

基于空间注意力机制的快速生成对抗网络

吴天宝, 徐 芳

西南石油大学理学院, 四川 成都

收稿日期: 2023年1月6日; 录用日期: 2023年2月3日; 发布日期: 2023年2月13日

摘 要

针对实际应用中数据量不足导致的图像生成质量低下以及模型收敛缓慢的问题, 提出了一种基于注意力机制的快速生成对抗网络模型。模型引入简化的注意力机制模块后能够获取图像的全局信息。模型由带有注意力机制的残差模块配合跳层连接机制为主。同时, 对抗学习使得模型生成的图像更加逼真。为了加快收敛速度, 元数据作为模型的输入为模型提供基本图像的基本信息。实验结果表明, 对比现有的模型, 该模型在小数据集上不仅收敛速度更快, 成像也更加逼真。在数据量足够的数据集上对比现有的模型, 能够生成质量非常接近的图像, 并且大大缩短了收敛的时间, 减少了占用的内存。

关键词

生成对抗网络, 图像生成, 自注意力机制, 卷积神经网络

Rapid Generation of Adversarial Networks Based on Spatial Attention Mechanism

Tianbao Wu, Fang Xu

School of Science, Southwest Petroleum University, Chengdu Sichuan

Received: Jan. 6th, 2023; accepted: Feb. 3rd, 2023; published: Feb. 13th, 2023

Abstract

Aiming at the problems of low image generation quality and slow model convergence caused by insufficient data in practical applications, a fast generation adversarial network model based on attention mechanism is proposed. The model can obtain the global information of the image by introducing the simplified attention mechanism module. The model is mainly composed of a residual module with an attention mechanism and a layer-jumping connection mechanism. At the same time, adversarial learning makes the image generated by the model more realistic. In order to speed up the convergence, metadata as the input of the model provides the basic information of the basic image for the model. Experimental results show that the proposed model not only converges faster on small data sets, but also has more realistic imaging. Compared with the existing

modules on the data set with sufficient data, this model can generate images of very similar quality, and greatly shorten the time of convergence and reduce the memory consumption.

Keywords

Generate Adversarial Network, Image Generation, Self-Attention Mechanism, Convolutional Neural Network

Copyright © 2023 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

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



Open Access

1. 引言

生成式对抗网络(Generative adversarial network, GAN)自 Ian Goodfellow [1]等人提出后, 越来越受到学术界和工业界的重视。随着对抗生成网络的发展, 其在图像与视频的生成[2] [3] [4]、图像翻译[5]、图像修复[6]等领域都取得了巨大的成功。然而, 这些成功却有着严苛的条件, 需要大量的数据作为支撑, 这导致了生成式对抗网络在小数据集上取得的结果很差。并且获得高分辨率的图像模型往往需要大显存和强算力的显卡, 例如使用大型的服务器或者专业的显卡比如 NVIDIA V100 等, 对设备的要求较高。因此迫切需要开发轻便且适用于小数据集的对抗神经网络。

在有限的数下训练 GAN 通常会导致过度拟合和不稳定性问题[7]。由于数据量过少, 判别器(Discriminator, D)无法获取更多的特征来分辨图像的真假, 同时对于生成器而言无法提供有效的特征信息指导其生成图像, 导致生成的图片质量低下。如何在有限的数下获取足够的数特征以此来改善图片的质量, 这仍然是待解决的问题。该问题的最新方法包括使用不同的数扩充训练集; 使用新的正则项来约束模型的输出; 使用新的架构获取更多的特征指导图像生成。由于数据集扩充的使用场景有限且对于图像质量提升也是有限的, 现在主要的做法是提升模型的特征获取能力增加其容量, 然后以数扩充作为辅助。

为了保持模型轻便性的同时又提升模型对于全局视野的感受能力以及图片的质量, 本文引入注意力机制[8]。注意力机制最初应用于自然语言处理中, 加入注意力机制的模型能够更好地关注语句的前后顺序。加入了注意力机制的模型相比于传统的卷积模块为主导的网络而言能更好地获取图片全局关系。本文将使用空间注意力机制模块, 该模块仅使用两个池化层和单个卷积层。池化层的加入增加了模型的感受范围。注意力机制模块获取的空间信息可以通过残差模块进入到更深层的网络中, 减少了空间信息的损失。为了更好地保持模型的收敛性, 本文将在生成器(Generator, G)中使用元数据作为生成器的输入。元数据中包含低维度的空间信息可以为目标图像提供基本信息(例如, 对象的类别和基本纹理等), 这些基本的信息可以提高模型的收敛速度, 并且能够稳定训练过程。基于注意力机制和元数据代替一般数据输入这两种方法本文提出一种适用于小数据集下的生成模型, 称为基于空间注意力机制的快速生成模型。该模型在 FASTGAN 的基础上使用带有空间注意力机制的残差模块代替 FASTGAN 的残差块, 在输入生成器的噪声中添加了元数据。相较于 FASTGAN 模型, 本文的模型在小型数据集上的发挥更为出色, 能够直观地提升图片的质量。本文的模型在提升图片质量的同时几乎不增加计算的负担。

本文的主要贡献有 3 个方面:

- 1) 本文引入空间注意力机制, 重新设计了残差模块。
- 2) 使用新的残差模块与元数据, 有效地改善了图像的质量。在有限的数据集上对于图像质量有明显

的提升, 在数据量足够的数据集上生成的图像包含更多的细节。

3) 本文的模型相较于 FASTGAN 几乎不增加计算成本。

2. 相关工作

生成对抗性网络旨在生成可信和真实的图像, 自开创以来取得了巨大进展[1]。然而有一个问题始终没能解决, GAN 的训练容易出现模式崩溃和不稳定[9]。为了解决这个问题, 本文通过设计新的优化目标和网络架构来稳定训练并提高综合质量的模型。WGAN [10]和 f-GAN [11]最小化了真实分布和生成分布的 Wasserstein 距离, 而不是 JS 散度。BigGAN [12]和 StyleGAN [13] [14]系列在制作逼真图像方面取得了突破性进展。SAGAN [15]首次引入注意力机制模块辅助图像的生成。然而, 这些模型的体积较大, 当给定有限的数据时会导致模型不收敛、成像质量较差或者无法生成、过拟合等问题。

在有限的数下提高合成质量仍然是一个未被解决的问题, 这一问题近年来引起了广泛的关注。训练数据不足会导致鉴别器过度拟合, 从而降低生成图像的质量。解决这种数据匮乏的一个直接方法是通过各种扩充来扩展训练集。除了采用传统的增强技术(例如, 翻转、裁剪), ADA [16]和 DiffAug [17]分别提出了自适应和可微增强来放大训练数据。APA [7]基于自适应伪增强的过拟合程度来欺骗 D。InsGen [18]将实例识别作为辅助任务, 以鼓励 D 区分每个单独的图像, 这提高了鉴别器的鉴别力。

DCGAN [19]使用卷积层来代替线性层, 大幅改善了 GAN 的画质。Lecam [20]在整个训练过程中使鉴别器的输出正规化。StackGAN [21]采用分段式结构, 同时使用多个判别器来稳定获取特征。FastGAN [22]采用跳跃层信道激励模块和自监督鉴别器来稳定和加速训练。

FASTGAN 提供了一个非常好的轻量级网络架构, 其结构简洁易懂。FASTGAN 在 DCGAN 的基础上设计了新的残差块, 延用了跳层连接的设计。这使得模型能够获得快速的梯度流。FASTGAN 在 RTX3060 等入门显卡上都可以运行。

该模型最大的优势在于给定很小的数据集也能快速的收敛, 当样本的数据量少于 1000 幅图像时(多数时候只有 100 幅图像时), 该模型也能收敛, 并生成较高质量的图像。但是在实际使用过程中, 由于追求更快的收敛速度从而导致模型对于全局信息和高频信息的感知能力不足, 例如使用 FFHQ 数据集生成 1024*1024 高分别率图片时片总会存在图片坐上角的头发的缺失或者爆炸。并且还会出现左右脸有明显的差异, 不符合人的感官。(如图 1 所示)

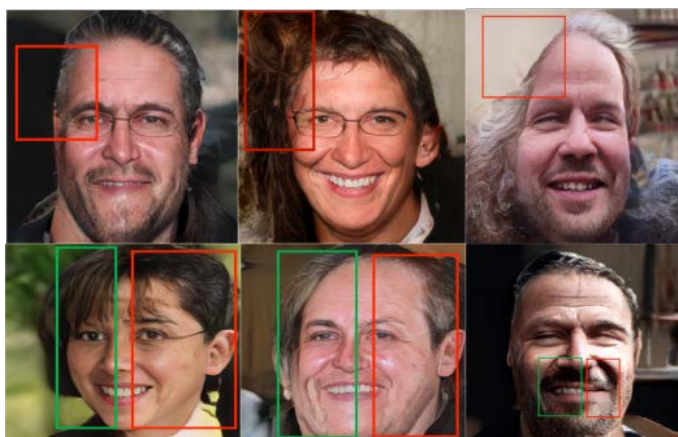


Figure 1. FASTGAN trained 50,000 FFHQ images on a RTX2080-Ti GPU for 24 hours at 1024*1024 resolution

图 1. FASTGAN 在 1024*1024 的分辨率下使用五万张 FFHQ 图片在 RTX2080-Ti GPU 上训练 24 个小时得到的结果

3. 思考与改进

FASTGAN 是一个轻量级架构, 其获得的特征信息更多的是依靠快速的梯度流, 这样的方式使模型在生成高分辨率的图片时能够快速收敛, 并获得质量更好的图像。FastGAN 使用跳跃层连接增加梯度流的强度同时使用改进后的跳层激励模块(Skip-Layer Excitation module, SLE), 通过跳连接这样的方式, 实现了为不同卷积层在激活之前的元素添加。在激活之前应用通道上的乘法(可以在不同分辨率上相乘), 消除了繁重的卷积计算。其次, 在先前的 GAN 工作中, 跳过连接仅在相同的分辨率内使用。相比之下, 由于不再需要相等的空间维度, 我们在具有更长范围(例如, 8 和 128、16 和 256)的分辨率之间执行跳过连接。这两种设计使得 SLE 继承了残差模块(ResBlock)的优点, 具有快捷的梯度流, 同时不需要额外的计算负担。

FastGAN 通过重新设计残差块来增加信息的流速, 以此来解决模型卷积层数量太少(高分辨率层仅使用两层卷积)所导致的生成高分辨率图像收敛缓慢和特征信息丢失的问题。然而 SLE 模块对于信息的承载能力和稳固性却被忽略了, 在快速的梯度流下会使得空间信息丢失。为了改善以上缺点, 本文首先考虑模型的体积和收敛速度。我们沿用 FastGAN 的架构, 在其基础上中进行两点改进。

3.1. 改进一

对 SLE 模块进行改进。京东方团队 2019 年提出的一种单卷积层加上注意力机制实现超分辨率图片的模型。2019 年提出的 SAGAN 中使用注意力机制在远距离依赖(long-range dependency)下来实现了图片生成任务。生成器产出的图片可以在判别器中检查图像中部和远端之间的详细特征是否一致。我们的目标是通过使用注意力机制来增加表征能力, 即专注于重要特征, 抑制不必要的特征。由于卷积运算通过将跨通道和空间信息混合在一起提取信息特征, 因此我们希望通过注意力机制模块来强调沿空间轴这个主要维度的有意义的特征。注意力机制模块通过学习强调或抑制一些特征来有效地帮助网络中的信息流动。在 SLE 模块上增加空间注意力机制模块(Spatial Attention Module), 得到新的模块, 称为跳层注意力激励模块(Skip-Layer Attention Excitation module, SLAE)。

图 2 的架构中, FastGAN 直接将低分辨率和高分辨率的特征图直接输入 SLE 模块中。在 SLAE 模块中低分辨率下的特征图要先经过空间注意力机制模块后保存低纬度的空间信息, 然后使用乘法将提取到的信息进行整合, 借由跳连接输入到高维的图像中, 这样可以减少空间信息的损失。(如图 3 所示)

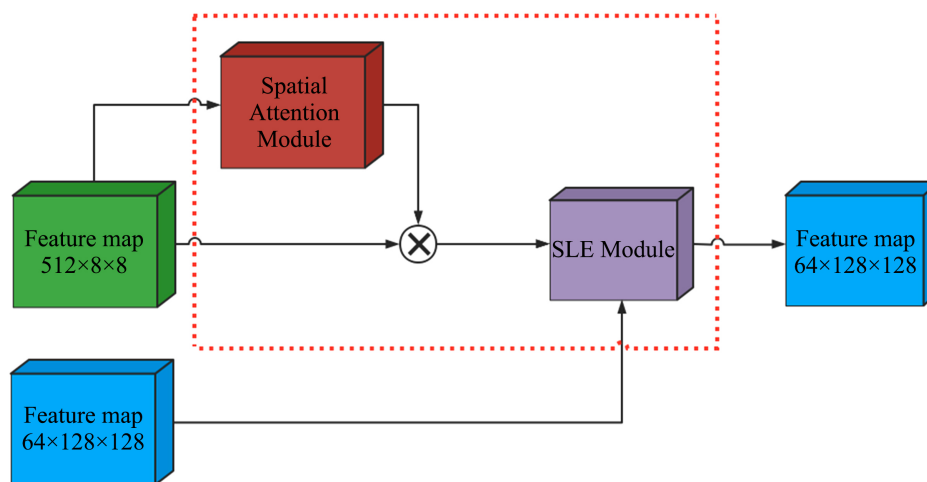


Figure 2. Skip-Layer Attention Excitation module

图 2. 跳层注意力激励模块

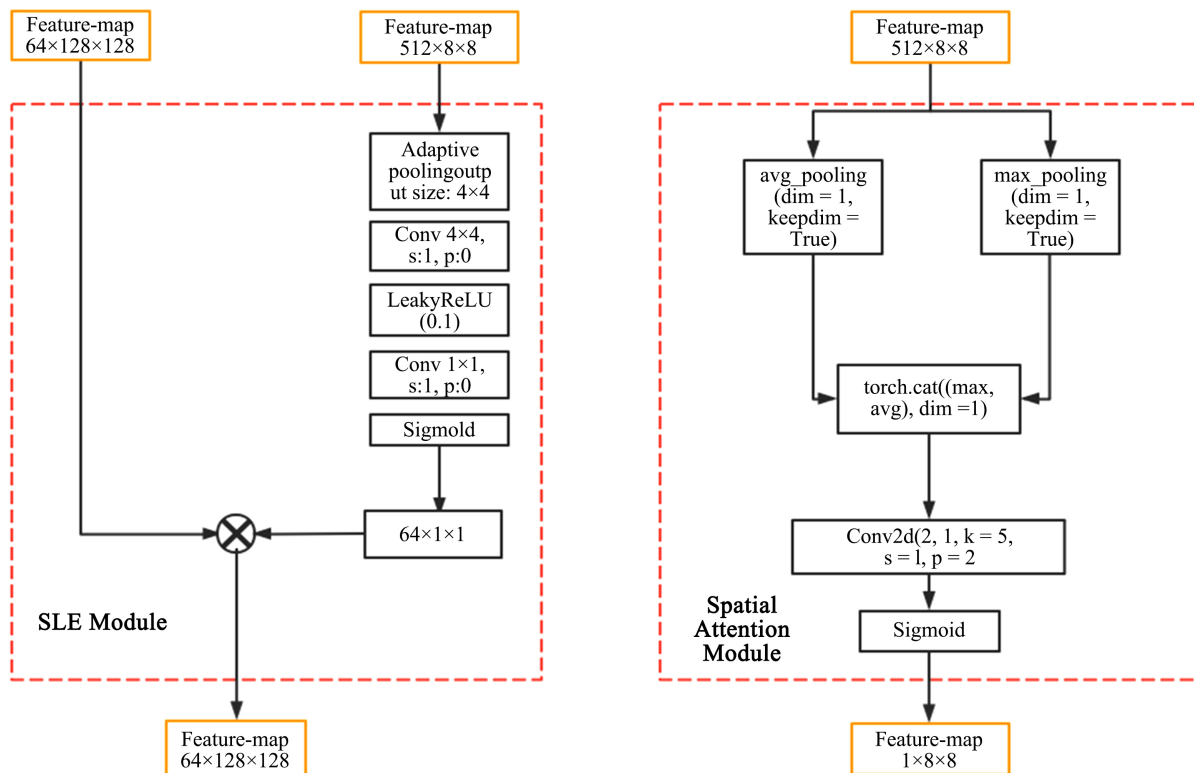


Figure 3. Skip-Layer Attention Excitation module and spatial attention module

图 3. 跳层激励模块与空间注意力模块

3.2. 改进二

对输入生成器的数据做一些改进，为了加快训练过程，本文将元数据(metadata)纳入模型中，为生成器提供关于目标图像的基本信息，其中该信息可能包含生成器最终生成真实图像所需要的特征分布的线索，这可以帮助其提前了解要合成的对象类型。

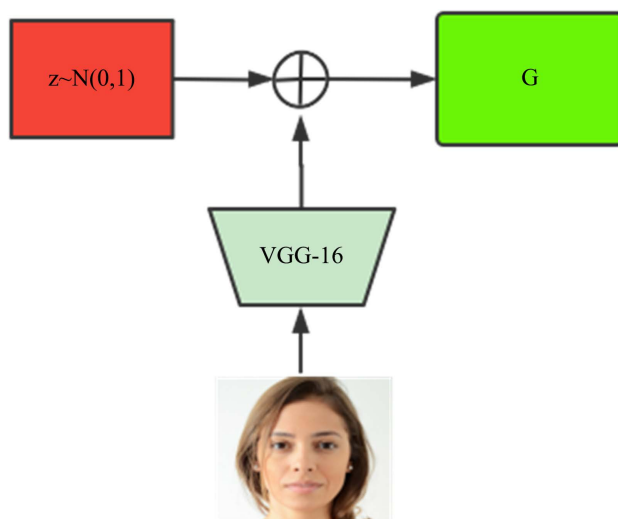


Figure 4. Metadata structure input structure diagram

图 4. 元数据结构输入结构图

此外, 为了防止所提供的元数据强制模型实现来自元图像的身份转换, 我们使用预训练的 VGG-16 网络的深层从给定的元图像中提取特征, 把这些深层的特征作为元数据, 因此元数据仅保留汇总的空间信息。之所以这样设计, 主要是因为正常的空间特征包含了太多关于目标图像的细节, 例如对象的颜色、形状、姿势和位置等。因此, 根据平均每个通道中的所有空间特征, 可以过滤这些细节, 并且只保留基本信息, 确保模型的良好多样性。具体过程见图 4。

对要进行迭代的图片使用 VGG-16 的深层(本文使 VGG-16 的 relu5_3 输出特征, 可根据数据图片分辨率大小适当地加深或者减少)来进行特征的提取然后相加到初始的噪声中。

3.3. 损失函数

采用对抗性损失函数的较链版本来训练我们的模型。

$$L_D = -\mathbb{E}_{x \sim I_{real}} [\min(0, -1 + D(x))] - \mathbb{E}_{\hat{x} \sim G(z)} [\min(0, -1 - D(\hat{x}))] + L_{recons}$$

$$L_G = -\mathbb{E}_{z \sim N} [D(G(x))]$$

使用重建损失来鼓励鉴别器提取更具有代表性的特征。

$$L_{recons} = \mathbb{E}_{f \sim D_{encode}(x), x \sim I_{real}} [\|g(f) - \mathcal{T}(x)\|]$$

其中 f 是来自 D 的中间特征图, 函数 G 包含对 f 和解码器的处理, \mathcal{T} 函数表示对来自真实图像 I_{real} 的样本 x 的处理。

4. 实验

设备: 特别指出由于 StyleGAN2 在 1024*1024 分辨率下训练时的显存要求, FASTGAN 原文中是在 RTX TITAN GPU 上训练 StyleGAN2。其余模型都在 RTX2080-TI 上运行, RTX2080-TI 和 TITAN 具有相似的性能, 模型在两个 GPU 上运行的时间相同。为了控制变量, 本文仍然使用这两张 GPU。

数据集: 为了体现改进的方法的有效性以及进一步控制变量, 使用了与 FASTGAN 相同的数据集。在 256×256 分辨率下, 测试了动物脸狗和猫、100 镜头奥巴马、熊猫和脾气暴躁的猫。在 1024×1024 分辨率下, 测试了 Flickr FaceHQ (FFHQ)、牛津花、WikiArt 的艺术画、Unsplash 的自然景观照片、Pokemon、动漫脸、头骨和贝壳。这些数据集旨在涵盖具有不同特征的图像: 写实照片、图形插图和类似艺术的图像。

评估指标: 1) 我们采用 Fréchet 起始距离(FID)来度量模型生成的图片的质量, FID 量化生成图像和真实图像的分布之间的距离。对于少于 1000 张图像的数据集(大多数只有 100 张图像), 我们让 G 生成 5000 张图像, 并计算合成图像和整个训练集之间的 FID。2) 学习感知相似性(LPIPS)提供了两个图像之间的感知距离。以上两种指标均为数值越低成像质量越高。

使用 LPIPS 报告生成器在给定真实图像上执行潜在空间反向跟踪时的重建质量, 并测量自动编码性能。发现没有必要涉及其他指标, FID 不太可能与其他指标不一致, 改进后的模型和比较模型之间存在显著的性能差距。

比较模型: 将我们的模型与: 1) 最先进的(SOTA)无条件模型 StyleGAN2, 2) FASTGAN 模型, 进行比较。由于 StyleGAN2 需要更多的计算成本(cc)来训练, 因此我们把 FASTGAN 作为基线模型(Baseline), 这样更能方便直观的体现模型间性能的差异。本文将改进后的模型与 StyleGAN2 在绝对图像合成质量上进行了比较, 而不考虑计算成本, 并在可比较的范围内使用基线模型作为参考。

Table 1. Calculation cost comparison of models

表 1. 模型的计算成本比较

		StyleGAN2	Baseline	Ours
分辨率: 256*256 批量大小: 8	训练时间(hour/10 k iter)	3.8	1	1.03
	训练所需显存(GB)	18	6.5	6.7
	模型参数量(million)	108.843	47.363	47.563
分辨率: 1024*1024 批量大小: 8	训练时间(hour/10 k iter)	7	1.7	1.7
	训练所需显存(GB)	36	10	10.5
	模型参数量(million)	109.229	47.413	47.734

以上的计算不包含提取元数据的过程, 因为元数据在一次提取后可以重复的使用, 并且可以与模型的迭代分开进行。StyleGAN2 模型由于其庞大的显存需求在 RTX TITAN GPU 上计算获得表 1 的结果, 其余均在 RTX2080-TI 上计算获得。本文的模型延续了轻巧的特性, 对比 StyleGAN2, 其参数量只有其二分之一。对比基线模型迭代时间与参数量几乎没有变化。

Table 2. FID comparison of a few sample data sets at 256*256 resolution

表 2. 256*256 分辨率下少数样本数据集的 FID 比较

		Animal Face-Dog	Animal Face-Cat	Obama	Panda	Grumpy-cat	
图片数量		389	160	100	100	100	
RTX2080TI 上训练时间	20 小时	StyleGAN2	58.85	42.44	46.87	12.06	27.08
		StyleGAN2finetune	61.03	46.07	35.75	14.5	29.34
		Baseline	50.66	35.11	41.05	10.03	26.65
		Baseline + Spatial Attention	44.32	33.21	37.26	9.32	23.13
	5 小时	Baseline + metadata	48.32	34.82	39.26	10.12	25.82
		Baseline + metadata + Spatial Attention (ours)	43.15	32.13	36.79	9.2	22.21

Table 3. FID comparison of a few sample data sets at 1024*1024 resolution

表 3. 1024*1024 分辨率下少数样本数据集的 FID 比较

		Art Paintings	FFHQ	Flower	Pokemon	Anime Face	Skull	Shell	
图片数量		1000	1000	1000	800	120	100	100	
RTX TITAN 上训练时间	20 小时	StyleGAN2	74.56	25.66	45.23	190.23	127.98	241.37	
		StyleGAN2 finetune	N/A	N/A	36.72	60.12	61.23	107.68	220.45
		Baseline	45.08	24.45	25.66	57.19	59.38	130.05	155.47
	5 小时	Baseline + metadata + Spatial Attention (ours)	43.16	21.12	21.39	51.16	53.74	102.35	131.56

通过表 1、表 2 与表 3, 容易发现模型在增加了一层卷积的前提下几乎不会增加模型的训练时间, 在相同的时间内获得了更好的成像质量, 说明模型并不是单纯的依靠卷积块的堆叠来提升性能。通过对比消融实验的结果, 发现 SLAE 模块对于小数据集的效果是显著的, 数据量少的数据集获得的提升更为明显。



Figure 5. Model adds metadata to the comparison diagram
图 5. 模型加入元数据的对比图

在图 5 中, 仅在基线模型上加入元数据的情况下, 容易看出图像收敛的更快了, 能够更快速的得到人脸的特征(语句需要简练)。在未加入元数据的情况下, 虽然对人脸的明显特征能很好的捕捉, 但是对于背景以及细节的刻画很抽象, 缺少对于全局信息的捕捉。

本文的模型在小样本上能获得更优良的性能, 为了进行更彻底的评估, 我们还使用更充分的训练样本在数据集上测试模型。在两个 TITANRTXGPU 上的批量大小为 16 训练三个模型, StyleGAN2 训练 5 天, 基线模型与本文的模型只训练 24 小时。

Table 4. FID comparison of a few sample data sets at 1024*1024 resolution

表 4. 1024*1024 分辨率下少数样本数据集的 FID 比较

图片数量	Art Paintings			FFHQ			Nature Photograph		
	2 k	5 k	10 k	2 k	5 k	10 k	2 k	5 k	10 k
StyleGAN2	70.02	48.36	41.23	18.38	10.45	7.86	67.12	41.27	39.05
Baseline	44.57	43.27	42.53	19.01	17.93	16.45	52.47	45.07	43.65
Baseline + metadata + Spatial Attention (ours)	41.87	40.37	39.57	19.56	16.38	15.12	49.47	45.56	41.37

Table 5. The LPIPS value of the generator

表 5. 生成器的 LPIPS 值

模型	迭代次数	Cat	Dog	FFHQ	Art
		256*256		1024*1024	
Baseline	20 k iter	1.821	1.918	2.425	2.624
Baseline	40 k iter	1.897	1.986	2.342	2.601
Ours	40 k iter	1.793	1.853	2.403	2.578

通过表 4 与表 5, 不难发现, 我们的模型相比于数据量充足的数据集有不小的提升, 在数据量足够的情况下依然没有办法超过 StyleGAN2 模型, 这受限于模型的体积容量。但是, 我们的模型花费的时间成本更少。通过表 5, 可以发现改进的 SLAE 模块提升了生成器重建图片的质量。

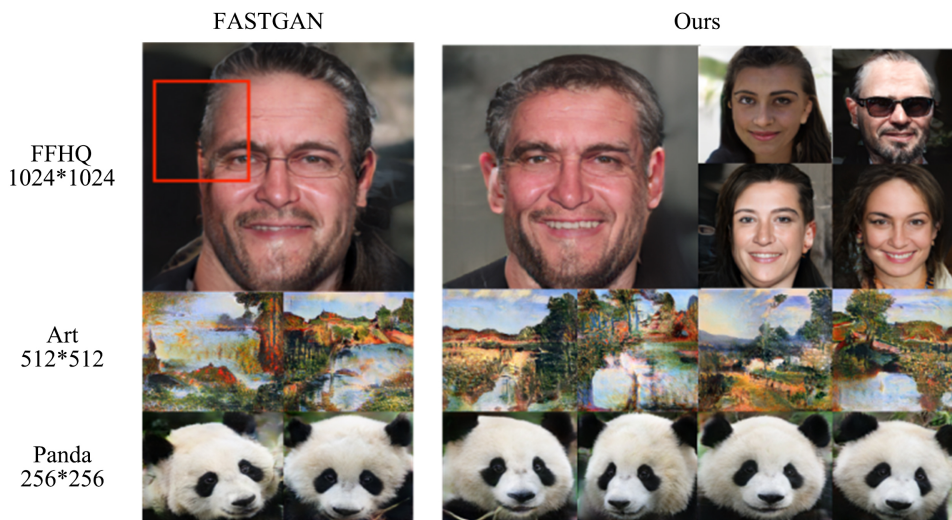


Figure 6. Partial results are displayed at each resolution

图 6. 各个分辨率下的部分结果展示

5. 结论

本文提出两种方法来提升模型的信息承载能力, 改进后的模型能够获取更多空间上的信息。在小样

本数据集上的作用更为明显, 对于依赖空间结构的人脸或者动物的脸等数据集, 能够使其产生的图片更加符合人的实际观感, 进一步提升了图片的质量。虽然在大型数据集上我们的模型用较小的计算代价获得了质量不错的图片(图 6 展示部分结果), 但是在大型数据集下模型受限于本身的体积, 当数据量到达一定的等级后模型的提升幅度无法与数据的增量相匹配。总体来说该模型更适合于数据量较少、设备有限、时间成本要求较低等情况。

参考文献

- [1] Goodfellow, I., Pouget-Abadie, J., Mirza, M., et al. (2014) Generative Adversarial Nets. *Proceedings of the 27th International Conference on Neural Information Processing Systems*, Volume 2, 2672-2680.
- [2] Karras, T., Aittala, M., Laine, S., Härkönen, E., Hellsten, J., Lehtinen, J. and Aila, T. (2021) Alias-Free Generative Adversarial Networks. *Advances in Neural Information Processing Systems*, **34**, 214-233.
- [3] Wang, Z.W., She, Q. and Ward, T.E. (2021) Generative Adversarial Networks in Computer Vision: A Survey and Taxonomy. *ACM Computing Surveys (CSUR)*, **54**, 1-38. <https://doi.org/10.1145/3439723>
- [4] Wang, L., Ho, Y.-S. and Yoon, K.-J. (2019) Event-Based High Dynamic Range Image and Very High Frame Rate Video Generation Using Conditional Generative Adversarial Networks. *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, Long Beach, 16-20 June 2019, 10081-10090. <https://doi.org/10.1109/CVPR.2019.01032>
- [5] Tov, O., Alaluf, Y., Nitzan, Y., et al. (2021) Designing an Encoder for Stylegan Image Manipulation. *ACM Transactions on Graphics*, **40**, 1-14. <https://doi.org/10.1145/3450626.3459838>
- [6] Liu, H.Y., Wan, Z.Y., Huang, W., et al. (2021) Pd-gan: Probabilistic Diverse Gan for Image Inpainting. *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, Nashville, 19-25 June 2021, 9371-9381. <https://doi.org/10.1109/CVPR46437.2021.00925>
- [7] Jiang, L.M., Dai, B., Wu, W. and Loy, C.C. (2021) Deceive D: Adaptive Pseudo Augmentation for GAN Training with Limited Data. *35th Conference on Neural Information Processing Systems (NeurIPS 2021)*, 6-14 December 2021, 153-171.
- [8] Xiao, T., Xu, Y., Yang, K., et al. (2015) The Application of Two-Level Attention Models in Deep Convolutional Neural Network for Fine-Grained Image Classification. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, Boston, 7-12 June 2015, 842-850.
- [9] Jabbar, A., Li, X. and Omar, B. (2021) A survey on Generative Adversarial Networks: Variants, Applications, and Training. *ACM Computing Surveys*, **54**, 1-49. <https://doi.org/10.1145/3463475>
- [10] Arjovsky, M., Chintala, S. and Bottou, L. (2017) Wasserstein Generative Adversarial Networks. *International Conference on Machine Learning*, Volume 70, 214-223.
- [11] Nowozin, S., Cseke, B. and Tomioka, R. (2016) F-GAN: Training Generative Neural Samplers Using Variational Divergence Minimization. *30th Conference on Neural Information Processing Systems (NIPS 2016)*, Barcelona, 5-10 December 2016, 432-654.
- [12] Donahue, J. and Simonyan, K. (2019) Large Scale Adversarial Representation Learning. *33rd Conference on Neural Information Processing Systems (NeurIPS 2019)*, Vancouver, 8-14 December 2019, 12-65.
- [13] Karras, T., Aila, T., Laine, S. and Lehtinen, J. (2018) Progressive Growing of GANs for Improved Quality, Stability, and Variation. *International Conference on Learning Representations*, Vancouver, 30 April-3 May 2018, 254-337.
- [14] Karras, T., Laine, S. and Aila, T. (2019) A Style-Based Generator Architecture for Generative Adversarial Networks. *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, Long Beach, 16-20 June 2019, 4401-4410. <https://doi.org/10.1109/CVPR.2019.00453>
- [15] Zhang, H., Goodfellow, I., Metaxas, D. and Odena, A. (2019) Self-Attention Generative Adversarial Networks. *Proceedings of the International Conference on Machine Learning*, Long Beach, 9-15 June 2019, 7354-7363.
- [16] Karras, T., Laine, S., Aittala, M., et al. (2020) Analyzing and Improving the Image Quality of Stylegan. *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, Seattle, 13-19 June 2020, 8110-8119. <https://doi.org/10.1109/CVPR42600.2020.00813>
- [17] Zhao, H.Y., Liu, Z.J., Lin, J., et al. (2020) Differentiable Augmentation for Data Efficient Gan Training. *34th Conference on Neural Information Processing Systems (NeurIPS 2020)*, Vancouver, 568-572.
- [18] Yang, C.Y., Shen, Y.J., Xu, Y.H. and Zhou, B.L. (2021) Data-Efficient Instance Generation from Instance Discrimination. *35th Conference on Neural Information Processing Systems (NeurIPS 2021)* 6-14 December 2021, 9378-9390.

- [19] Radford, A., Metz, L. and Chintala, S. (2015) Unsupervised Representation Learning with Deep Convolutional Generative Adversarial Networks. *4th International Conference on Learning Representations, ICLR 2016*, San Juan, 2-4 May 2016, 689-956.
- [20] Tseng, H.-Y., Jiang, L., Liu, C., *et al.* (2021) Regularizing Generative Adversarial Networks under Limited Data. *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, Nashville, 19-25 June 2021, 7921-7931. <https://doi.org/10.1109/CVPR46437.2021.00783>
- [21] Zhang, H., Xu, T., Li, H., *et al.* (2017) StackGAN: Text to Photorealistic Image Synthesis with Stacked Generative Adversarial Networks. *IEEE International Conference on Computer Vision, ICCV 2017*, Venice, 22-29 October 2017, 5907-5915. <https://doi.org/10.1109/ICCV.2017.629>
- [22] Liu, B.C., Zhu, Y.Z., Song, K.P. and Elgammal, A. (2021) Towards Faster and Stabilized GAN Training for High-Fidelity Few-Shot Image Synthesis. *International Conference on Learning Representations, ICLR 2021*, Vienna, 4 May 2021, 135-983.