

# AutoGAN: 神经架构搜索生成对抗网络

Xinyu Gong<sup>1</sup> Shiyu Chang<sup>2</sup> Yifan Jiang<sup>1</sup> Zhangyang Wang<sup>1</sup>

<sup>1</sup>Department of Computer Science & Engineering, Texas A&M University

<sup>2</sup>MIT-IBM Watson AI Lab

fxy gong, yifanjiang97, atlaswangg@tamu.edu

shiyu.chang@ibm.com

## 摘要

神经架构搜索 (NAS) 见证了图像分类和 (最近) 分割任务的成功。在本文中, 我们提出了第一个关于将 NAS 算法引入生成对抗网络 (GAN) 的初步研究, 称为 AutoGAN。NAS 和 GAN 的结合面临着独特的挑战。我们为生成器架构变量定义搜索空间, 并使用 RNN 控制器指导搜索, 通过参数共享和动态重置来加速搜索过程。采用初始分数作为奖励, 并引入多级搜索策略以渐进方式执行 NAS。实验验证了 AutoGAN 在无条件的图像生成任务上的有效性。具体而言, 与目前最先进的手工制作的 GAN 相比, 我们方法发现的架构实现了极具竞争力的性能, 例如, 分别在 CIFAR-10 上达到了 12.42 的最新 FID 分数和在 STL-10 上达到了 31.01 分。最后, 我们还讨论了当前局限和 AutoGAN 的未来潜力。该代码可通过 <https://github.com/TAMU-VITA/AutoGAN> 获得。

## 1. 介绍

生成对抗网络 (GAN) 自其起源以来一直在发展[11]。他们最引人注目的成功之一在于用各种卷积结构生成逼真的自然图像[47,5,61,26,13,67]。为了提高生成图像的质量, 已经作出了许多尝试, 包括修改判别损失函数[3,68], 强制执行正则化项[6,13,42,5], 引入注意力机制[61,64], 并采用渐进式训练[26]。

然而, GAN 的骨干架构设计受到的关注相对较少, 并且往往被认为是影响 GAN 性能的一个不太重要的因素[39,32]。大多数早期的 GAN 主要使用相对较浅的生成器和判别器架构

大多由于 GAN 训练中臭名昭著的不稳定性。最近, 一些最先进的 GAN [13,42,64,5]采用了深度残差网络骨干生成器, 以便更好地生成高分辨率图像。在大多数其他计算机视觉任务中, 很多进展都来自网络架构的改进设计, 例如图像分类[15,16,52,51,35], 分割[8,9,49]和姿势估计[43]。因此, 我们提倡增强骨干网络设计对于进一步改善 GAN 也很重要。

近年来, 人们对自动设计复杂的神经网络架构产生了浓厚的兴趣。神经架构搜索 (NAS) 已经成功地开发并评估了图像分类的任务[71,46], 最近还进行了图像分割[7,36], 并且发现的架构优于人工设计的模型。然而, 天真地将现有的 NAS 想法从图像分类/分割移植到 GAN 是不可能的。首先, 即使是手工设计的架构, GAN 的训练也是出了名的不稳定, 容易崩溃[50]。融合 NAS 进入训练过程无疑会增加难度。作为另一个重要挑战, 虽然验证准确性为图像分类中的 NAS 提供了奖励的一种自然选择, 但其实选择用于评估和指导 GAN 搜索过程的良好度量标准并不那么简单。

本文介绍了一种专为 GAN 量身定制的架构搜索方案, 称为 AutoGAN。据我们所知, AutoGAN 描述了首次将 NAS 与 GAN 结合使用的尝试, 并且属于首次尝试将 NAS 扩展到图像分类之外。我们的技术创新总结如下:

- 我们定义搜索空间以捕获 GAN 架构变化。最重要的是, 我们使用 RNN 控制器[71]来指导架构搜索。基于[46]中的参数共享策略, 我们在搜索过程中进一步引入了参数动态重置策略, 以提高训练速度。
- 我们使用 Inception score (IS) [50]作为奖励, 在基于强化学习的 AutoGAN 优化中。发现的模型也展示出

其他 GAN 指标下的有说服力表现, 例如 Frechet 初始距离 (FID) [50]。

- 我们进一步向 AutoGAN 引入了多级架构搜索 (MLAS), 这是由渐进式 GAN 训练[26]推动的。MLAS 使用波束搜索以自下而上的顺序方式在多个阶段执行搜索[37]。

我们进行了各种实验来验证 AutoGAN 的有效性。我们发现的架构产生了非常有前景的结果, 这些结果比目前手工设计的 GAN 更好或更有可比性。在 CIFAR-10 数据集上, AutoGAN 获得的初始分数为 8.55, FID 分数为 12.42。此外, 我们还表明 CIFAR-10 上的发现架构甚至在 STL-10 图像生成任务上具有竞争力, 其中 9.16 的初始得分和 31.01 FID 得分, 展示了强大的可迁移性。在这两个数据集上, AutoGAN 建立了新的最先进的 FID 分数。我们的许多实验发现都与之前的 GAN 制作经验相吻合, 并阐明了对 GAN 生成器设计的新见解。

## 2. 相关工作

### 2.1. 神经架构搜索

神经架构搜索 (NAS) 算法旨在找到最佳的神经网络架构, 而不是使用手工制作的特定模型。之前关于 NAS 的工作在图像分类任务上取得了巨大成功[30]。最近的工作进一步将 NAS 算法扩展到密集和结构化的预测[7,36]。值得一提的是, NAS 也适用于 CNN 的压缩和加速[17]。但是, 目前没有为生成模型开发任何 NAS 算法。

NAS 算法由三个关键组件组成: 搜索空间, 优化算法和代理任务。对于搜索空间, 通常有两个类别: 直接搜索整个架构 (宏搜索), 或搜索单元并以预定义的方式堆叠它们 (微搜索)。对于优化算法, 流行的选项包括强化学习 [4,71,69,72], 进化算法[60], 贝叶斯优化[24], 随机搜索[7]和基于梯度的优化方法[38,1]。对于代理任务, 它被设计用于在训练期间有效地评估发现的架构的性能。例子包括早期停止[7,72], 使用低分辨率图像[57,10,29], 使用替代模型进行性能预测[37], 采用小骨干[7]或利用共享参数[46]等。

大多数 NAS 算法[71,46]通过一次通过控制器的生成网络架构 (宏搜索) 或单元 (微搜索)。最近的一项工作[37]使用波束搜索在图像分类任务中引入了 NAS 的多级搜索。架构搜索将从较小单元开始

并且保留了顶级表现的候选架构。对于更大的单元, 下一轮搜索将继续基于它们。

### 2.2. 生成对抗网络

GAN 有一个生成器网络和一个判别器网络, 可以互相玩最小-最大双人游戏。它在许多生成和合成任务中取得了巨大的成功, 例如文本到图像的翻译[66,65,61,48], 图像到图像的翻译[22,70,63]和图像增强[33,31,23]。然而, GAN 的训练经常被发现非常不稳定[50], 并且通常遭受非收敛、模式崩溃和对超参数的敏感等问题。许多努力都致力于缓解这些问题, 例如 Wasserstein 损失[3], 谱归一化[42], 渐进式训练[26]和自注意力层[64], 仅举几例。

## 3. 技术方法

GAN 由两个竞争网络组成: 生成器和判别器。众所周知, 两个架构必须在学习能力方面进行微妙的平衡。因此, 要构建 AutoGAN, 第一个问题是: 如何在 GAN (生成器和判别器, 下文表示为 G 和 D) 中一起构建两个网络? 一方面, 如果我们使用预先固定的 D (或 G) 并仅搜索 G (或 D), 则很容易引起 D 或 G [18,2]之间的不平衡 (特别是在 NAS 的早期阶段), 导致更新缓慢或碎片式的学习。另一方面, 尽管可能联合搜索 G 和 D, 但经验及实验观察到这种双向 NAS 将进一步恶化原始的 GAN 训练的不稳定性, 导致高度振荡的训练曲线并且通常会无法收敛。作为权衡, 我们建议使用 NAS 仅搜索 G 的体系结构, 同时通过遵循给定的例程来堆叠预定义的块来使 D 变得更深。增长 D 的细节将在补充中更详细解释。

基于此, AutoGAN 遵循[71]的基本思想, 使用循环神经网络 (RNN) 控制器从其搜索空间中选择块, 以构建 G 网络。基本方案如图 1 所示。我们进行多重创新, 以解决训练 GAN 的具体任务所带来的独特挑战。接下来我们从三个关键方面介绍 AutoGAN: 搜索空间, 代理任务和优化算法。

### 3.1. 搜索空间

AutoGAN 基于多级架构搜索策略, 其中生成器由多个单元组成。这里我们使用  $(s+5)$  元素元组  $(skip_1, \dots, skip_s, C, N, U, SC)$  来对第  $s$  个单元格进行分类,

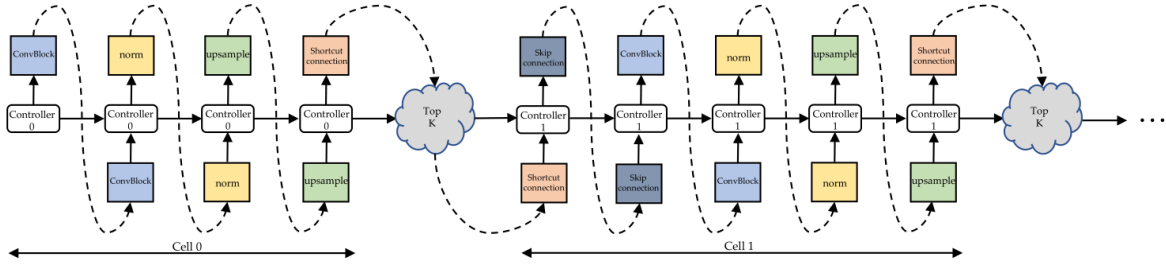


图 1: RNN 控制器的运行方案。在每个时间步骤，控制器将要解码的隐藏矢量输出到具有其对应的 softmax 分类器的操作中。请注意，使用 MLAS，我们使用不同的控制器来搜索每个单元格。一旦完成一个单元的搜索过程，控制器就会对  $M$  个候选架构进行采样，然后从中挑选出前  $K$  个。前  $K$  个架构的控制器输出将作为下一个单元控制器的输入。

其中  $s$  是从 0 开始的单元格索引（第 0 个单元格没有  $skip_0$  连接）：

- $skip_i$  是二进制值，指示当前第  $s$  个单元是否从第  $(i-1)$  个单元获取跳跃连接作为其输入， $i = 1, \dots, s$ 。请注意，每个单元可以从其他前面的单元格中获取多个跳跃连接。
- $C$  是基本卷积块类型，包括预激活[16]和激活后卷积块。
- $N$  代表该块的标准类型，包括三个选项：批量标准化[21]，实例标准化[55]，无标准化。
- $U$  代表上采样操作，其在当前图像生成 GAN 中是标准的，包括双线性上采样，最近邻上采样和步幅-2 反卷积。
- $SC$  是指示单元内快捷方式的二进制值。

图 2 说明了 AutoGAN 生成器搜索空间。上采样操作  $U$  还将确定跳入特征映射的上采样方法。

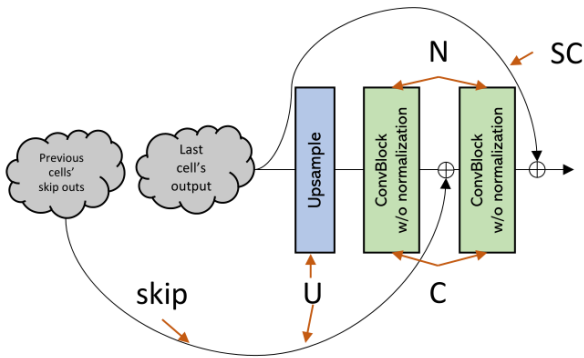


图 2: AutoGAN 中生成器单元的搜索空间。

### 3.2 代理任务

初始得分 (IS) [50] 和 FID 得分[18]是 GAN 的两个主要评估指标。由于 FID 得分的计算时间要长得多，我们选择每个派生子模型的 IS 作为通过强化学习更新控制器的奖励。

参数共享[40,73]显示了有效提高 NAS 的效率[46]。基于[46]中的参数共享，我们进一步为 AutoGAN 引入了一个参数动态重置策略。据观察，GAN 的训练变得不稳定，经过长时间的训练后会出现模式崩溃[5]。继续训练共享崩溃的模型可能是浪费时间。根据经验，我们观察到当模式崩溃发生时，训练损失（铰链损失）的方差通常变得非常小。

基于这一观察，我们设置了一个移动窗口来存储生成器和判别器的最新训练损失值。一旦那些存储的训练损失的标准偏差小于预定阈值，就终止当前迭代的共享 GAN 的训练。在当前迭代更新控制器之后，将重新初始化共享 GAN 模型的参数。请注意，我们不会重新初始化 RNN 控制器的参数，因此它可以继续引导架构搜索并继承历史知识。通过动态重置，搜索过程变得更加高效。

### 3.3 优化算法

AutoGAN 中有两组参数：RNN 控制器的参数（表示为  $\theta$ ）；搜索到的生成器和相应的判别器中的共享 GAN 参数（表示为  $\omega$ ）。算法 1 简要概述了训练过程，作为两阶段之间的交替过程。

第一阶段将训练共享 GAN 的  $\omega$  一段时期， $\theta$  固定。对于每次训练迭代

在共享 GAN 的训练过程中, 候选架构将由 RNN 控制器进行采样。一旦记录的训练损失的标准偏差降至阈值以下, 动态重置标识  $F_{DR}$  将被设置为 True, 并且共享 GAN 的训练过程将立即终止。注意, 在当前时期的控制器训练完成之前, 不会重新初始化共享 GAN。第二阶段训练  $\theta$  并且  $\omega$  固定: 控制器将首先对共享生成器的  $K$  个子模型进行采样。他们的 IS 将被计算为奖励。RNN 控制器将通过有移动平均基线的强化学习方法进行更新。在训练了  $u_{stage}$  迭代次数之后, 将从已有的架构中挑选出排前的  $K$  个架构。同时, 新的控制器将被初始化以进行下一阶段的架构搜索。

---

**Algorithm 1:** Pseudo codes for AutoGAN searching.
 

---

```

iters = 0;
stage = 0;
 $F_{DR} = False$ ;
while iters < 90 do
    train (generator, discriminator,  $F_{DR}$ );
    train (controller);
    if iters %  $u_{stage} == 0$  then
        save the top  $K$  architectures;
        generator = grow (generator);
        discriminator = grow (discriminator);
        controller = new (controller);
        stage += 1;
    end
    if  $F_{DR} == True$  then
        // dynamic reset
        initialize (generator);
        initialize (discriminator);
         $F_{DR} = False$ ;
    end
    iters += 1;
end
  
```

---

### 3.3.1 训练共享 GAN

在此阶段, 我们修正 RNN 控制器的策略  $\pi(a, \theta)$  并通过标准的 GAN 训练更新共享参数  $\omega$ 。具体来说, 我们使用铰链对抗性损失以交替方式训练[42,5,53,64]:

$$\mathcal{L}_D = E_{x \sim q_{data}} [\min(0, -1 + D(x))] + E_{z \sim p(z)} [\min(0, -1 - D(G(z)))] \quad (1)$$

$$\mathcal{L}_G = E_{z \sim p(z)} [\min(0, D(G(z)))] \quad (2)$$

我们进一步向 AutoGAN 引入了多级架构搜索 (MLAS), 其中生成器 (以及相应的判别器) 将逐渐增长。

MLAS 使用波束搜索[37]以自下而上的方式进行单元搜索。当搜索下一个单元时, 我们将使用不同的控制器, 从当前候选单元中选择前  $K$  个波束并基于它们开始下一轮搜索。

### 3.3.2 训练控制器

在这个阶段我们修复  $\omega$  并更新控制器的策略参数。我们将奖励  $\mathcal{R}(a, \omega)$  定义为采样子模型  $a$  的 IS。RNN 控制器使用 Adam 优化器[28]通过 REINFORCE [59]更新, 具有移动平均基线。此外, 我们还添加了一个熵项来鼓励探索。

我们使用 LSTM [20]控制器。对于每个时间步长, LSTM 将输出隐藏矢量, 该矢量将被其相应的 softmax 分类器解码和分类。LSTM 控制器以自动渐进的方式工作, 最后一步的输出将被输入下一步, 并将每个时间步的输出中对 GAN 的每个单元的操作进行采样。具体而言, 当将新单元添加到现有模型以增加输出图像分辨率时, 将初始化新控制器。将保存先前的排前的  $K$  模型的架构和相应的隐藏向量。它们隐藏的矢量将作为输入馈入新控制器, 以搜索下一个单元的操作。

### 3.3.3 架构推导

我们将首先从学习策略  $\pi(a, \theta)$  中抽取几个生成器架构。然后, 将为每个模型计算奖励  $R$  (初始得分)。然后, 我们将根据最高奖励挑选 top-K 模型, 并从头开始训练。之后, 我们再次评估他们的初始分数, 并且具有最高初始分数的模型成为我们最终的派生生成器架构。

## 4. 实验

**数据集。**在本文中, 我们采用 CIFAR-10 [29]作为 AutoGAN 的主要测试平台。它由 50,000 个训练图像和 10,000 个测试图像组成, 每个图像的分辨率为  $32 \times 32$ 。我们使用该训练集训练 AutoGAN, 无需任何数据扩充。

我们还采用 STL-10 数据集来显示 AutoGAN 发现的架构的迁移能力。当使用 STL-10 进行训练时, 我们采用 5,000 图像训练集和 100,000 图像未标记集。所有图像都调整为  $48 \times 48$ , 没有任何其他数据增加。

**训练细节。**在训练共享 GAN 时, 我们遵循谱归一化 GAN [42]的训练设置。生成器和判别器的学习率设置为  $2e-4$ , 使用铰链损耗, Adam 优化器[28],



判别器的批次大小为 64, 生成器的批次大小为 128。谱归一化仅在判别器上强制执行。我们使用 Adam [28] 训练我们的控制器, 学习率为  $3.5e-4$ 。我们还将控制器输出概率的熵添加到奖励中, 加权  $1e-4$ , 以奖励探索。AutoGAN 搜索 90 个迭代周期。对于每次迭代, 共享 GAN 将被训练 15 个时期, 并且控制器将被训练 30 个时期。我们将动态重置方差阈值设置为  $1e-3$ 。我们使用与共享 GAN 相同的训练设置训练发现的体系结构, 并进行 50,000 次生成器迭代。

#### 4.1. CIFAR 10 上的结果

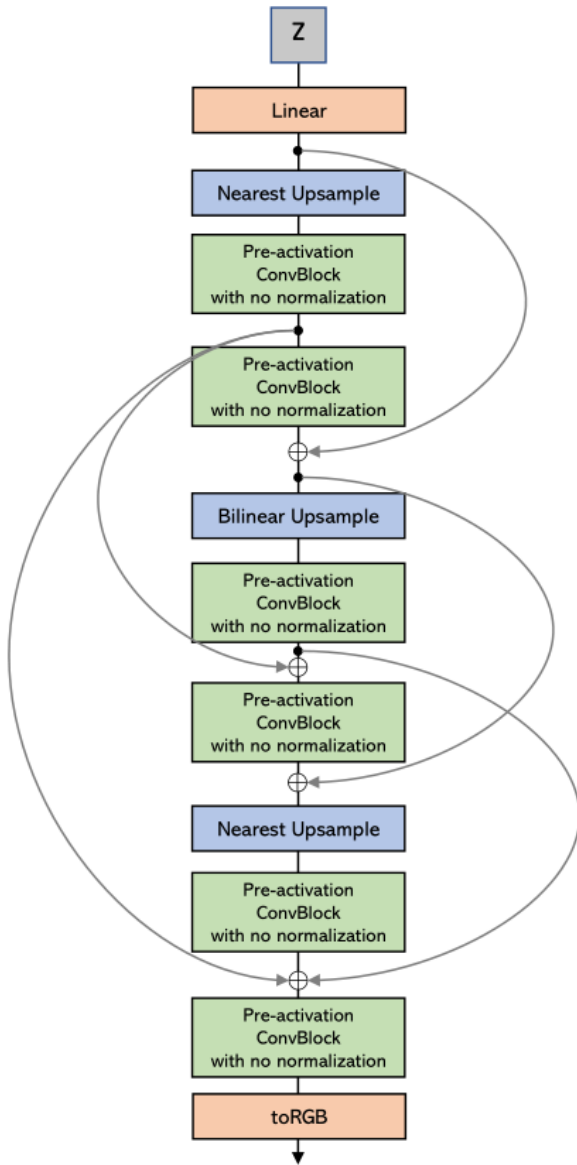


图 3: AutoGAN 在 CIFAR-10 上发现的 AutoGAN (生成器) 架构。

AutoGAN 在 CIFAR-10 训练集上发现的生成器体系结构如图 3 所示。对于无条件 CIFAR-10 图像生成 (没有使用类标签) 的任务, 可以总结一些值得注意的观察结果:

- 发现的架构有 3 个卷积块。与激活后的卷积块相比, AutoGAN 显然更喜欢预激活卷积块。
- AutoGAN 非常喜欢最近邻上采样 (也是双线性上采样) 去反卷积。这似乎与之前的经验[45]一致, 即反卷积可能会产生棋盘效应, 而最近邻采样可能会更受青睐。
- 有趣的是, AutoGAN 似乎不喜欢使用任何标准化。
- AutoGAN 显然赞成 (更密集的) 跳跃连接。特别是, 它能够发现中长跳跃 (跳跃 2-4 个卷积层), 并实现多尺度特征融合。

我们将 AutoGAN 与 CIFAR-10 数据集上手工制作的 GAN 最近发布的结果进行比较, 如表 2 所示。所有结果都是从那些原始论文中收集的: 因此, 其他的结果是在他们手动调整的最佳训练设置下获得的。在初始得分方面, AutoGAN 略微接近 Progressive GAN [26], 并且超越许多最新的强势竞争对手, 如 SNGAN [42], 提升的 MMD-GAN [56], Dist-GAN [54], MGAN [19], 和 WGAN-GP [13]。就 FID 而言, AutoGAN 优于所有当前最先进的模型。CIFAR-10 生成结果的可视化示例如图 4 所示。

请注意, AutoGAN 的当前搜索空间只能覆盖 SN-GAN, 而不能覆盖其他搜索空间。例如, WGAN-GP 使用 Wasserstein 损失 (具有梯度限幅), 并且 MGAN 采用多判别器结构 (而我们的判别器非常简单且未被搜索)。通过使用相同的构建块组, AutoGAN 能够胜过手工制作的 SN-GAN: 公平的比较方式为生成器结构的重要性和搜索算法的有效性提供了直接证据。

虽然尚未配备任何明确的模型复杂正则化, 但 AutoGAN 展示了 NAS 的预期模型参数效率。图 3 中最顶尖的架构具有 1.77G FLOP, 而其性能明显优于 SN-GAN (1.69G FLOP), 并且与 Progressive GAN (6.39G FLOP) 相当。

我们还会在表 2 底部报告候选前 2 名和前 3 名发现的架构的初始分数和 FID 分数。他们相应的架构如图 5 所示。我们可以看到我们所有的前三个架构都与当前最先进的模型实现了有竞争力的性能。

表 1: CIFAR-10 上无条件图像生成任务的初始分数和 FID 分数。

我们获得了最新的 FID 分数 12.42。

Method	Inception score	FID
DCGAN [47]	$6.64 \pm .14$	-
Improved GAN [50]	$6.86 \pm .06$	-
LRGAN [62]	$7.17 \pm .17$	-
DFM [58]	$7.72 \pm .13$	-
ProbGAN [14]	7.75	24.60
WGAN-GP, ResNet [13]	$7.86 \pm .07$	-
Splitting GAN [12]	$7.90 \pm .09$	-
SN-GAN [42]	$8.22 \pm .05$	$21.7 \pm .01$
MGAN [19]	$8.33 \pm .10$	26.7
Dist-GAN [54]	-	$17.61 \pm .30$
Progressive GAN [26]	<b><math>8.80 \pm .05</math></b>	-
Improving MMD GAN [56]	8.29	16.21
AutoGAN-top1 (Ours)	$8.55 \pm .10$	<b>12.42</b>
AutoGAN-top2	$8.42 \pm .07$	13.67
AutoGAN-top3	$8.41 \pm .11$	13.87

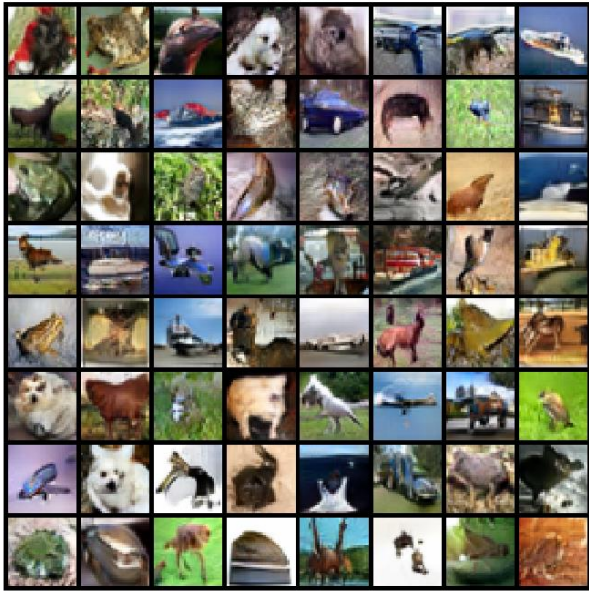


图 4: 生成的 AutoGAN 的 CIFAR-10 结果。它们是随机采样而不是采摘挑选的结果。

## 4.2 STL-10 的可转移性

我们试图回答的下一个问题是：发现的架构是否可能过度拟合数据集？换句话说，当我们使用另一个数据集重新训练其权重（架构结构固定）时，相同的架构是否会依然有效？为了解决这个问题，我们在 CIFAR-10 上采用 AutoGAN 发现的架构，并在 STL-10 训练和无标签数据集上重新训练其权重，用于 STL-10 上的无条件图像生成任务。请注意，STL-10 的图像分辨率高于 CIFAR-10。



图 5: 前 2 个和前 3 个发现的体系结构。

我们将最近公布的结果与 STL-10 数据集中手工制作的 GAN 进行比较，结果如表 2 所示。我们的结果令人惊讶地展示出：尽管 AutoGAN 在初始得分方面略微落后于提升的 MMD-GAN，后者实现了 31.01 的最佳 FID 结果。STL-10 生成结果的可视化示例如图 4 所示。如果我们从头开始重新执行架构搜索，我们希望在 STL-10 上获得更高的结果，并将其留待将来工作。

表 2: STL-10 上无条件图像生成的初始分数和 FID 分数。AutoGAN 在 CIFAR-10 上使用探索架构。

Method	Inception score	FID
D2GAN [44]	7.98	-
DFM [58]	$8.51 \pm .13$	-
ProbGAN [14]	$8.87 \pm .095$	46.74
SN-GAN [42]	$9.10 \pm .04$	$40.1 \pm .50$
Dist-GAN [54]	-	36.19
Improving MMD GAN [56]	<b>9.34</b>	37.63
AutoGAN (Ours)	$9.16 \pm .12$	<b>31.01</b>

## 4.3. 消融研究与分析

### 4.3.1 验证代理任务

我们方法的代理任务是直接评估共享 GAN 的子模型上的 IS，这大大提高了训练速度。验证有效性在

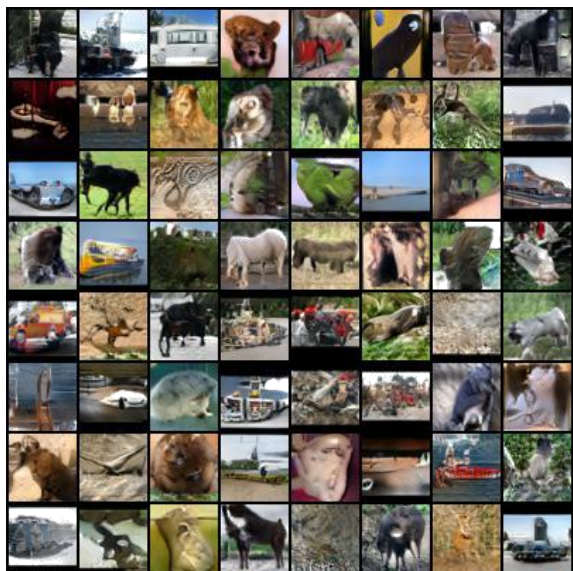


图 6: 生成的 AutoGAN 的 STL-10 结果。它们是随机抽样而不是采摘挑选的结果。

设计代理任务上的效果, 我们从头开始训练衍生架构 30 个时期并使用 CIFAR-10 上的 AutoGAN 评估它们的 IS。我们绘制了由代理任务 (Proxy) 提供的评估值与图 7 中的真实评估 (Real) 之间的相关性。我们可以观察到彼此之间的正相关, Spearman's 相关系数排在 0.779, 证明代理任务提供了真实评估的公平近似。我们还可以观察到代理评估往往低估了 Inception 分数, 这是由于我们的代理任务中有未完成和共享的训练导致的。

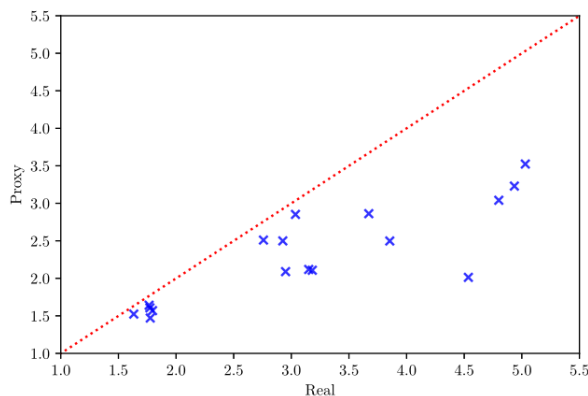


图 7: CIFAR-10 上实际评估和代理任务性能的相关图。

### 4.3.2 与使用 FID 作为奖励相比较

除了 IS 之外, GAN 的另一个重要评估指标是 FID 得分 [50]: 较小的 FID 表示更好的生成

质量。如果我们使用 FID 的倒数而不是 IS 作为 AutoGAN 控制器的奖励怎么样? 为了回答这个问题, 在 CIFAR-10 上, 我们从头开始搜索另一个模型, 使用 FID 的倒数作为奖励 (其他不变), 并与使用 IS 奖励训练的 AutoGAN 进行比较。在搜索过程中, 对于这两种模型, 我们会在训练过程中定期评估 IS 和 FID 值。图 8 描绘了曲线, 从中我们可以看到, 虽然两种奖励都可以随着搜索的进行而增加, 但是 FID 驱动搜索 (红色曲线) 显示了与 IS 驱动搜索 (蓝色曲线) 相比在两种指标下的性能。然而, 计算 FID 比计算 IS 更耗时, 因为它需要计算大矩阵的协方差和均值。因此, 我们选择 IS 作为 AutoGAN 的默认奖励。

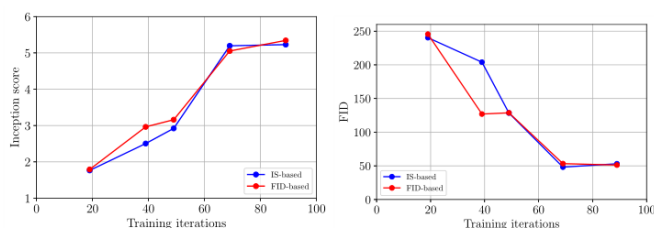


图 8: 使用 IS 和 FID (倒数) 作为 AutoGAN 奖励的比较。具有基于 IS 的奖励的模型以蓝色绘制, 而具有基于 FID 的奖励的模型以红色绘制。上图显示了两种模型的 IS 值随着训练的进行情况, 以及 FID 值的变化情况。

### 4.3.3 参数动态重置

众所周知, GAN 的训练过程非常不稳定, 经过长时间的训练后容易出现模式崩溃 [5]。继续训练共享折叠模型可能是浪费时间。因此, 我们引入了参数动态重置来缓解这个问题。为了进行比较, 我们在 CIFAR-10 上进行了两次 AutoGAN 实验, [46] 中提出了参数共享策略, 并提出了动态复位加参数共享策略。我们在两个训练过程中评估 IS。如图 9 所示, 我们可以看到它们实现了相当的性能。但是, 动态重置的训练过程只需要 43 个小时, 而没有动态重置则需要 72 个小时。

### 4.3.4 多级架构搜索

我们的 AutoGAN 框架默认采用多级架构搜索 (MLAS) 策略。为了比较, 我们使用单级架构搜索 (SLAS) 在 CIFAR-10 上进行另一个 AutoGAN 实验, 其中整个架构将通过单个控制器一次派生。我们评估 IS 并与下列进行比较



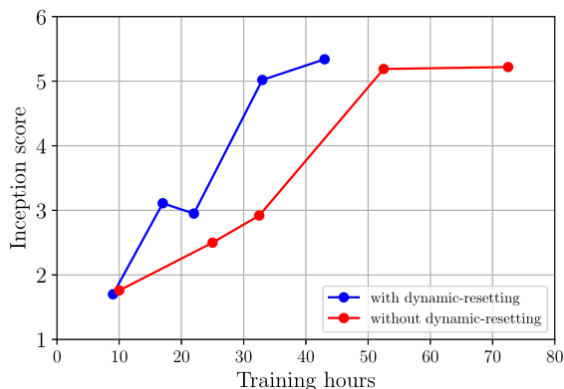


图 9: 将 AutoGAN 与使用 (蓝色) 和不使用 (红色) 动态重置进行比较。动态重置可以提高训练效率, 同时实现可比性能。

即 MLAS 的训练。我们可以看到 SLAS 在开始时获得了较高的初始分数, 但 MLAS 的初始分数逐渐增长, 最终优于 SLAS。此外, 训练 SLAS 比训练 MLAS 慢得多, 因为生成器输出图像将始终伴随最终分辨率。图 10 展示了 MLAS 的明显和一致的优点。

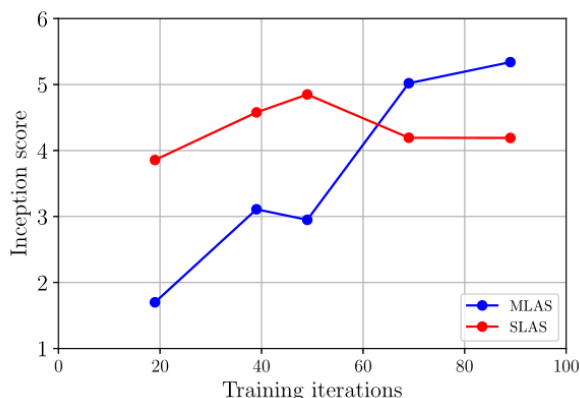


图 10: AutoGAN 与 MLAS (蓝线) 和 SLAS (红线) 训练过程的比较。

#### 4.3.5 与随机搜索相比较

我们在[34]中实现了两个随机搜索算法: 一个具有权重共享, 另一个没有权重共享 (早期停止)。我们在 CIFAR-10 上使用这两种算法重新搜索了 AutoGAN, 并将搜索时间限制为 48 小时。结果, 发现的具有权重共享的体系结构实现了  $IS = 8.09$  和  $FID = 17.34$ , 而另一个实现了  $IS = 7.97$  和  $FID = 21.39$ 。两者都不如我们提出的搜索算法, 并符合其效果。

## 5. 结论, 局限和讨论

AutoGAN 首次尝试将 NAS 引入 GAN。它能够在 CIFAR-10 和 STL-10 数据集上识别出高效的架构, 从而在当前最先进的手工制作的 GAN 模型中实现具有竞争力的图像生成结果。消融研究进一步揭示了每种成分的益处。

由于 GAN 训练本身的高度不稳定性 and 超高灵敏度, AutoGAN 似乎比 NAS 更具挑战性。回想一下, 在 AutoML 的初始阶段, 它只能设计与人类专家设计的神经网络相当的小型神经网络, 这些结果仅限于小型学术数据集, 如 CIFAR-10 和 Penn Treebank [71]。同样, 尽管取得了初步的成功和承诺, 但毫无疑问, AutoGAN 还有很大的发展空间。

为了使 AutoGAN 比最先进的手工设计的 GAN 更具竞争力, 我们指出了一些需要继续努力的具体项目:

- AutoGAN 的当前搜索空间有限, 并且一些强大的 GAN 被排除在可搜索范围之外。需要通过更多的构建块来扩大它, 这在 GAN 文献中是有效的。参考最近的 GAN 基准研究[39,32], 我们考虑扩大我们的搜索空间, 注意力/自注意力[64], 基于风格的生成器[27], 相对判别器[25]和各种损失, 如 Wasserstein 损失[3], 等等。
- 到目前为止, 我们尚未对更高分辨率的图像合成测试 AutoGAN, 例如 ImageNet。虽然相同的算法原则上可直接应用, 但计算成本会变得非常高。例如, 对 CIFAR-10 的搜索已经需要 43 个小时。关键的挑战在于如何进一步提高搜索算法的效率。
- 作为一种可能的策略, NAS 在图像分类中通常将迁移学习从低分辨率图像应用到更高分辨率[72]。看看类似的想法如何应用于 GAN 的图像生成任务可能是有趣的, 尽管具有挑战性, 因为它在保存/合成细节方面比分类要求更高。
- 我们没有释放寻找更好的判别器的潜力。我们可能会在生成器和判别器之间进行交替搜索, 这可能会使 AutoGAN 变得更具挑战性。
- 最终, AutoGAN 将需要能够联合内部标签, 例如条件 GAN (CGAN)[41]和半监督 GAN [50]。



## 参考文献

- [1] Karim Ahmed and Lorenzo Torresani. MaskConnect - Connectivity Learning by Gradient Descent. ECCV, cs.CV, 2018.
- [2] Martin Arjovsky and Leon Bottou. Towards principled methods for training generative adversarial networks. arXiv preprint arXiv:1701.04862, 2017.
- [3] Martin Arjovsky, Soumith Chintala, and Leon Bottou. Wasserstein generative adversarial networks. In International Conference on Machine Learning, pages 214–223, 2017.
- [4] Bowen Baker, Otthrist Gupta, Nikhil Naik, and Ramesh Raskar. Designing neural network architectures using reinforcement learning. arXiv preprint arXiv:1611.02167, 2016.
- [5] Andrew Brock, Jeff Donahue, and Karen Simonyan. Large scale gan training for high fidelity natural image synthesis. arXiv preprint arXiv:1809.11096, 2018.
- [6] Andrew Brock, Theodore Lim, James M Ritchie, and Nick Weston. Neural photo editing with introspective adversarial networks. arXiv preprint arXiv:1609.07093, 2016.
- [7] Liang-Chieh Chen, Maxwell Collins, Yukun Zhu, George Papandreou, Barret Zoph, Florian Schroff, Hartwig Adam, and Jon Shlens. Searching for efficient multi-scale architectures for dense image prediction. In Advances in Neural Information Processing Systems, pages 8713–8724, 2018.
- [8] Liang-Chieh Chen, George Papandreou, Iasonas Kokkinos, Kevin Murphy, and Alan L Yuille. Deeplab: Semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected crfs. IEEE transactions on pattern analysis and machine intelligence, 40(4):834–848, 2018.
- [9] Liang-Chieh Chen, George Papandreou, Florian Schroff, and Hartwig Adam. Rethinking atrous convolution for semantic image segmentation. arXiv preprint arXiv:1706.05587, 2017.
- [10] Patryk Chrabaszcz, Ilya Loshchilov, and Frank Hutter. A downsampled variant of imagenet as an alternative to the cifar datasets. arXiv preprint arXiv:1707.08819, 2017.
- [11] Ian Goodfellow, Jean Pouget-Abadie, Mehdi Mirza, Bing Xu, David Warde-Farley, Sherjil Ozair, Aaron Courville, and Yoshua Bengio. Generative adversarial nets. In Advances in neural information processing systems, pages 2672–2680, 2014.
- [12] Guillermo L Grinblat, Lucas C Uzal, and Pablo M Granitto. Class-splitting generative adversarial networks. arXiv preprint arXiv:1709.07359, 2017.
- [13] Ishaan Gulrajani, Faruk Ahmed, Martin Arjovsky, Vincent Dumoulin, and Aaron C Courville. Improved training of wasserstein gans. In Advances in Neural Information Processing Systems, pages 5767–5777, 2017.
- [14] Guang-He Lee Yonglong Tian Hao He, Hao Wang. Progan: Towards probabilistic gan with theoretical guarantees. In ICLR, 2019.
- [15] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. Deep residual learning for image recognition. In Proceedings of the IEEE conference on computer vision and pattern recognition, pages 770–778, 2016.
- [16] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. Identity mappings in deep residual networks. In European conference on computer vision, pages 630–645. Springer, 2016.
- [17] Yihui He, Ji Lin, Zhijian Liu, Hanrui Wang, Li-Jia Li, and Song Han. Amc: Automl for model compression and acceleration on mobile devices. In European Conference on Computer Vision, pages 815–832. Springer, 2018.
- [18] Martin Heusel, Hubert Ramsauer, Thomas Unterthiner, Bernhard Nessler, Gunter Klambauer, and Sepp Hochreiter. Gans trained by a two time-scale update rule converge to a nash equilibrium. arXiv preprint arXiv:1706.08500, 12(1), 2017.
- [19] Quan Hoang, Tu Dinh Nguyen, Trung Le, and Dinh Phung. Mgan: Training generative adversarial nets with multiple generators. 2018.
- [20] Sepp Hochreiter and Jürgen Schmidhuber. Long short-term memory. Neural computation, 9(8):1735–1780, 1997.
- [21] Sergey Ioffe and Christian Szegedy. Batch normalization: Accelerating deep network training by reducing internal co-variate shift. arXiv preprint arXiv:1502.03167, 2015.
- [22] Phillip Isola, Jun-Yan Zhu, Tinghui Zhou, and Alexei A Efros. Image-to-image translation with conditional adversarial networks. In 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pages 5967–5976. IEEE, 2017.
- [23] Yifan Jiang, Xinyu Gong, Ding Liu, Yu Cheng, Chen Fang, Xiaohui Shen, Jianchao Yang, Pan Zhou, and Zhangyang Wang. Enlightengan: Deep light enhancement without paired supervision. arXiv preprint arXiv:1906.06972, 2019.
- [24] Haifeng Jin, Qingquan Song, and Xia Hu. Auto-keras: Efficient neural architecture search with network morphism, 2018.
- [25] Alexia Jolicoeur-Martineau. The relativistic discriminator: a key element missing from standard gan. arXiv preprint arXiv:1807.00734, 2018.
- [26] Tero Karras, Timo Aila, Samuli Laine, and Jaakko Lehtinen. Progressive growing of gans for improved quality, stability, and variation. ICLR, 2018.
- [27] Tero Karras, Samuli Laine, and Timo Aila. A style-based generator architecture for generative adversarial networks. arXiv preprint arXiv:1812.04948, 2018.
- [28] Diederik P Kingma and Jimmy Ba. Adam: A method for stochastic optimization. arXiv preprint arXiv:1412.6980, 2014.
- [29] Alex Krizhevsky and Geoffrey Hinton. Learning multiple layers of features from tiny images. Technical report, Cite-seer, 2009.
- [30] Alex Krizhevsky, Ilya Sutskever, and Geoffrey E Hinton. Imagenet classification with deep convolutional neural networks. In Advances in neural information processing systems, pages 1097–1105, 2012.
- [31] Orest Kupyn, Volodymyr Budzan, Mykola Mykhailych, Dmytro Mishkin, and Jiří Matas. Deblurgan: Blind motion deblurring using conditional adversarial networks. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pages 8183–8192, 2018.
- [32] Karol Kurach, Mario Lucic, Xiaohua Zhai, Marcin Michalski, and Sylvain Gelly. The gan landscape: Losses, archi-

- tures, regularization, and normalization. arXiv preprint arXiv:1807.04720, 2018.
- [33] Christian Ledig, Lucas Theis, Ferenc Huszar, Jose Caballero, Andrew Cunningham, Alejandro Acosta, Andrew Aitken, Alykhan Tejani, Johannes Totz, Zehan Wang, et al. Photo-realistic single image super-resolution using a generative adversarial network. In 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pages 105–114. IEEE, 2017.
  - [34] Liam Li and Ameet Talwalkar. Random search and reproducibility for neural architecture search. arXiv, 2019.
  - [35] Tsung-Yi Lin, Piotr Dollar, Ross Girshick, Kaiming He, Bharath Hariharan, and Serge Belongie. Feature pyramid networks for object detection. In CVPR, volume 1, page 4, 2017.
  - [36] Chenxi Liu, Liang-Chieh Chen, Florian Schroff, Hartwig Adam, Wei Hua, Alan Yuille, and Li Fei-Fei. Auto-deeplab: Hierarchical neural architecture search for semantic image segmentation. arXiv preprint arXiv:1901.02985, 2019.
  - [37] Chenxi Liu, Barret Zoph, Maxim Neumann, Jonathon Shlens, Wei Hua, Li-Jia Li, Li Fei-Fei, Alan Yuille, Jonathan Huang, and Kevin Murphy. Progressive Neural Architecture Search. pages 19–34, 2018.
  - [38] H Liu, K Simonyan, Y Yang arXiv preprint arXiv 1806.09055, and 2018. Darts: Differentiable architecture search. arxiv.org.
  - [39] Mario Lucic, Karol Kurach, Marcin Michalski, Sylvain Gelly, and Olivier Bousquet. Are gans created equal? a large-scale study. In Advances in neural information processing systems, pages 700–709, 2018.
  - [40] Minh-Thang Luong, Quoc V Le, Ilya Sutskever, Oriol Vinyals, and Lukasz Kaiser. Multi-task sequence to sequence learning. arXiv preprint arXiv:1511.06114, 2015.
  - [41] Mehdi Mirza and Simon Osindero. Conditional generative adversarial nets. arXiv preprint arXiv:1411.1784, 2014.
  - [42] Takeru Miyato, Toshiki Kataoka, Masanori Koyama, and Yuichi Yoshida. Spectral normalization for generative adversarial networks. arXiv preprint arXiv:1802.05957, 2018.
  - [43] Alejandro Newell, Kaiyu Yang, and Jia Deng. Stacked hour-glass networks for human pose estimation. In European Conference on Computer Vision, pages 483–499. Springer, 2016.
  - [44] Tu Nguyen, Trung Le, Hung Vu, and Dinh Phung. Dual discriminator generative adversarial nets. In Advances in Neural Information Processing Systems, pages 2670–2680, 2017.
  - [45] Augustus Odena, Vincent Dumoulin, and Chris Olah. De-convolution and checkerboard artifacts. Distill, 1(10):e3, 2016.
  - [46] H Pham, M Y Guan, B Zoph, Q V Le, J Dean arXiv preprint arXiv, and 2018. Efficient Neural Architecture Search via Parameter Sharing. arxiv.org.
  - [47] Alec Radford, Luke Metz, and Soumith Chintala. Unsupervised representation learning with deep convolutional generative adversarial networks. arXiv preprint arXiv:1511.06434, 2015.
  - [48] S Reed, Z Akata, X Yan, L Logeswaran arXiv preprint arXiv, and 2016. Generative adversarial text to image synthesis. jmlr.org, 2016.
  - [49] Olaf Ronneberger, Philipp Fischer, and Thomas Brox. U-net: Convolutional networks for biomedical image segmentation. In International Conference on Medical image computing and computer-assisted intervention, pages 234–241. Springer, 2015.
  - [50] Tim Salimans, Ian Goodfellow, Wojciech Zaremba, Vicki Cheung, Alec Radford, and Xi Chen. Improved techniques for training gans. In Advances in Neural Information Processing Systems, pages 2234–2242, 2016.
  - [51] Christian Szegedy, Wei Liu, Yangqing Jia, Pierre Sermanet, Scott Reed, Dragomir Anguelov, Dumitru Erhan, Vincent Vanhoucke, and Andrew Rabinovich. Going deeper with convolutions. In Proceedings of the IEEE conference on computer vision and pattern recognition, pages 1–9, 2015.
  - [52] Christian Szegedy, Vincent Vanhoucke, Sergey Ioffe, Jon Shlens, and Zbigniew Wojna. Rethinking the inception architecture for computer vision. In Proceedings of the IEEE conference on computer vision and pattern recognition, pages 2818–2826, 2016.
  - [53] Dustin Tran, Rajesh Ranganath, and David M Blei. Deep and hierarchical implicit models. CoRR, abs/1702.08896, 2017.
  - [54] Ngoc-Trung Tran, Tuan-Anh Bui, and Ngai-Man Cheung. Dist-gan: An improved gan using distance constraints. In Proceedings of the European Conference on Computer Vision (ECCV), pages 370–385, 2018.
  - [55] Dmitry Ulyanov, Andrea Vedaldi, and Victor Lempitsky. In-stance normalization: The missing ingredient for fast stylization. arXiv preprint arXiv:1607.08022, 2016.
  - [56] Wei Wang, Yuan Sun, and Saman Halgamuge. Improving mmd-gan training with repulsive loss function. arXiv preprint arXiv:1812.09916, 2018.
  - [57] Zhangyang Wang, Shiyu Chang, Yingzhen Yang, Ding Liu, and Thomas S Huang. Studying very low resolution recognition using deep networks. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pages 4792–4800, 2016.
  - [58] David Warde-Farley and Yoshua Bengio. Improving generative adversarial networks with denoising feature matching. (2017). In ICLR, 2017.
  - [59] Ronald J Williams. Simple statistical gradient-following algorithms for connectionist reinforcement learning. Machine learning, 8(3-4):229–256, 1992.
  - [60] Lingxi Xie and Alan Yuille. Genetic cnn. In 2017 IEEE International Conference on Computer Vision (ICCV), pages 1388–1397. IEEE, 2017.
  - [61] Tao Xu, Pengchuan Zhang, Qiuyuan Huang, Han Zhang, Zhe Gan, Xiaolei Huang, and Xiaodong He. Attngan: Fine-grained text to image generation with attentional generative adversarial networks. arXiv preprint, 2017.
  - [62] Jianwei Yang, Anitha Kannan, Dhruv Batra, and Devi Parikh. Lr-gan: Layered recursive generative adversarial networks for image generation. arXiv preprint arXiv:1703.01560, 2017.
  - [63] Shuai Yang, Zhangyang Wang, Zhaowen Wang, Ning Xu, Jiaying Liu, and Zongming Guo. Controllable artistic text style transfer via shape-matching gan. arXiv preprint arXiv:1905.01354, 2019.

- [64] Han Zhang, Ian Goodfellow, Dimitris Metaxas, and Augustus Odena. Self-attention generative adversarial networks. arXiv preprint arXiv:1805.08318, 2018.
- [65] Han Zhang, Tao Xu, Hongsheng Li, Shaoqing Zhang, Xiao-gang Wang, Xiao-lei Huang, and Dimitris Metaxas. Stack-gan++: Realistic image synthesis with stacked generative adversarial networks. arXiv preprint arXiv:1710.10916, 2017.
- [66] Han Zhang, Tao Xu, Hongsheng Li, Shaoqing Zhang, Xiao-gang Wang, Xiao-lei Huang, and Dimitris N Metaxas. Stack-gan: Text to photo-realistic image synthesis with stacked generative adversarial networks. In Proceedings of the IEEE International Conference on Computer Vision, pages 5907–5915, 2017.
- [67] Xiaofeng Zhang, Zhangyang Wang, Dong Liu, and Qing Ling. Dada: Deep adversarial data augmentation for extremely low data regime classification. In ICASSP 2019-2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), pages 2807–2811. IEEE, 2019.
- [68] Junbo Zhao, Michael Mathieu, and Yann LeCun. Energy-based generative adversarial network. arXiv preprint arXiv:1609.03126, 2016.
- [69] Zhao Zhong, Junjie Yan, Wei Wu, Jing Shao, and Cheng-Lin Liu. Practical block-wise neural network architecture generation. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pages 2423–2432, 2018.
- [70] Jun-Yan Zhu, Taesung Park, Phillip Isola, and Alexei A Efros. Unpaired Image-to-Image Translation using Cycle-Consistent Adversarial Networks. arXiv.org, Mar. 2017.
- [71] Barret Zoph and Quoc V Le. Neural architecture search with reinforcement learning. arXiv preprint arXiv:1611.01578, 2016.
- [72] Barret Zoph, Vijay Vasudevan, Jonathon Shlens, and Quoc V Le. Learning Transferable Architectures for Scalable Image Recognition. CVPR, 2018.
- [73] Barret Zoph, Deniz Yuret, Jonathan May, and Kevin Knight. Transfer learning for low-resource neural machine translation. arXiv preprint arXiv:1604.02201, 2016.