

GAN 生成对抗网络在简单图像生成中的初步探索

黄云磊

浙江清桦智控科技有限公司 浙江杭州

【摘要】生成对抗网络（GAN）在图像生成领域展现强大潜力。探索其在简单图像生成的应用，分析网络结构与训练机制对生成效果的影响。生成器与判别器相互博弈的过程，通过对抗训练优化参数，实现对简单图像数据分布的学习与拟合。研究不同激活函数、损失函数组合下的生成性能，采用多层感知机与卷积神经网络构建基础模型，对比实验结果表明，卷积结构结合合适损失函数能显著提升图像生成质量与多样性。详细剖析训练过程中的模式崩溃、梯度消失等问题及应对策略，为后续复杂图像生成研究提供理论与实践参考，同时拓展 GAN 在计算机视觉基础任务中的应用边界。

【关键词】生成对抗网络；简单图像生成；对抗训练；图像质量；模型优化

【收稿日期】2025 年 5 月 14 日

【出刊日期】2025 年 6 月 5 日

【DOI】10.12208/j.jer.20250258

A preliminary exploration of generative adversarial networks (GANs) in simple image generation

Yunlei Huang

Zhejiang Qinghua Intelligent Control Technology Co., Ltd, Hangzhou, Zhejiang

【Abstract】Generative Adversarial Networks (GANs) have shown significant potential in the field of image generation. This study explores the application of GANs in simple image generation, analyzing how network structure and training mechanisms influence the generation outcomes. The process involves a mutual game between the generator and discriminator, where parameters are optimized through adversarial training to learn and fit the distribution of simple image data. The study examines the generation performance under different activation functions and loss function combinations, using multi-layer perceptrons and convolutional neural networks as the basic models. Comparative experiments show that combining a convolutional structure with an appropriate loss function can significantly enhance the quality and diversity of image generation. The study also delves into issues such as mode collapse and gradient disappearance during training, along with their corresponding strategies, providing theoretical and practical references for future research on complex image generation. Additionally, it expands the application boundaries of GANs in basic computer vision tasks.

【Keywords】Generative adversarial networks; Simple image generation; Adversarial training; Image quality; Model optimization

引言

随着计算机视觉技术发展，图像生成成为重要研究方向。生成对抗网络以独特的对抗训练机制，为图像生成提供全新思路。在简单图像生成场景中，其能快速学习图像特征分布，高效生成具有一定真实感的图像样本。网络训练稳定性、生成图像质量及多样性等问题亟待解决。探索 GAN 在简单图像生成中的应用，剖析其技术原理与优化策略，对推动图像生成技术发展、拓展其在图像编辑、虚拟数据生成等领域应用具有重要意义，也为更复杂图像生成研究奠定基础。

1 问题剖析

生成对抗网络（GAN）在涉足简单图像生成领域时，犹如踏入布满荆棘的丛林，面临着重重挑战。在网络训练的核心环节，生成器与判别器之间的动态平衡恰似走钢丝般微妙而艰难。判别器作为图像真伪的“守门人”，若其能力过强，生成器就如同被禁锢在牢笼中的探索者，难以接触到真实数据分布的“养分”。它输出的图像会陷入低质量的困境，像素模糊、结构扭曲，无法捕捉到真实图像的神韵与细节。反之，当生成器的能力过度膨胀，判别器则会迅速失去存在意义，就像失去猎物的猎手。此时，生成器生成的图像会陷入模式崩溃的泥沼，不断重复生成少数几种相似的图像，丧

失了图像的多样性，无法反映真实世界的丰富性。

除了生成器与判别器的博弈困境，梯度消失与梯度爆炸问题如同隐藏在暗处的幽灵，时刻威胁着网络训练的稳定性。在深度神经网络的训练过程中，梯度是引导网络参数更新的关键信号。当梯度消失发生时，网络参数的更新变得极为缓慢，甚至停滞不前，就像一辆失去动力的汽车，难以抵达理想的训练终点。而梯度爆炸则如同脱缰的野马，使得网络参数在更新过程中剧烈波动，数值不断增大，最终导致网络无法收敛，训练过程陷入混乱^[1]。这两种情况都会让生成器难以精准地学习到真实数据分布的规律，生成的图像自然难以符合预期。

图像世界的多样性决定了没有一套通用的网络结构与参数配置能够“包打天下”。不同类型的简单图像，无论是 MNIST 数据集中简洁的手写数字，还是 CIFAR-10 数据集中色彩丰富的小图像，都有着独特的特征与分布规律。设计网络结构时，需要考虑如何合理搭建网络层数、选择合适的激活函数、确定卷积核的大小等^[2]。参数配置方面，学习率的大小直接影响网络训练的速度与稳定性，batchsize 的选择关乎网络对数据的学习效率。若不能根据图像类型进行针对性的设计与调整，网络在生成图像时就如同“盲人摸象”，难以生成高质量、符合需求的图像。

2 解决策略

为突破上述困境，我们从网络结构设计与训练方法优化这两大关键维度出发，探索解决之道。在网络结构的革新上，卷积神经网络（CNN）凭借其独特的优势，成为替代传统多层感知机的不二之选。CNN 中的卷积层就像一个个“视觉神经元”，能够敏锐地捕捉图像的局部特征。无论是手写数字的笔画走势，还是 CIFAR-10 图像中物体的轮廓细节，卷积层都能将其转化为抽象的特征表示。多层卷积与池化操作的组合，宛如搭建起一座通往图像本质的阶梯。卷积层不断提取更高级的特征，池化层则对特征进行筛选与降维，让网络能够在保留关键信息的高效地捕捉图像的空间结构信息，为生成高质量图像奠定基础。

批归一化层的引入，堪称网络训练中的“稳定器”。在网络训练过程中，每一层的输入数据分布会随着参数的更新而发生变化，这种现象被称为“内部协变量转移”。它会使网络训练变得困难，增加训练时间，甚至导致网络难以收敛^[3]。批归一化层通过对每层输入进行标准化处理，将数据的均值和方差控制在一定范围内，有效地缓解了内部协变量转移问题。这就如同为

网络训练铺平了道路，让梯度能够更加顺畅地传递，极大地缓解了梯度消失问题，使得网络能够以更快的速度收敛到最优解，提高训练效率与稳定性。

在训练方法的优化上，损失函数的改进是核心突破口。传统 GAN 的损失函数在衡量生成图像与真实图像的差异时，存在梯度不稳定的缺陷，容易引发模式崩溃等问题。WassersteinGAN (WGAN) 的损失计算方式则另辟蹊径，引入 Wasserstein 距离，它就像一把更精准的“尺子”，能够更合理地度量两个概率分布之间的差异^[4]。这种改进使得损失函数的梯度更加稳定，避免了梯度消失和模式崩溃的风险，让生成器在训练过程中能够更准确地学习真实数据分布。调整生成器与判别器的训练次数比例也是关键策略。通过合理分配两者的训练机会，就像调节天平的砝码，让生成器与判别器在对抗过程中保持相对均衡的态势，形成良性竞争，共同推动图像生成质量的提升。

3 实验验证

为检验改进策略的有效性，我们精心搭建了实验环境，选择 MNIST 手写数字数据集与 CIFAR-10 简单图像数据集作为实验对象。MNIST 数据集以其简洁清晰的手写数字图像，成为验证图像生成算法基础性能的经典选择；而 CIFAR-10 数据集涵盖多种类别的小图像，色彩丰富、细节多样，能够更全面地考验算法在复杂场景下的表现。在 MNIST 数据集的实验中，我们分别运用传统 GAN、WGAN 以及经过改进后的网络模型展开图像生成训练。训练过程中，三种模型如同三位不同风格的“画家”，在数字图像的画布上挥洒笔墨。

在对生成图像质量的评估环节，我们采用峰值信噪比（PSNR）、结构相似性指数（SSIM）等指标进行量化分析。PSNR 从像素层面衡量图像失真程度，SSIM 则从结构和纹理等方面评估图像的相似性。传统 GAN 生成的手写数字图像，如同被蒙上一层薄纱，数字边缘模糊不清，细节部分也因失真而难以辨认；WGAN 在一定程度上有所改善，但仍存在数字形态不够自然流畅的问题^[5]。而改进后的网络模型，凭借优化后的网络结构与训练方法，生成的手写数字图像如同精雕细琢的艺术品，边缘清晰锐利，每一个笔画的走势都精准呈现，细节部分也生动逼真，仿佛能感受到书写者笔尖的力度。

在 CIFAR-10 数据集的实验中，改进策略的优势同样得到充分验证。面对色彩斑斓、形态各异的图像，传统模型生成的图像常常出现色彩偏差，物体形态扭曲变形，如同梦境中虚幻模糊的景象。WGAN 虽然在色

彩表现上有所进步，但在物体细节刻画上仍显不足^[6]。改进后的网络模型则展现出强大的实力，生成的图像在色彩还原度上近乎完美，物体的形态栩栩如生，无论是飞翔的小鸟羽毛的纹理，还是汽车的轮廓线条，都与真实图像相差无几。这充分证明，改进后的策略能够有效提升图像生成的质量与多样性，适应不同类型图像的生成需求。

4 效果评估

对改进后的 GAN 在简单图像生成任务中的效果，我们从多个维度进行了全面而深入的评估。首先，从视觉效果层面观察，生成的图像堪称一场视觉盛宴。以 MNIST 数据集为例，生成的手写数字不再是生硬、呆板的像素组合，而是具有了手写字体特有的灵动与韵味，每个数字都仿佛是真实书写留下的痕迹。在 CIFAR-10 数据集中，生成的图像色彩鲜艳且自然，物体的形态、位置与比例都与真实图像高度契合，无论是蓝天白云下的房屋，还是绿茵草地上的动物，都如同从现实世界中截取的画面，细节丰富，充满真实感，完全能够满足简单图像生成的基本需求。

在定量分析方面，我们通过多项指标与传统模型进行对比，结果令人振奋。在图像生成质量评估中，改进后的模型在 PSNR、SSIM 等指标上均大幅超越传统模型，这意味着其生成的图像在像素精度、结构相似度等关键维度表现更为出色。在多样性评估中，传统模型生成的图像常常陷入模式单一的困境，而改进后的模型生成的图像在内容、风格等方面呈现出丰富的变化，能够生成多样化的图像样本，真实地反映出数据集中图像的多样性特征^[7]。在训练效率的考量上，改进后的训练策略如同为网络注入了一剂“强心针”。网络收敛速度显著加快，原本需要漫长时间才能完成的训练过程，如今能够在更短的时间内达到理想的训练效果，大大提高了研究与应用的效率。

综合来看，改进后的 GAN 在简单图像生成任务中实现了质的飞跃。它不仅在图像生成质量与多样性上取得了显著提升，还大幅提高了训练效率。这一成果为后续复杂图像生成研究提供了宝贵的经验与可借鉴的方法^[8]。无论是优化网络结构的思路，还是改进训练方法的策略，都为进一步探索图像生成的奥秘奠定了坚实基础，激励着研究者们向更具挑战性的图像生成领域继续迈进。

5 结语

GAN 在简单图像生成中的探索取得一定成果，通过优化网络结构与训练方法，有效提升图像生成质量与多样性，解决了训练过程中的部分关键问题。未来，可进一步探索将 GAN 与其他深度学习模型融合，拓展在复杂场景图像生成中的应用。深入研究网络训练机制，优化对抗策略，以应对更具挑战性的图像生成任务。探索 GAN 在跨模态图像生成、图像修复等领域的应用，推动图像生成技术向更智能化、多样化方向发展。

参考文献

- [1] 魏志宏,孔旭东,孔燕,等.基于全局和局部特征信息的生成对抗网络在海马体分割中的应用[J].光谱学杂志,2025,42(02):143-153.
- [2] 许睿,贺玉洁,万航,等.基于时序生成对抗网络的水质特征数据生成与时变过程预测[J/OL].水生态学杂志,1-12[2025-06-23].
- [3] 傅鹏,宋晓霞,赵心宇.基于 Transformer-CNN 的生成对抗网络矿井图像去噪方法[J/OL].数学的实践与认识,1-12[2025-06-23]
- [4] 周建新,郭强,霍彤明,等.生成对抗网络与模糊 PID 的层流冷却温度控制研究[J/OL].热加工工艺,1-6[2025-06-23].
- [5] 吕朋蓬,卜强生,郭野,等.基于梯度惩罚生成对抗网络的配电网缺失数据修复方法[J].太阳能学报,2025,46(05):185-192.
- [6] 唐岚,黄力文,王成磊,等.基于联邦学习框架的改进生成对抗网络场景生成方法[J/OL].昆明理工大学学报(自然科学版),1-13[2025-06-23].
- [7] 郭亮,徐行,刘开勇,等.数字孪生环境下基于生成对抗网络的钻井液流变性能预测方法[J].钻井液与完井液,2025,42(03):359-367.
- [8] 李朝霞,吴宏涛.基于生成对抗网络的图像生成方法研究[J].信息与电脑,2025,37(10):1-5.

版权声明：©2025 作者与开放获取期刊研究中心(OAJRC)所有。本文章按照知识共享署名许可条款发表。

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

