【ICML2019】IanGoodfellow自注意力GAN的代码与PPT

专知 今天

【导读】谷歌研究人员Han Zhang和Ian Goodfellow在ICML2019提出的”自注意力生成对抗网络”（SAGAN），将自注意力机制引入到卷积GAN中，作为卷积的补充，在ImageNet多类别图像合成任务中取得了最优的结果。最近在ICML2019作者公开PPT和论文代码。

【SAGAN的PPT下载】

> **后台回复“SAGAN” 就可以获取《自注意力生成对抗网络》PPT下载链接~**

论文PPT

在这篇论文中，我们提出自注意力生成对抗网络（ Self-Attention Generative Adversarial Network ，SAGAN）。SAGAN允许对图像生成任务进行注意力驱动、长相关性的建模。传统的卷积GAN生成的高分辨率细节仅作为在低分辨率特征图上的空间局部点的函数。在SAGAN中，可以使用来自所有特征位置的线索来生成细节。此外，鉴别器可以检查图像的远端部分的高度详细的特征彼此一致。此外，最近的研究表明，生成器条件会影响GAN的性能。利用这些发现，我们将谱归一化到GAN生成器中，并发现这改进了训练动态。我们提出的SAGAN达到了state-of-the-art的结果，将Inception score从当前最高的36.8提高到52.52，并且在具有挑战性的ImageNet数据集上将Frechet Inception distance从27.62降低到18.65。注意力层的可视化表明，生成器利用与对象形状相对应的邻域，而不是固定形状的局部区域。

What did we do?

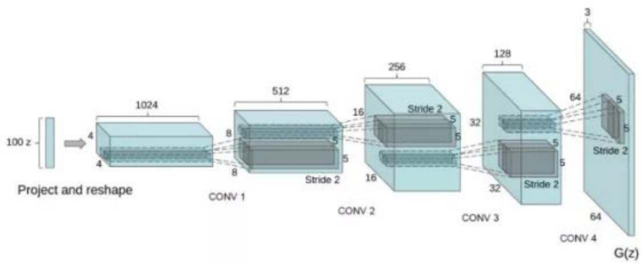
- Add Self-Attention blocks to Generator and Discriminator
- Spectral Normalization (Miyato et al., ICLR, 2018) in both Generator and Discriminator
- Different learning rate for Generator and Discriminator (TTUR: Heusel et al., NIPS, 2017)

What were the results?

Model	Inception Score	Intra-FID	FID
AC-GAN	28.5	260	\
SNGAN-projection	36.8	92.4	27.62
Our SAGAN	52.52	83.7	18.65

Comparison of SAGAN and AC-GAN (A. Odena et al., ICLR, 2017), SNGAN-projection (T. Miyato et al., ICLR, 2018) on ImageNet

What's wrong with convolutions?



DCGAN (Radford et al, ICLR, 2016)

Improved GAN
(Salimans et al, NIPS, 2016)

+ Excel at synthesizing image classes with few structural constraints
- Fail to capture geometric or structural patterns

What is Self-Attention?

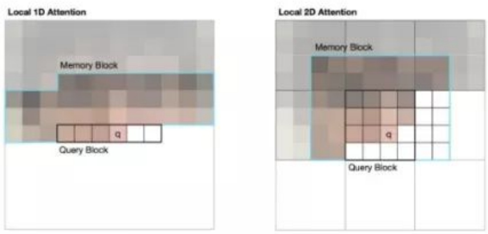


Image Transformer
(Parmar et al, ICML, 2018)

Non-local Neural Networks
(Wang et al, CVPR, 2018)

+ Models long-range dependencies more efficiently

Self-Attention in GANs

$$y_i = \frac{1}{C(x)} \sum_j f(x_i, x_j)\, g(x_j)$$

Self-Attention in GANs

$$\underbrace{\overbrace{y_i}^{\text{single output location}}}_{} := \frac{1}{\underbrace{C(x)}_{\sum_j f(x_i, x_j)\ \text{Normalizing constant}}} \sum_j \overbrace{f(x_i, x_j)}^{\text{"relevance" of } x_j \text{ to } x_i} \underbrace{g(x_j)}_{\text{learned function of a location}}$$

Self-Attention in GANs

$$y_i := \frac{1}{C(x)} \sum_j f(x_i, x_j)\, g(x_j)$$

$$f(x_i, x_j) := \exp(\langle \theta(x_i) | \phi(x_j) \rangle)$$

$$\Rightarrow \frac{1}{C(x)} \sum_j f(x_i, x_j) \text{ is a "softmax"}$$

$g(x_j)$ is just an embedding lookup

代码实现

TensorFlow实现

https://github.com/brain-research/self-attention-gan

依赖环境

python 3.6

TensorFlow 1.5

数据

Download Imagenet dataset and preprocess the images into tfrecord files as instructed in improved gan. Put the tfrecord files into ./data

训练

The current batch size is 64x4=256. Larger batch size seems to give better performance. But it might need to find new hyperparameters for G&D learning rate. Note: It usually takes several weeks to train one million steps.

CUDA_VISIBLE_DEVICES=0,1,2,3 python train_imagenet.py --generator_type test --discriminator_type test --data_dir ./data

评价

```
CUDA_VISIBLE_DEVICES=4 python eval_imagenet.py --generator_type test --data_dir ./data
```

专知《深度学习:算法到实战》课程全部完成！550+位同学在学习，现在报名，限时优惠！网易云课堂人工智能畅销榜首位！

点击“**阅读原文**”，了解报名**专知《深度学习:算法到实战》**课程

阅读原文