

# 基于 WGAN 的图像修复应用

文/哈文全

## 摘要

近年来深度学习在图像处理、音频处理以及 NLP 领域取得了令人瞩目的成绩,特别在图像处理领域,深度学习已然成为主流方法。本文介绍深度学习方法在图像修复领域的应用,通过使用 WGAN 来生成破损图像中缺失的部分。通过从可用的大批量的数据集中生成对应的概率分布的生成模型,进而通过比较上下文内容和先验损失函数,来近似的获取破损图像中缺失部分的内容从而达到生成指定区域图像内容的目的。

【关键词】图像修复 语义修复 深度学习 WGAN

GAN 对于生成式模型的发展具有非常重要的意义。GAN 的训练过程创新性地两个神经网络的对抗作为训练准则并且可以使用反向传播进行训练,大大改善了生成式模型的训练难度和训练效率。对抗训练方法摒弃了直接对真实数据的复制或平均,增加了生成样本的多性。GAN 在生成样本的实践中,生成的样本易于人类理解。GAN 解决了生成模型的一些问题,但是 GAN 也存在一些缺点。GAN 采用对抗学习的准则,理论上还不能判断模型的收敛性和均衡点的存在性。并且在实际操作过程中,这种平衡性与同步性是很难把握的,导致了训练过程的不稳定性。另外,GAN 生成的样本虽然存在多样性,但是存在崩溃模式现象,可能生成多样性,但是站在人们的角度看差别不大的样本。而本文旨在立足于 GAN 的基础上,使用 Wasserstein-GAN (采用 Earth-Mover (EM) 距离代替 JS 散度来度量真实样本和生成样本分布之间的距离)来弥补原始 GAN 的训练不稳定以及易崩溃导致的多样性不足问题。

## 1 把图像解析成概率分布中的样本点

在本文中我们主要关注两种类型的信息:



图 1

环境信息和知觉信息。我们从一种概率密度分布入手:正态分布 (PDF)。图像和统计学之间的关键联系在于,我们可以把图像解析成一个高维概率分布中的样本点。当我们知道一些值并想修复所有缺失值时,我们就把它当作了一个最大化问题来讨论,修复得到的就是可能性最大的图像。如果直接观察正态分布中那些样本点,使用 GAN 来得到最适合的概率统计模型并且用它来拟合数据就可以了。

## 2 WGAN

本文所用方法是在 DCGAN 的基础上使用 Wasserstein 距离来规避 GAN 的缺点,Wasserstein 的距离又叫做 Earth-Mover (EM) 距离,定义如下:

$$W(P_r, P_g) = \inf_{\gamma \sim \prod(P_r, P_g)} \mathbb{E}_{(x, y) \sim \gamma} [\|x - y\|]$$

W 距离相比 KL 散度、JS 散度的优越性在于,即使两个分布没有重叠,W 距离任然能够反映它们的远近。而其优点可以概括如下:

- (1) 彻底解决 GAN 训练不稳定的问题;
  - (2) 基本解决 collapse mode 多样性的问题,确保了生成样本的多样性;
  - (3) 训练过程中终于有一个像交叉熵、准确率这样的数值来指示训练的进程,这个数值越小代表 GAN 训练得越好,代表生成器产生的图像质量越高;
  - (4) 以上一切好处不需要精心设计的网络架构,最简单的多层全连接网络就可以做到。
- 总结一下, WGAN 彻底解决了训练不稳定性问题,同时基本解决了崩溃模式现象。

## 3 在输入图像上添加Mask

首先定义一个指定大小的二进制掩码,用来给输入图像添加遮盖,作为破损图像以供生成模型使用。

```
1.self.mask = tf.placeholder(tf.float32, self
```

```
image_shape, name='mask')
```

## 4 补全图像

语境损失:为了得到和输入图像相同的上下文,需要确保  $y$  已知像素对应位置的  $G(z)$  尽可能相似。所以当  $G(z)$  的输出和  $y$  已知位置图像不相似的时候,需要对  $G(z)$  进行惩罚。为此,我们用  $G(z)$  减去  $y$  中对应位置的像素,然后得到它们不相似的程度:  $\mathcal{L}_{\text{contextual}}(Z) = \|M \odot G(Z) - M \odot y\|$ 。感知损失:为了重建一个看起来真实的图像,需要确保判别器判定图像看起来是真实的。  $\mathcal{L}_{\text{perceptual}}(Z) = \log(1 - D(G(Z)))$ 。最后,将语境损失和感知损失组合起来,就可以找到  $\hat{z}$  了。  $\mathcal{L}(Z) = \mathcal{L}_{\text{contextual}}(Z) + \lambda \mathcal{L}_{\text{perceptual}}(Z)$

$$\hat{z} = \arg\min_Z \mathcal{L}(Z)$$

我们通过对梯度  $\nabla_Z \mathcal{L}(Z)$  进行梯度下降,可以迭代地求出  $\arg\min_Z \mathcal{L}(Z)$ 。最后,对比真实图片,添加二进制掩码后的图片,以及使用 WGAN 生成图像修复的图片,如图 1。

## 参考文献

- [1] Ian J. Goodfellow, Jean Pouget-Abadie, Mehdi Mirza. Generative Adversarial Networks: Cornell University Library, 2014.
- [2] Martin Arjovsky, Soumith Chintala, Léon Bottou. Wasserstein GAN: Machine Learning (stat.ML), 2017.
- [3] Raymond A. Yeh, \*, Chen Chen \*: Semantic Image Inpainting with Deep Generative Models : CV, 2017.

## 作者单位

上海交通大学 上海市 200030