

协作式生成对抗网络

张龙¹ 赵杰煜¹ 叶绪伦¹ 董伟¹

摘要 生成对抗网络 (Generative adversarial nets, GANs) 将生成模型与判别模型进行了巧妙结合, 采用无监督的训练方式, 通过相互对抗共同提高, 其在学术界掀起了一股新的机器学习热潮. GANs 的学习目标是可以完整拟合任意真实样本的数据分布, 然而在实际当中, 真实样本分布的复杂程度难以预计, 容易发生模式坍塌 (Mode collapse) 等问题, 从而导致结果冗余, 模型不收敛等. 为提高无监督条件下的 GANs 生成能力, 减少或消除模式坍塌, 本文提出一种全新的协作式生成网络结构, 通过构建多个生成模型, 引入协作机制, 使得生成模型在训练过程中能够相互学习, 共同进步, 从而提高模型对真实数据的拟合能力, 进一步提高生成质量. 通过在三组不同类型的数据集上进行实验, 分析对比结果后发现新模型在二维图像生成方面, 特别是人脸图片, 有着显著的效果, 协作机制不仅可以加快模型收敛速度, 提高训练效率, 还能消除损失函数噪声, 在三维模型生成方面也产生了一定的影响. 通过调整模型参数, 模式坍塌问题也得到了遏制. 本文还设计了一种动态学习方法, 动态调节模型的学习速率, 有效减少了过大或过小的梯度惩罚.

关键词 生成对抗网络, 协作式, 模式坍塌, 生成模型, 无监督学习

引用格式 张龙, 赵杰煜, 叶绪伦, 董伟. 协作式生成对抗网络. 自动化学报, 2018, 44(5): 804–810

DOI 10.16383/j.aas.2018.c170483

Co-operative Generative Adversarial Nets

ZHANG Long¹ ZHAO Jie-Yu¹ YE Xu-Lun¹ DONG Wei¹

Abstract Generative adversarial nets (GANs) combine the generative model with the discriminative model. With unsupervised training methods, the two types of models mutually improve through the adversarial process. It sets off a new machine learning boom in academia. The final goal of GANs learning is to fit any real-world data distribution. In practice, however, the real-world data distribution is difficult to estimate. The major problem is mode collapse, which may lead to redundancy and non-convergence. To improve the unsupervised generator and eliminate the risk of mode collapse, this paper proposes a novel co-operative network structure for GANs. Multiple generative models are constructed with a co-operative mechanism. It can help generative models to work together and learn from each other during training. In this way, the fitting ability of generators is largely enhanced, furthermore, the quality of generated data is eventually upgraded. Experiments are conducted on three different types of benchmark datasets. Results show that the new model significantly improves image generation, especially for human face pictures. Additionally, the co-operative mechanism can speed up the convergence, improve network's learning efficiency and deduct loss function noise. It also plays a certain role in 3D model generation and suppress the problem of mode collapse. In order to solve the inconsistency between generation model and discriminative model, a dynamic learning method is developed which can dynamically adjust learning frequency. It ultimately reduces unnecessary gradient penalties.

Key words Generative adversarial nets (GANs), co-operative, mode collapse, generative model, unsupervised learning

Citation Zhang Long, Zhao Jie-Yu, Ye Xu-Lun, Dong Wei. Co-operative generative adversarial nets. *Acta Automatica Sinica*, 2018, 44(5): 804–810

在过去的十年里, 深度学习方面的研究取得了巨大进展. 它的成功影响了许多学术领域, 其相关应用例如自动驾驶、机器翻译、人脸识别等, 也极大地改善了人类的生活. 深度学习相关算法构建了一个

类似人脑结构的多层神经网络, 可以对输入信息进行特征抽取与合成等操作, 进而形成更加抽象的高维特征. 大量的实验数据证明, 深度学习算法的结果已经大大超越了传统机器学习算法, 同时可以被应用于诸多领域例如图像识别、信息检索、自然语言处理、可视数据分析等^[1–7]. 在深度学习的发展过程中, 出现了许多优秀的广为人知的网络模型, 例如 Hinton 等提出的深度置信网络 (Deep belief network, DBN) 模型^[4], LeCun 等提出的卷积神经网络 (Convolutional neural network, CNN) 模型^[8]等, 大量的研究者对这些模型进行了扩展与应用, 取得了傲人的成果.

最近几年, 一种新的网络模型又掀起了一股深度学习研究热潮, 是由 Goodfellow 等于 2014 年

收稿日期 2017-09-07 录用日期 2018-02-07
Manuscript received September 7, 2017; accepted February 7, 2018

国家自然科学基金 (61571247), 浙江省自然科学基金 (LZ16F030001), 浙江省国际合作项目 (2013C24027) 资助
Supported by National Natural Science Foundation of China (61571247), National Natural Science Foundation of Zhejiang Province (LZ16F030001), and International Cooperation Projects of Zhejiang Province (2013C24027)

本文责任编辑 李力
Recommended by Associate Editor LI Li

1. 宁波大学信息与工程学院 宁波 315211
1. Faculty of Information Science and Engineering, Ningbo University, Ningbo 315211

提出的生成对抗网络^[9-10] (Generative adversarial nets, GANs). 该网络将生成模型与判别模型进行了巧妙的结合, 生成模型对输入参数进行建模并产生数据, 判别模型则对该数据的真伪进行鉴别, 二者由此竞争, 相互促进, 生成模型产生的数据愈来愈逼真, 最后达到以假乱真的程度, 生成模型的能力也愈来愈强, 对于真假数据的判断明察秋毫. GANs 的训练过程是一个全自动的非指导性学习过程, 几乎无需人工干预. 与常用的深度学习模型 Variational autoencoders^[11] 相比, 生成对抗网络不需要预先假设数据服从一个先验分布, 优势显著, 而且在复杂图像生成方面的实验取得了显著的成绩, 因此 GANs 成为了 2016 年度 NIPS 会议最热议的主题, 被 Yann LeCun 教授称为当前深度学习研究取得的“最激动人心的进展”.

本文基于最新的 GANs 研究进展, 结合进化算法思想, 提出一种全新的协作式生成对抗网络模型, 可以显著提高图像生成质量, 并通过实验与其他方法进行了对比. 在后续的内容中, 第 1 节简要的介绍了相关工作; 第 2 节详细的描述了新模型的网络结构以及优化算法; 第 3 节是实验设计与结果; 第 4 节进行了总结.

1 相关工作

由于 GANs 是全自动的非指导性学习, 其生成图像的质量与分辨率很难达到要求. 因此在当前的诸多研究工作中, 如何提高生成图像的质量一直都是热门课题. Denton 等^[12] 应用了拉普拉斯金字塔算法, Radford 等^[10] 将深度卷积神经网络^[13] 与 GANs 进行了有效结合, 可以生成较高分辨率的人脸图像, Liu 等^[14] 构建了耦合式生成对抗网络 (Coupled generative adversarial network, CoGAN), 通过共享网络权值可以同时图像的两种属性进行学习, 这种共享机制保证了两者之间的相关性. 还有一些研究者们通过引入监督信息, 诞生了各式各样的半监督 GANs 模型. Mirza 等^[15] 首先提出了条件生成对抗网络 (Conditional generative adversarial nets, CGAN), 将标签数据的类别信息作为模型输入与生成图像一起作为判别条件. 类别信息的加入可以显著提高生成图像质量^[16], 而更加详细的图像信息 (例如说明文字、边框位置等) 则更能增强模型的生成效果^[17]. 在此基础上, 基于辅助编码器 (Auxiliary decoder) 的半监督 GANs 网络^[18-20] 也被相继提出. 著名网站 arXiv 上出现了大量相关应用的文章, 半监督学习将 GANs 的研究热度推向顶峰.

监督信息虽然可以显著提高模型的精度, 但对网络的生成能力并没有进行扩展, 而该能力恰恰是生成对抗网络的核心. 因此本文将提高无监督条

件下的 GANs 生成能力作为首要研究目标. 与此同时我们也认识到, 模式坍塌 (Mode collapse) 问题^[18, 21-22] 严重影响了网络的生成结果. 究其原因, 是由于网络的学习能力有限, 在实际训练中不能完好地拟合真实数据分布, 从而导致一些模式 (Mode) 的缺失. 解决方法不外乎两种: 1) 想办法提高网络的学习能力, 目前已有文献 [18, 22] 介绍了许多通用的方法; 2) 通过调整网络结构, 强化网络拟合不同模式的能力, 这方面的研究还比较少. Liu 等^[14] 的 CoGAN 可以强化网络学习两种指定模式的能力. Ghosh 等^[21] 则认为不同模式之间必然存在着较大的差异, 其对 CoGAN 进行了改造, 提出 MAD (Multi-agent diverse)-GAN, 通过最大化 k 个生成器间的差异, 强制网络去学习真实数据的不同模式.

与 Ghosh 的想法不同, 我们认为真实数据的不同模式之间既存在差异性也存在着相似性. 因此本文在生成器间引入一种协作机制, 使其相互之间进行学习, 在保留全局相似的同时具备局部差异, 实验证明这种全新的无监督协作式生成对抗网络不仅可以有效提高模型的生成能力, 也从另一个角度减少了模式坍塌的可能.

2 协作式生成对抗网络

2.1 生成对抗网络

生成对抗网络由一个生成器网络 G 与一个判别器网络 D 构成. 生成器的目标是拟合“真”数据 (训练样本), 并产生“假”数据, 而判别器的目标则是进行真假数据区分. 生成器与判别器的网络结构皆为多层感知器. 给定真实样本集 $\{x^1, \dots, x^n\}$, 假设 p_x 为其数据分布, 从另一个预先定义好的分布 p_z 中进行随机采样得到噪声集 $\{z^1, \dots, z^m\}$. 令生成器的输入为 z , 输出的“假”数据可表示为 $G(z)$. 判别器的输入依次为“真”“假”数据 (可根据实验情况调整比例), 输出为一个一维标量, 代表输入为真的概率, 根据输入的不同表示为 $D(x)$ 与 $D(G(z))$. 理想情况下, $D(x) = 1$ 且 $D(G(z)) = 0$. 网络优化过程可以描述为一个“二元极大极小”问题, 目标函数如下:

$$\min_G \max_D E_{x \sim p_x} \ln D(x) + E_{z \sim p_z} \ln(1 - D(G(z))) \quad (1)$$

如果将 $G(z)$ 的数据分布表示为 p_G , 那么该“二元极大极小”问题存在全局最优解, 即 $p_G = p_x$ ^[10].

生成器与判别器的训练过程是交替进行的, 更新一方的参数时, 另一方的参数固定住不更新. 通常情况下, 判别器的学习能力要强于生成器, 为了保持两者同步, 常用做法是训练生成器 k 次, 再训练判别器一次. 而通过实验发现, 两者的学习能力并不是固定比例, 而是会随着时间变化而变化, 因此在后续的

实验中本文设计了一种动态学习方法, 通过观测损失函数值的变化来保持两者同步.

2.2 协作式生成对抗网络

生成器与判别器是“对抗”关系, 此消彼长, 共同进步, 最终目标是使得生成器能够完好拟合真实样本的数据分布. 由于缺乏监督信息的指导, 该拟合过程充满了随机性, 在实际当中, 受限于网络的学习能力, 通常只能拟合出真实数据分布的一部分, 从而导致一些模式的缺失, 即模式坍塌 (Mode collapse)^[18, 21]. 如图 1 所示, 模式坍塌会导致训练结果出现冗余, 生成图像质量差等问题. 通过对真实数据进行分析不难发现, 不同模式之间存在着显著的差异, 例如人脸中的男性与女性, 场景中的白天与晚上等, 同时也存在着联系, 例如五官结构、物体形状、位置等. 强调差异而忽略联系, 或者反之, 我们认为都不算是好的解决方案, 寻求两者间的平衡是解决问题的关键.

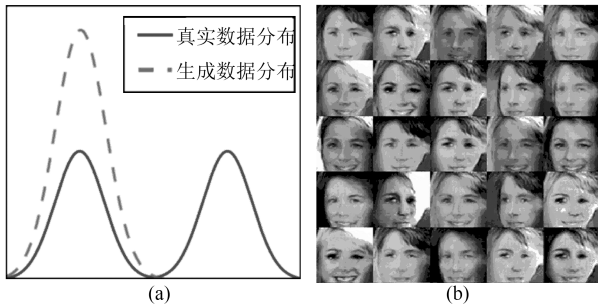


图 1 生成对抗网络中的模式坍塌问题 ((a) 生成数据分布无法完好拟合真实数据分布; (b) 模式坍塌导致生成数据冗余 (重复图像过多))

Fig. 1 Mode collapse problem in GANs ((a) synthetic data distribution cannot fit real data distribution in good shape; (b) mode collapse leads to synthetic data redundancy (too many duplicate images))

由此本文设计了如图 2 (c) 所示的网络结构. 通

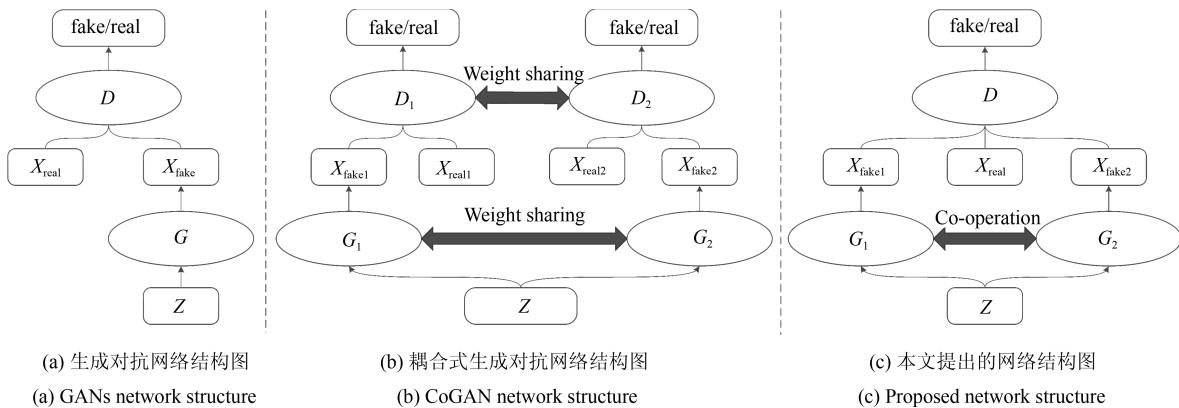


图 2 网络结构图
Fig. 2 Network structure

过构建两个 (或更多) 生成器, 共享一个输入数据 (进行协作的基础) 以及一个判别器, 同步进行训练, 训练方法与经典 GANs 相同. 此外生成器之间相互学习, 该步骤我们称为“协作”, 互为指导, 共同进步. “协作”穿插在正常训练之中, 速率可以根据实际情况进行调整, 例如训练生成器两次, 协作一次. 从数据分布的角度看, 如图 3 所示, 经典对抗式训练可以拉近真实分布与生成分布之间的距离, 而协作式训练则可以拉近不同生成器生成分布之间的距离. 这种做法不但可以提高模型收敛速度, 而且增加生成器的数量可以增强模型的学习能力, 降低模式坍塌的可能.

由于生成器间共享输入与判别器网络, 可能会造成生成器生成分布重合的现象. 这样不仅无法达到预期的目标, 还造成了额外的网络负荷. 为避免该现象, 本文在设计生成器时采取了不同的网络结构, 并进行了随机权值初始化. 重合问题在实际训练过程中未曾出现, 不同生成器产生的结果始终保持着一定的差异. 判别器的目标函数为

$$\max E_{x \sim p_x} \ln D(x) + E_{z \sim p_z} \ln(1 - D(G_1(z))) + E_{z \sim p_z} \ln(1 - D(G_2(z))) \quad (2)$$

对生成器而言, $E_{x \sim p_x} \ln D(x)$ 是不受影响的, 因此其目标函数为

$$\max E_{z \sim p_z} \ln D(G_1(z)) + E_{z \sim p_z} \ln D(G_2(z)) + \lambda L \quad (3)$$

$$L = -\|G_1(z) - G_2(z)\|_2 \quad (4)$$

其中, λ 为常数, 协作因子 L 选择了二范数来拉近生成器之间的距离 (从实验结果来看二范数略优于一范数). $D(G_1(z))$ 与 $D(G_2(z))$ 分别为生成器 G_1 和 G_2 生成数据的判别结果, 定义参数

$$s = D(G_1(z)) - D(G_2(z)) \quad (5)$$

当 $s > 0$ 时, 判别器 D 得出的结果中 G_1 的得分较高, 也就是说 $G_1(z)$ 的图像真实度更高, 此时应拉近 G_2 到 G_1 的距离, 这项操作可以通过固定住 G_1 参数, 计算协作因子 L , 惩罚 G_2 的网络连接权值完成; 而 $s \leq 0$ 时则完全相反, 需固定住 G_2 , 对 G_1 进行惩罚. 惩罚力度与 s 的大小相关. 通过这种方式, 判别得分较高的生成器对较差的产生吸引的作用力. 由于网络的随机性, 整个训练过程中 G_1 与 G_2 交替作用, 相互协助, 最终收敛于真实数据分布. 综上所述, 我们将这样的网络结构称之为“协作式生成对抗网络”.

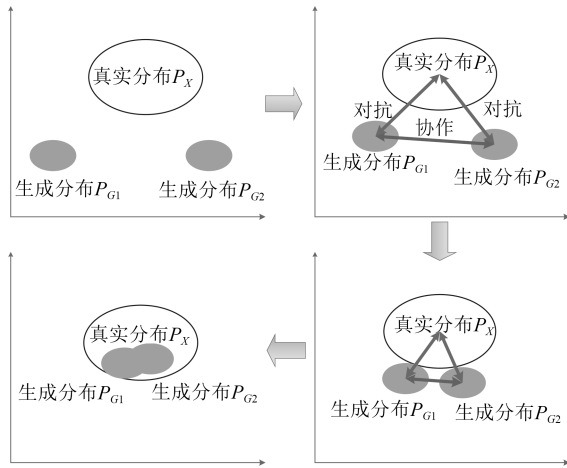


图3 本文提出的网络拟合过程

Fig.3 Fitting process for proposed networks

3 实验结果

针对本文提出的网络结构, 我们分别在 MNIST 手写体 (灰度图)、CelebA 人脸图片 (RGB 图) 以

及 ModelNet40 (三维模型) 等数据集上进行了实验. 运行环境为 Tensorflow 1.0, 显卡为 NVIDIA GFORCE GTX 1080.

3.1 MNIST 手写体

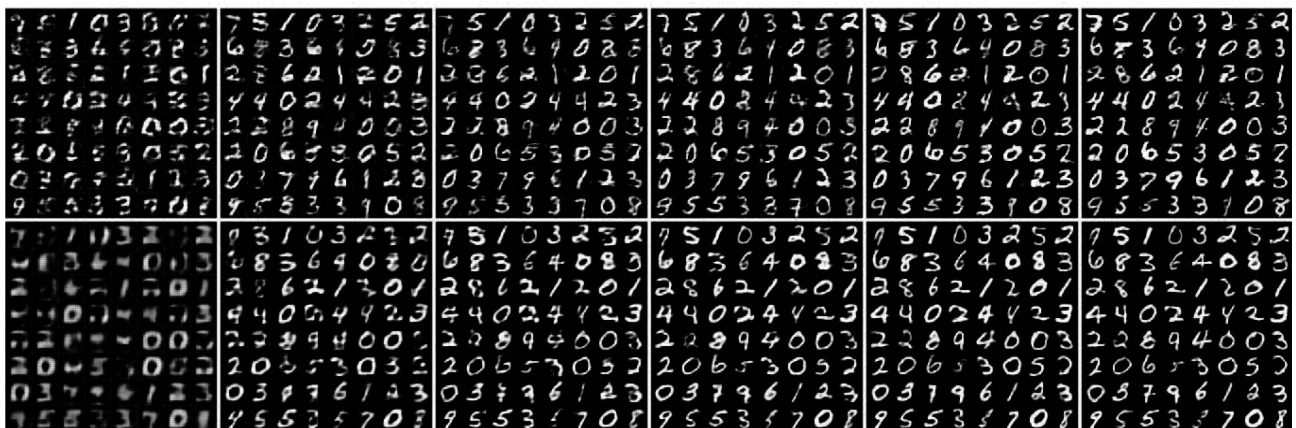
MNIST 手写体数字数据集包含从 0 到 9 的 10 类共 7 万个手写体数字图片^[23-24]. 训练结果如图 4 所示. 由于协作因子的介入, 对初期的训练结果产生了干扰, 但在迭代 1000 次之后, 协作式生成对抗网络逐步超越了传统生成对抗网络, 并在迭代 2000 次后开始收敛, 验证了本文的网络结构不仅能够增强图像生成质量, 也能提高模型收敛速度.

3.2 CelebA 人脸

CelebA 数据集包含 202599 张姿态不同、背景杂乱的人脸照片^[25-26]. 我们构建了一个生成器与判别器都是 5 层的深度卷积生成对抗网络 (Deep convolutional generative adversarial nets, DCGAN)^[9], 输入是一个 100 维的向量, 随机采样于均匀分布. 每层卷积模板的数量分别为 1024, 512, 256, 128, 3, 卷积核大小为 4×4 , 步长为 2, 生成器的输出为分辨率 64×64 的人脸图片. 训练过程中 mini-batch 设置为 64, 一个回合共 3166 个 batch.

对于式 (3) 中的常数 λ 取值问题, 通过反复的实验发现, 较小的 λ 对模型影响不大, 生成结果并没有明显改进, 而较大的 λ 则会造成生成器间的距离过于接近, 生成结果难以区分, 因此根据实验对象的不同还需要人工对该常数反复调整. 在针对 CelebA 数据集进行训练时, 最佳的 λ 取值应在 50~100 之间.

训练结果如图 5 所示, 在训练刚开始的第一个



(a) 迭代 200 次 (b) 迭代 500 次 (c) 迭代 1000 次 (d) 迭代 2000 次 (e) 迭代 4000 次 (f) 迭代 6000 次
 (a) 200 iterations (b) 500 iterations (c) 1000 iterations (d) 2000 iterations (e) 4000 iterations (f) 6000 iterations

图4 MNIST 手写体数据集训练结果 (上层采用标准生成对抗网络, 下层采用协作式生成对抗网络)

Fig.4 Training results on MNIST handwritten digits dataset (upper layer implements standard GANs, lower layer implements co-operative GANs)

回合里图 5(a)~(c), 协作式生成对抗网络展现出了微弱的优势, 人脸的轮廓已经形成, 可以模糊地分辨出不同人脸的面部特征. 从第二个回合开始图 5(d)~(h), 清晰的人脸逐步展现出来, 并且在第 4 个回合 (f), 生成的人脸已趋近真实. 反观左侧生成图像, 不仅学习速度慢, 6 个回合后约半数图像仍存在重大瑕疵.

图 6 中进一步展示了 DCGAN^[9]、MADGAN^[21] 以及本文提出网络的生成结果对比, 三者的网络层数、过滤模板数量以及激活方法保持一致. 可以直观地看出, 本文提出的方法在该数据集上生成的人脸质量显著优于前两者, 大部分图像可以“以假乱真”, 人眼几乎无法区分其真伪. 与此同时, 我们还发现图 6(a) 中方框标识的人脸相似度过高, 其原因可能是由上文中提到的“模式坍塌”问题引起的, 该现象在本文提出网络的生成结果中并未出现.

另外, 协作因子的加入可以在一定程度上减少噪声, 缩小损失函数的波动范围, 这样可以防止过大或过小的梯度惩罚. 图 7 中的数据揭示了这一现象, 损失函数值维持在一个稳定的比例可以保证“对抗”的有效性, 提高网络的学习能力.

3.3 ModelNet40 三维模型

ModelNet^[27-28] 是三维领域知名的大型数据集, 它包含 127 915 个 CAD 三维模型. ModelNet40 是其子集, 包含 12 312 个标定类别的三维模型, 分为 40 个类. 为了验证新模型在三维物体生成上同样适用, 首先, 将 ModelNet40 中的三维网格模型进行了体素化操作; 然后对第 3.2 节中的网络结构进行修改, 使其能够处理三维体素数据, 具体参数参照了 3DGAN^[29], 输入为一个 200 维向量, 随机采样于均匀分布, 生成器输出为 $64 \times 64 \times 64$ 的三维体素模

型, mini-batch 定义为 5 (数字越小效果越好, 训练速度也相对较慢).

在实验过程中, 生成器与判别器的学习速率并不成固定比例, 而是随着训练时间变化而变化, 因此常用的运行生成器 k 次再运行判别器 1 次的做法并不适用. 根据对抗特性以及损失函数值的变化幅度, 本文加入了一个动态控制学习速度的方法, 具体为:

当判别器的损失函数值小于生成器时, 判别器获胜, 训练生成器 m 次直到生成器获胜, 再训练判别器 n 次直到判别器获胜, 如此反复进行直到训练结束, 其中 $0 < m, n < \theta$ (θ 为常数).

图 8 中展示了协作式生成对抗网络的部分训练结果, 从整体结构来看, 生成的三维体素模型已经足够刻画物体的结构属性, 局部特征也表现得比较好. 但从细节来看, 每个模型都存在不同程度上的体素块缺失或盈余, 越复杂的细节越无法刻画完整. 这大概是因为体素化后的三维模型本身就缺失了大量细节信息, 以体素为基础而构建的生成模型存在先天不足; 其次, 相较于二维图片, 三维模型包含的信息量大大增加, 需要更大体量的网络结构或者更好的特征来刻画目标; 再次, 三维采集设备还不够先进与普及, 缺少大量的训练样本做支撑, 网络很容易达到过拟合.

4 总结

本文提出了一种全新的协作式生成对抗网络结构, 通过构建多个生成器, 引入协作机制, 生成器之间相互学习, 共同进步, 可以显著提高生成图像质量, 加快网络收敛速度, 去除噪声, 提高学习效率, 降低模式坍塌的可能性. 通过多组实验发现, 新的网络结构在人脸数据方面有着明显的优势, 同时在三维

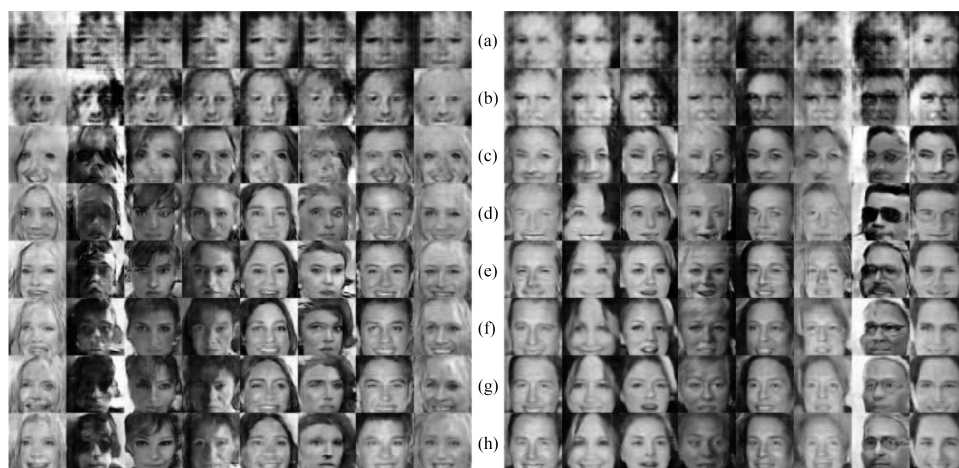


图 5 CelebA 人脸数据集训练结果 (左侧为深度卷积生成对抗网络, 右侧为协作式生成对抗网络, (a) 迭代 500 次; (b) 迭代 1 000 次; (c)~(h) 迭代 1~6 回合)

Fig. 5 Training results on CelebA human faces dataset (left side is trained by DCGAN, right side is trained by ours after, (a) 500 iterations; (b) 1 000 iterations; (c)~(h) 1~6 epochs)

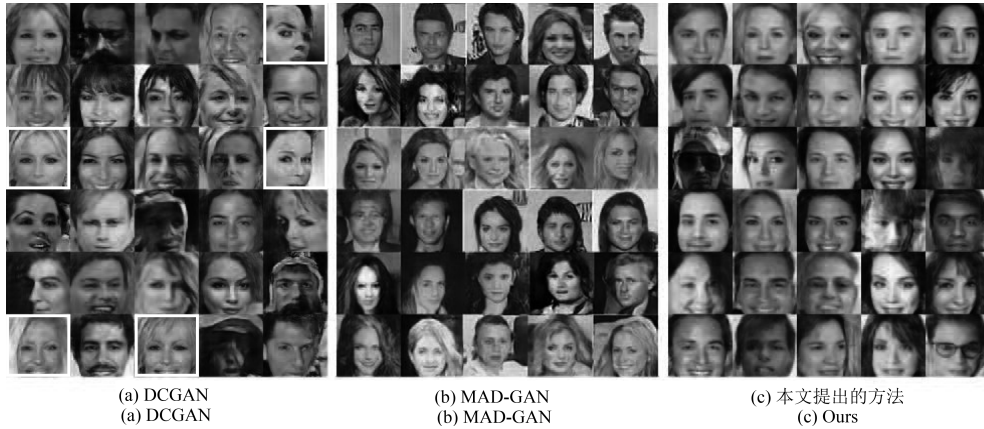


图 6 CelebA 数据集生成结果对比

Fig. 6 Comparison of synthetic data with CelebA dataset

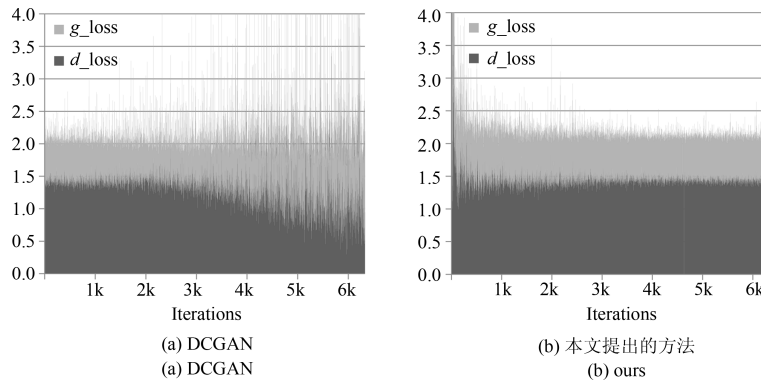


图 7 判别与生成模型的损失函数值变换情况

Fig. 7 Loss value changes of discriminator and generator models

模型生成方面也起到了一定的作用. 三维目标处理是目前研究的难点, 未来的工作重心会放在寻找替代体素、能够更好刻画三维模型的特征上面去, 同时网络结构优化与创新也同样重要.



图 8 协作式生成对抗网络在 ModelNet40 数据集的训练结果

Fig. 8 Results by co-operative GANs on ModelNet40 dataset

References

- 1 Hinton G E. To recognize shapes, first learn to generate images. *Progress in Brain Research*, 2007, **165**: 535–547
- 2 Taylor G W, Hinton G E, Roweis S. Modeling human motion using binary latent variables. In: *Proceedings of the*

- 19th International Conference on Neural Information Processing Systems. Canada: MIT Press, 2006. 1345–1352
- 3 Taylor G W, Hinton G E. Factored conditional restricted Boltzmann machines for modeling motion style. In: *Proceedings of the 26th Annual International Conference on Machine Learning*. Montreal, Quebec, Canada: ACM, 2009. 1025–1032
- 4 Hinton G E, Salakhutdinov R R. Reducing the dimensionality of data with neural networks. *Science*, 2006, **313**(5786): 504–507
- 5 Mohamed A, Dahl G E, Hinton G. Acoustic modeling using deep belief networks. *IEEE Transactions on Audio, Speech, and Language Processing*, 2012, **20**(1): 14–22
- 6 Hinton G, Deng L, Yu D, Dahl G E, Mohamed A R, Jaitly N, et al. Deep neural networks for acoustic modeling in speech recognition: the shared views of four research groups. *IEEE Signal Processing Magazine*, 2012, **29**(6): 82–97
- 7 Liu Y, Zhou S S, Chen Q C. Discriminative deep belief networks for visual data classification. *Pattern Recognition*, 2011, **44**(10–11): 2287–2296
- 8 Le Cun Y, Boser B, Denker J S, Howard R E, Hubbard W, Jackel L D, et al. Handwritten digit recognition with a back-propagation network. In: *Proceedings of Advances in Neural Information Processing Systems*. San Francisco, CA, USA: Morgan Kaufmann Publishers Inc., 1990. 396–404
- 9 Goodfellow I J, Pouget-Abadie J, Mirza M, Xu B, Warde-Farley D, Ozair S, et al. Generative adversarial nets. In: *Proceedings of the 27th International Conference on Neural Information Processing Systems*. Montreal, Canada: MIT Press, 2014. 2672–2680

- 10 Radford A, Metz L, Chintala S. Unsupervised representation learning with deep convolutional generative adversarial networks. In: Proceedings of the 4th International Conference on Learning Representations. Caribe Hilton, San Juan, Puerto Rico, 2016. 97–108
- 11 Xue T F, Wu J J, Bouman K L, Freeman W T. Visual dynamics: probabilistic future frame synthesis via cross convolutional networks. In: Proceedings of Advances in Neural Information Processing Systems. Barcelona, Spain: Curran Associates, Inc., 2016. 91–99
- 12 Denton E L, Chintala S, Szlam A, Fergus R. Deep generative image models using a Laplacian pyramid of adversarial networks. In: Proceedings of the 28th International Conference on Neural Information Processing Systems. Montreal, Canada: MIT Press, 2015. 1486–1494
- 13 Krizhevsky A, Sutskever I, Hinton G E. ImageNet classification with deep convolutional neural networks. In: Proceedings of Advances in Neural Information Processing Systems. Lake Tahoe, Nevada, USA: Curran Associates, Inc., 2012. 1097–1105
- 14 Liu M Y, Tuzel O. Coupled generative adversarial networks. In: Proceedings of Advances in Neural Information Processing Systems. Barcelona, Spain: Curran Associates, Inc., 2016. 469–477
- 15 Mirza M, Osindero S. Conditional generative adversarial nets. arXiv: 1411.1784, 2014
- 16 van den Oord A, Kalchbrenner N, Espeholt L, Kavukcuoglu K, Vinyals O, Graves A. Conditional image generation with PixelCNN decoders. In: Proceedings of Advances in Neural Information Processing Systems. Barcelona, Spain: Curran Associates, Inc., 2016. 4790–4798
- 17 Reed S, Akata Z, Mohan S, Tenka S, Schiele B, Lee H. Learning what and where to draw. In: Proceedings of Advances in Neural Information Processing Systems. Barcelona, Spain: Curran Associates, Inc., 2016. 217–225
- 18 Salimans T, Goodfellow I, Zaremba W, Cheung V, Radford A, Chen X, et al. Improved techniques for training GANs. In: Proceedings of Advances in Neural Information Processing Systems. Barcelona, Spain: Curran Associates, Inc., 2016. 2226–2234
- 19 Chen X, Chen X, Duan Y, Houthoof R, Schulman J, Sutskever I, et al. InfoGAN: interpretable representation learning by information maximizing generative adversarial nets. In: Proceedings of Advances in Neural Information Processing Systems. Barcelona, Spain: Curran Associates, Inc., 2016. 2172–2180
- 20 Odena A, Olah C, Shlens J. Conditional image synthesis with auxiliary classifier GANs. In: Proceedings of the 34th International Conference on Machine Learning. Sydney, Australia, 2017. 2642–2651
- 21 Ghosh A, Kulharia V, Nambodiri V, Torr P H S, Dokania P K. Multi-agent diverse generative adversarial networks. arXiv: 1704.02906, 2017
- 22 Arjovsky M, Bottou L. Towards principled methods for training generative adversarial networks. arXiv:1701.04862, 2017
- 23 LeCun Y, Cortes C, Burges C J C. The MNIST database of handwritten digits [Online], available: <http://yann.lecun.com/exdb/mnist>, June 3, 2017.
- 24 LeCun Y, Bottou L, Bengio Y, Haffner P. Gradient-based learning applied to document recognition. *Proceedings of the IEEE*, 1998, **86**(11): 2278–2324
- 25 Liu Z W, Luo P, Wang X G, Tang X O. Large-scale CelebFaces Attributes (CelebA) Dataset [Online], available: <http://mmlab.ie.cuhk.edu.hk/projects/CelebA.html>, July 20, 2017.
- 26 Liu Z W, Luo P, Wang X G, Tang X O. Deep learning face attributes in the wild. In: Proceedings of the 2015 IEEE International Conference on Computer Vision. Santiago, Chile: IEEE, 2015. 3730–3738
- 27 Princeton ModelNet [Online], available: <http://modelnet.cs.princeton.edu>, August 13, 2017.
- 28 Wu Z R, Song S R, Khosla A, Yu F, Zhang L G, Tang X O, et al. 3D ShapeNets: a deep representation for volumetric shapes. In: Proceedings of the 2015 IEEE Conference on Computer Vision and Pattern Recognition. Boston, USA: IEEE, 2015. 1912–1920
- 29 Wu J J, Zhang C K, Xue T F, Freeman B, Tenenbaum J. Learning a probabilistic latent space of object shapes via 3D generative-adversarial modeling. In: Proceedings of Advances in Neural Information Processing Systems. Barcelona, Spain: Curran Associates, Inc., 2016. 82–90



张 龙 宁波大学博士研究生. 2008 年获得瑞典布京理工学院硕士学位. 主要研究方向为神经网络与深度学习.

E-mail: 1401082013@nbu.edu.cn

(**ZHANG Long** Ph.D. candidate at Ningbo University. He received his master degree from Blekinge Institute of Technology, Sweden in 2008. His research interest covers neural networks and deep learning.)



赵杰煜 宁波大学教授. 主要研究方向为计算机图像处理, 机器学习, 神经网络. 本文通信作者.

E-mail: zhao_jieyu@nbu.edu.cn

(**ZHAO Jie-Yu** Professor at Ningbo University. His research interest covers computer image processing, machine learning, and neural networks. Corresponding author of this paper.)



叶绪伦 宁波大学博士研究生. 2016 年获得宁波大学硕士学位. 主要研究方向为非参聚类, 流形学习以及非负矩阵分解. E-mail: 1601082017@nbu.edu.cn

(**YE Xu-Lun** Ph.D. candidate at Ningbo University. He received his master degree from Ningbo University in 2016. His research interest covers

non-parametric clustering, manifold learning, and non-negative matrix factorization.)



董 伟 宁波大学硕士研究生. 2015 年获得辽宁科技大学学士学位. 主要研究方向为神经网络, 深度学习.

E-mail: 1511082629@nbu.edu.cn

(**DONG Wei** Master student at Ningbo University. He received his bachelor degree from University of Science and Technology Liaoning in 2015.

His research interest covers neural networks and deep learning.)