

CVPR2020之多码先验GAN：预训练模型如何使用？

原创 bryant8 机器学习与生成对抗网络 3月23日

点击上方“机器学习与生成对抗网络”，关注“星标”

获取有趣、好玩的前沿干货！

今天分享一篇港中文周博磊老师组的工作。



Image Processing Using Multi-Code GAN Prior

<https://arxiv.xilesou.top/pdf/1912.07116.pdf>

Image Processing Using Multi-Code GAN Prior

Jinjin Gu^{1,2}, Yujun Shen², Bolei Zhou²

¹The Chinese University of Hong Kong, Shenzhen ²The Chinese University of Hong Kong

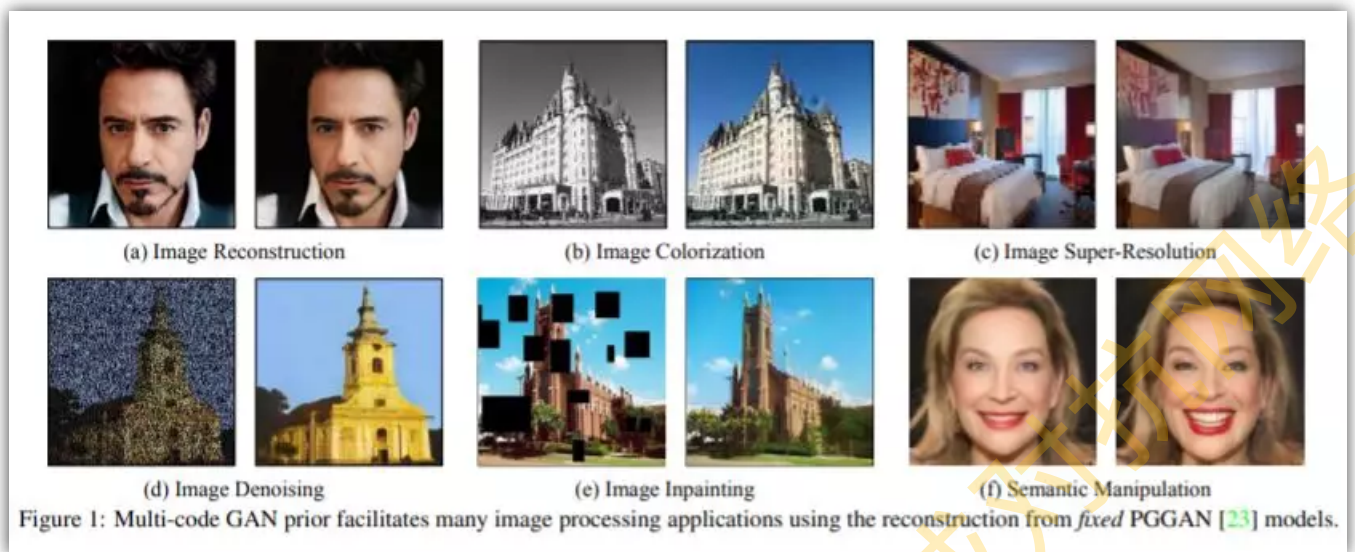
jinjingu@link.cuhk.edu.cn, {syll6, bzhou}@ie.cuhk.edu.hk

尽管生成对抗网络（GAN）在图像生成里取得了巨大成功，但是将训练好的GAN模型应用到真实的图像处理任务仍然具有挑战性。无监督的GAN中，生成器通常将潜在空间 Z 映射到图像空间 $image$ ，也就是没有地方提供去采纳一个真实图像作为输入，这导致训练好的模型无法进行很好地进行后续图像处理。

为了让训练好的GAN能用于处理图像，现有方法尝试以重新反向传播（寻找合适的 Z ）或者添加一个额外的编码器`encoder`将图像映射到潜在空间。但多数情况下两者的重建并不理想。

在这项工作中提出一种新的逆映射（ $image \rightarrow Z$ ）方法，将训练好的GAN作为一个有效的先验去处理多种图像处理任务。具体而言，给定需要做逆映射的GAN模型，使用多个潜码 Z 利用该GAN模型生成多个特征图（映射到生成器的某个中间层），然后计算它们的重要性系数最终组合并生成目标图像。这种多参数化潜码 Z 训练的方式可以显著提高图像的

重建质量。而高质量的图像重建可以使得训练好的GAN作为一种先验应用于许多真实场景下的应用，例如：图像上色、超分辨率、图像修复、图像语义物体操作编辑等等。同时论文还进一步分析了训练好的模型中每一层的表征属性，去解释每层所能表示的知识。



1、GAN逆映射（GAN inversion）

A well-trained generator $G(\cdot)$ of GAN can synthesize high-quality images by sampling codes from the latent space \mathcal{Z} . Given a target image \mathbf{x} , the GAN inversion task aims at reversing the generation process by finding the adequate code to recover \mathbf{x} . It can be formulated as

$$\mathbf{z}^* = \arg \min_{\mathbf{z} \in \mathcal{Z}} \mathcal{L}(G(\mathbf{z}), \mathbf{x}), \quad (1)$$

where $\mathcal{L}(\cdot, \cdot)$ denotes the objective function.

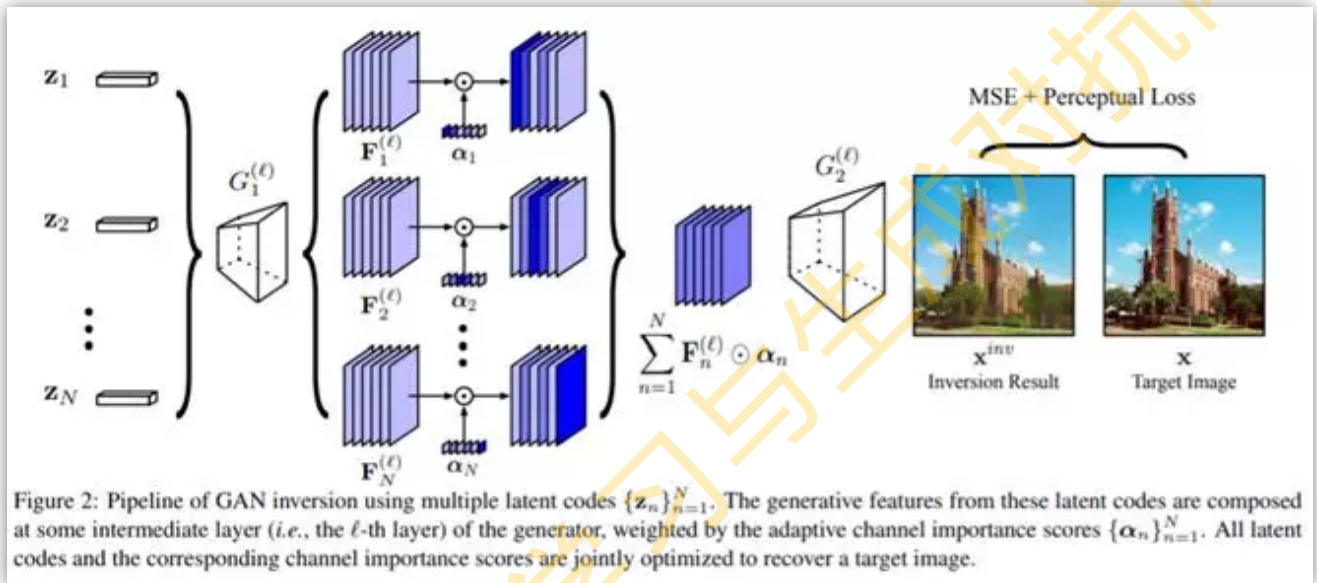
一个无监督、训练好的GAN可以通过从潜在空间 \mathcal{Z} 中采样然后合成高质量的图像，也就是 $\mathcal{Z} \rightarrow \text{image}$ 。而所谓的GAN逆映射指的是，找到一个合适 \mathcal{Z} 去恢复目标图像，也就是 $\text{image} \rightarrow \mathcal{Z}$ （ \mathcal{Z} 此时是一个待优化的参数）。

2、特征组合

作者认为单码（single latent code）去恢复重建目标图像的所有细节so hard，所以使用多码，期待它们可以更全面的去重建更细致的部分（将复杂任务细化、分而治之）。本文引进的所谓多码（multiple latent codes），如何将它们结合在一起优化是一个关键问题。

we propose to combine the latent codes by composing their intermediate feature maps. More concretely, the generator $G(\cdot)$ is divided into two sub-networks, *i.e.*, $G_1^{(\ell)}(\cdot)$ and $G_2^{(\ell)}(\cdot)$. Here, ℓ is the index of the intermediate layer to perform feature composition. With such a separation, for any \mathbf{z}_n , we can extract the corresponding spatial feature $\mathbf{F}_n^{(\ell)} = G_1^{(\ell)}(\mathbf{z}_n)$ for further composition.

作者的方法是将生成器“分为”两部分，G1和G2，L是中间某层的特征输出，提出“自适应通道重要性”（**Adaptive Channel Importance**）去做融合后重建。



$$\mathbf{F}_n^{(\ell)} = G_1^{(\ell)}(\mathbf{z}_n)$$

$$\mathbf{x}^{inv} = G_2^{(\ell)}\left(\sum_{n=1}^N \mathbf{F}_n^{(\ell)} \odot \alpha_n\right),$$

3、优化目标

最终整个框架是求解Z和 α

$$\{\mathbf{z}_n^*\}_{n=1}^N, \{\alpha_n^*\}_{n=1}^N = \arg \min_{\{\mathbf{z}_n\}_{n=1}^N, \{\alpha_n\}_{n=1}^N} \mathcal{L}(\mathbf{x}^{inv}, \mathbf{x}). \quad (4)$$

重建损失使用的是MSE和VGG的感知损失距离：

$$\mathcal{L}(\mathbf{x}_1, \mathbf{x}_2) = \|\mathbf{x}_1 - \mathbf{x}_2\|_2^2 + \|\phi(\mathbf{x}_1), \phi(\mathbf{x}_2)\|_1, \quad (5)$$

4、图像处理任务的应用

对于图像上色

- For image colorization task, with a grayscale image I_{gray} as the input, we expect the inversion result to have the same gray channel as I_{gray} with

$$\mathcal{L}_{color} = \mathcal{L}(\text{gray}(\mathbf{x}^{inv}), I_{gray}), \quad (6)$$

where $\text{gray}(\cdot)$ stands for the operation to take the gray channel of an image.



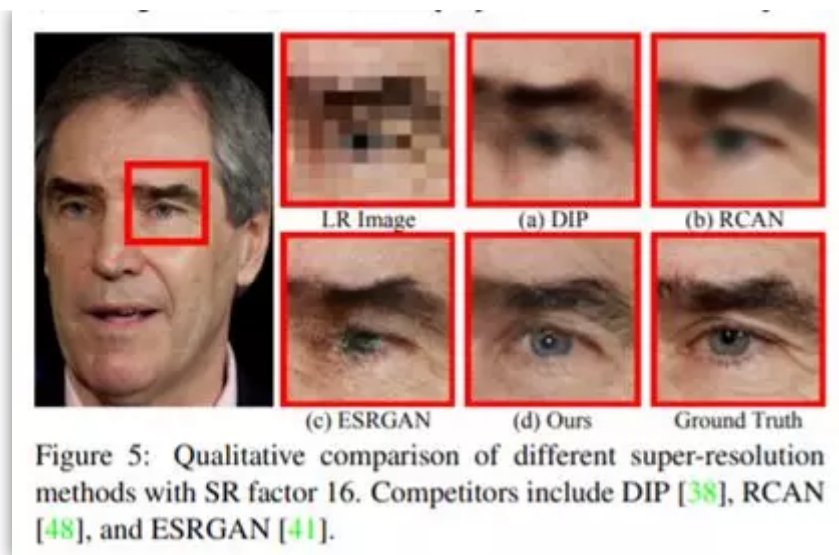
Figure 4: Qualitative comparison of different colorization methods, including (a) inversion by optimizing feature maps [3], (b) DIP [38], (c) Zhang *et al.* [46], and (d) our proposed multi-code GAN prior.

对于图像超分

- For image super-resolution task, with a low-resolution image I_{LR} as the input, we downsample the inversion result to approximate I_{LR} with

$$\mathcal{L}_{SR} = \mathcal{L}(\text{down}(\mathbf{x}^{inv}), I_{LR}), \quad (7)$$

where $\text{down}(\cdot)$ stands for the downsampling operation.



对于图像修复

- For image inpainting task, with an intact image I_{ori} and a binary mask m indicating known pixels, we only reconstruct the incompact parts and let the GAN model fill in the missing pixels automatically with

$$\mathcal{L}_{inp} = \mathcal{L}(x^{inv} \circ m, I_{ori} \circ m), \quad (8)$$

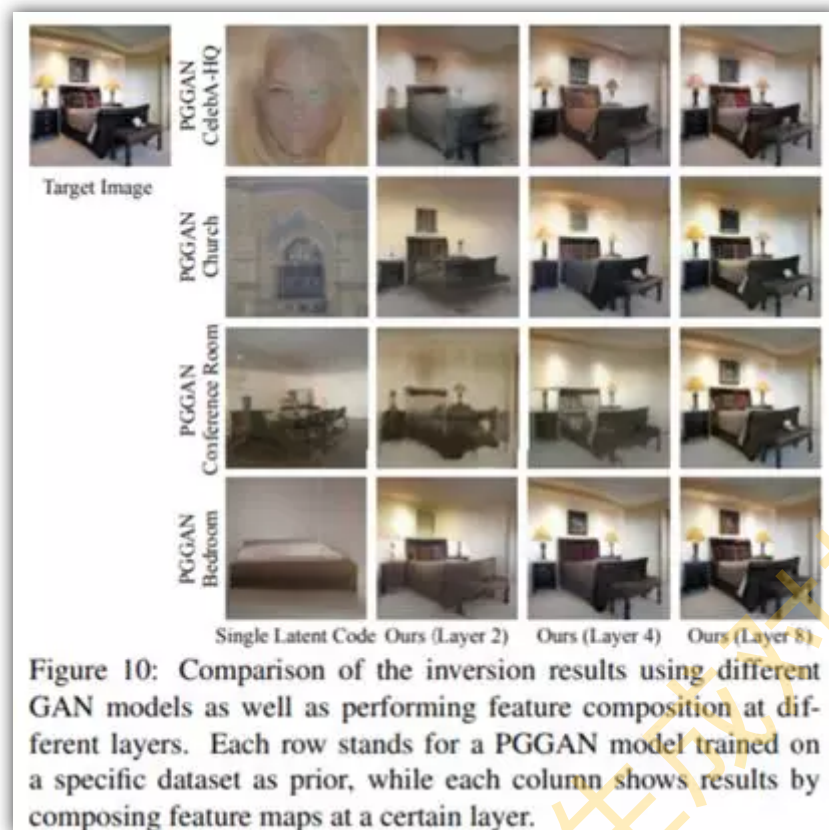
where \circ denotes the element-wise product.



5、GAN里的知识表征

分别使用face、church、conference room、bedroom四个数据集预训练PGGAN，并以此作为先验。当使用single latent code去重建时，很难“覆盖先验”，出来的结

果会呈现原训练集图像内容，而无法重建到目标图像。



结论

本文提出了一种新颖的GAN逆映射方法。基于预训练的无监督GAN，使用多码（multiple latent codes）去重建目标图像的方法。实验表明该方法可以有效利用预训练好的GAN进行多种图像处理任务。

公众号近期荐读：

GAN整整6年了！是时候要来捋捋了！

新手指南综述 | GAN模型太多，不知道选哪个好？

数百篇GAN论文已下载好！搭配一份生成对抗网络最新综述！

CVPR2020之MSG-GAN：简单有效的SOTA

CVPR2020之姿势变换GAN：图像里谁都会劈叉？

3月已半，GAN最近都在整些啥？

有点夸张、有点扭曲！速览这些GAN如何夸张漫画化人脸！

见微知细之超分辨率GAN！附70多篇论文下载！

天降斯雨，于我却无！GAN用于去雨如何？

脸部转正！GAN能否让侧颜杀手、小猪佩奇真容无处遁形？

容颜渐失！GAN来预测？

强数据所难！SSL（半监督学习）结合GAN如何？

弱水三千，只取你标！AL（主动学习）结合GAN如何？

异常检测，GAN如何gan？

虚拟换衣！速览这几篇最新论文咋做的！

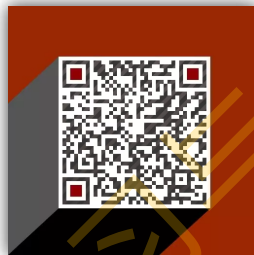
脸部妆容迁移！速览几篇用GAN来做的论文

【1】GAN在医学图像上的生成，今如何？

01-GAN公式简明原理之铁甲小宝篇

GAN&CV交流群，无论小白还是大佬，诚挚邀您加入！

一起讨论交流！长按备注【进群】加入：



更多分享、长按关注本公众号：

