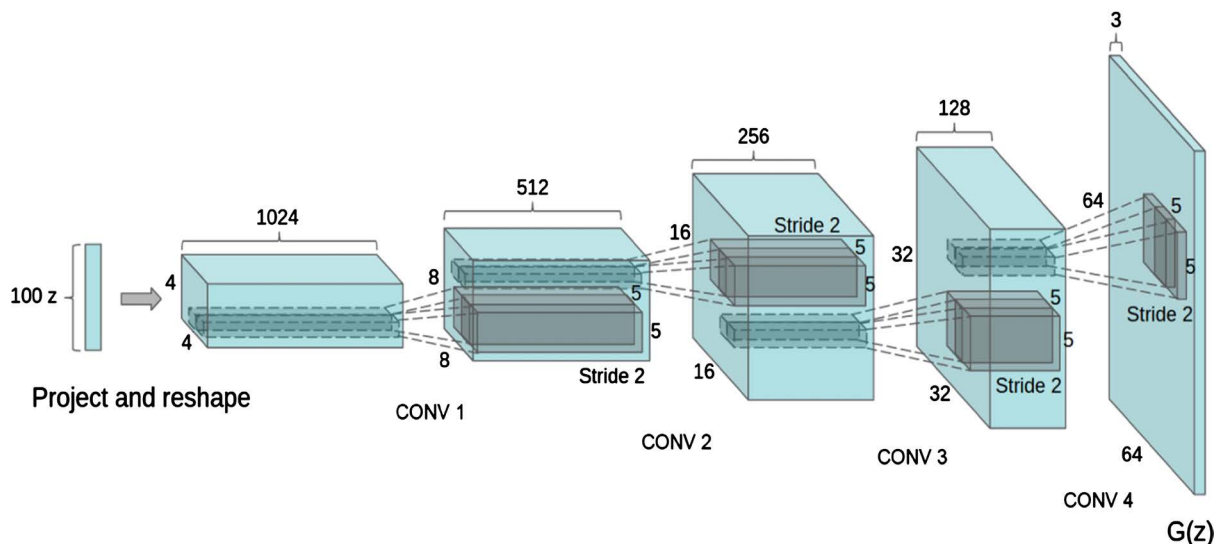


Figure 5. The generator network structure of DCGAN

图 5. 深度卷积生成对抗网络的结构示意图

训练大规模图像数据集，实现大规模景物理解，是深度卷积生成对抗网络(DCGAN, Deep Convolution Generative Adversarial Nets)所研究的课题之一。目前 DCGAN 可以在受限图像域上生成高质量的图片。

图 6 是 DCGAN 的应用实例，它生成了不同于卧室图片数据集的、看似真实却虚拟的卧室图片。由于对于图片生成的“真实性”的好坏没有唯一的客观评判标准，只能通过设置和不断优化判别网络的性能，以尽可能的提升视觉效果作为目标，因此具有一定的主观性。同时 DCGAN 也清楚地表明对输入语义属性的算术操作可以得到清晰的解释，这也是 DCGAN 研究的意义所在。



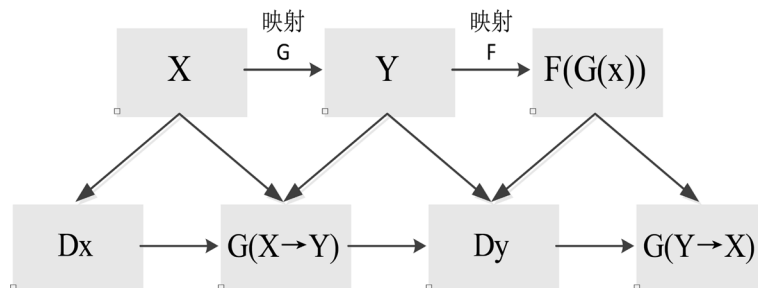
Figure 6. Virtual bedroom images generated from sample distribution by utilizing DCGAN

图 6. 利用 DCGAN 样本分布中产生的虚拟卧室图像

### 4.3. 循环生成对抗网络

Cycle GAN 是基于 CGAN 在图像——图像翻译方向上的拓展。Cycle GAN 解决了 CGAN 需要对

训练数据的难题，它提供了一种将一张图片从原域  $X$  映射到目标域  $Y$  的方法，记做  $G: X \rightarrow Y$ 。如图 7 所示。由于训练集拥有庞大的训练样本，这种映射又是缺少约束的，容易生成目标域  $Y$  中与图片不相干的图片，以此来引入它的逆映射  $F: Y \rightarrow X$  和一个循环一致性损失，最后使得  $F(G(X)) \approx X$ 。

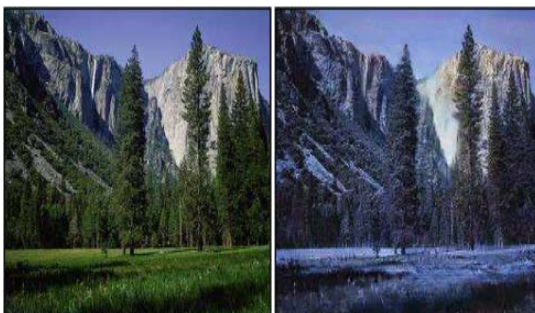


**Figure 7.** Schematic diagram of Cycle GAN  
**图 7.** 循环生成对抗网络的结构示意图

应用 Cycle GAN 可以实现对原始图片做风格转换[16]，如图 8 将一匹普通马转换成为了一匹斑马，又如图 9 则是将一张夏天风格的图片转化成为了冬天的格调。但是 Cycle GAN 对图像的边缘化处理不到位，例如在图 8 的生成图样中，并不能将处理马匹的细节和绿色的草原分离开来，所以在改变马的斑纹的同时，草地的颜色也黯淡下来。这都是生成对抗网络需要改进和加强的方面。



**Figure 8.** Transform a horse to a zebra by Cycle GAN  
**图 8.** 利用 Cycle GAN 普通马的图片转化为斑马



**Figure 9.** Transform summer to winter by Cycle GAN  
**图 9.** 利用 Cycle GAN 将夏天的景色转化为冬天风格

## 5. 总结

本文从监督学习和无监督学习的特点入手，通过比对无监督学习的主要生成模型，引入并详细地介绍了生成对抗网络模型。然后介绍了生成对抗网络在“真实”图像生成和图像翻译上有趣又实用的应用

案例,为生成对抗网络的进一步发展提供了直观的理解。总体来说,GAN及其衍生模型所生成的图片具有灵活、内容丰富等优点,既能满足某些场景下需要关注图片的细节,以此选择生成高清、高分辨率的图片;有时又能在满足一定分辨率需求的前提下,选择压缩图像节省空间的操作等。GAN在无监督训练过程中存在着难以设置初始化、不稳定、模式崩溃(Mode Collapse)和对3D图像不支持等一系列缺点,虽然在一定程度上限制了生成图样的多样性和通用性,同时为确保模型收敛增加了设计和优化网络结构的难度,但是它与监督学习、半监督学习结合起来加速训练,通过对生成器添加标签或加入控制条件等指导最终的生成结果,不仅可以弥补GAN本身的不足,还可以解决传统机器学习模型所不能解决的很多问题。因此,随着生成对抗网络理论的不间断突破、拓扑算法的扩展和各种应用的交叉完善,GAN的大家庭必定在人工智能领域大有可为。

## 参考文献

- [1] Halbritter, S. (2016) Generative Adversarial Networks. Hamburg University of Applied Science, Hamburg.
- [2] Goodfellow, I. (2016) Nips2016 Tutorial: Generative Adversarial Networks. arXiv preprint arXiv:1701.00160.
- [3] Goodfellow, I., Pouget-Abadie, J. and Mirza, M. (2014) Generative Adversarial Net. *NIPS 2014: Advances in Neural Information Processing Systems*, **27**, 2672-2680.
- [4] Cun, Y.L., Bengio, Y. and Hinton, G. (2015) Deep Learning. *Nature*, **521**, 436-444.
- [5] Kotsiantis, S.B. (2007) Supervised Machine Learning: A Review of Classification Techniques. *Informatica*, **31**, 249-268.
- [6] Goodfellow, I. and Bengio, Y. 深度学习[M]. 北京: 人民邮电出版社, 2017.
- [7] 殷瑞刚, 魏帅, 李晗, 于洪. 深度学习中的无监督学习方法综述[J]. 计算机系统应用, 2016, 25(8): 1-7.
- [8] 孙吉贵, 刘杰, 赵连宇. 聚类算法研究[J]. 软件学报, 2008, 19(1): 48-61.
- [9] 李彦冬, 郝宗波, 雷航. 卷积神经网络研究综述[J]. 计算机应用, 2016, 36(9): 2508-2515.
- [10] 百度百科.生成模型. 2018.  
<https://baike.baidu.com/item/%E7%94%9F%E6%88%90%E6%A8%A1%E5%9E%8B/6563656>
- [11] Ratliff, L.J., Burden, S.A. and Sastry, S.S. (2013) Characterization and Computation of Local Nash Equilibria in Continuous Games. In: *Proceedings of the 51st Annual Allerton Conference on Communication, Control, and Computing (Allerton)*, Monticello, IL, 2-4 October 2013, 917-924. <https://doi.org/10.1109/Allerton.2013.6736623>
- [12] Mirza, M. and Osindero, S. (2014) Conditional Generative Adversarial Nets. arXiv preprint arXiv: 1411.1784.
- [13] 黄锦池. 条件对抗网络. 2018. <https://blog.csdn.net/hjmce/article/details/54407981>
- [14] Isola, P., Zhu, J.-Y., Zhou, T.H. and Efros, A.A. (2017) Image-to-Image Translation with Conditional Adversarial Networks. *Conference on Computer Vision and Pattern Recognition*.
- [15] Radford, A. and Metz, L. (2016) Unsupervised Representation Learning with Deep Convolutional Generative Adversarial Networks. <https://arxiv.org/pdf/15-11.06434.pdf>
- [16] Zhu, J.-Y., Park, T., Isola, P. and Efros, A.A. (2017) Unpaired Image-to-Image Translation Using Cycle-Consistent Adversarial Networks. CoRR,abs/1703.10593.



**知网检索的两种方式：**

1. 打开知网页面 <http://kns.cnki.net/kns/brief/result.aspx?dbPrefix=WWJD>  
下拉列表框选择：[ISSN]，输入期刊 ISSN：2161-8801，即可查询
2. 打开知网首页 <http://cnki.net/>  
左侧“国际文献总库”进入，输入文章标题，即可查询

投稿请点击：<http://www.hanspub.org/Submission.aspx>

期刊邮箱：[csa@hanspub.org](mailto:csa@hanspub.org)