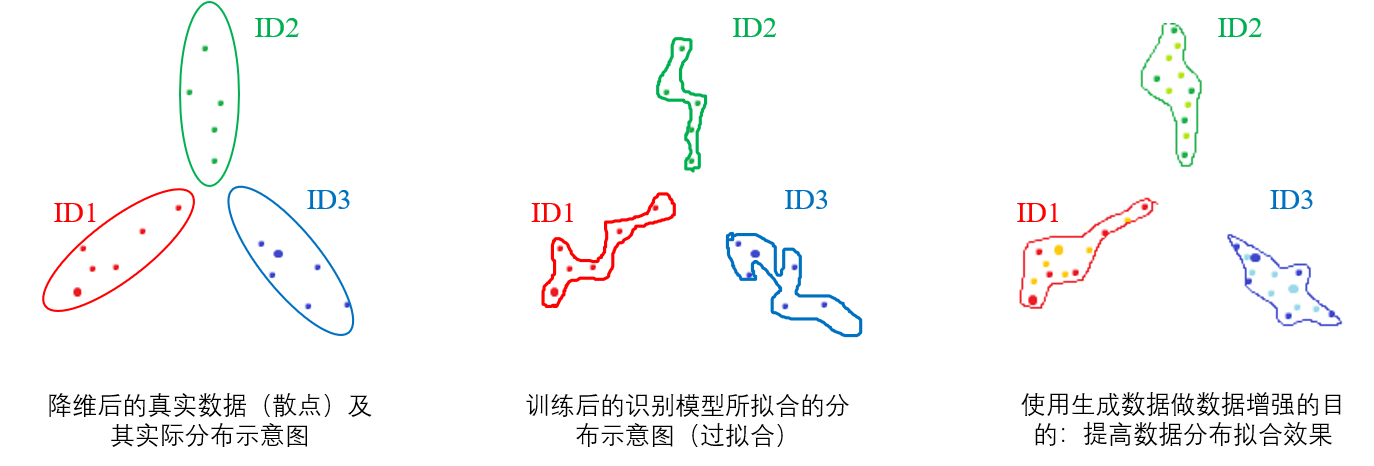
为什么行人生成能为行人重识别做数据增强

*初学者很难避免理解偏差，如果有什么说的不对的地方，还请谅解，我接受且欢迎您的斧正。*

我第一次正视这个问题是在半个月以前。交流群里有人说“GAN做数据增强不可行，要是生成模型能补出你想要的分布，那同规模和架构的识别（分类）模型理应也能识别出这样的分布，所以更应该去改进识别模型而不是生成模型。”

首先我是知道的，深度生成模型（GAN、VAE……）的生成，并不能做到无中生有，只能做到有中插值。从这个角度来说，使用深度生成模型做数据增强实际上是一个有待商榷的做法，它除了有可能扩大类别偏差外，对问题的优化也存在一个明显的极限——随着数据分布的逐渐稠密，识别模型学会的分布越来越准确，这种“分布内插值”的增强方法将越来越失去意义。

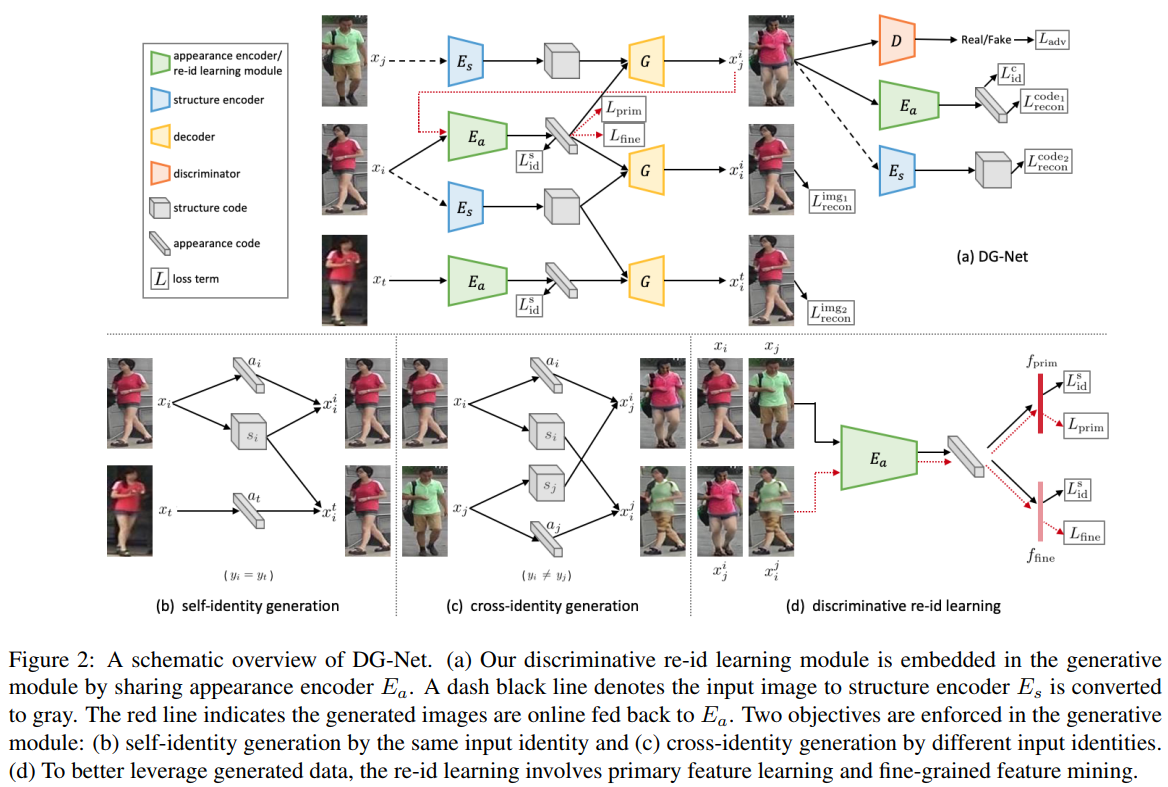
我的看法是，尽管存在这样的一个极限，但生成模型的插值功能正是其作为数据增强的重点所在。生成模型做数据增强的特点在于其生成数据分布贴近真实分布的，能将稀疏数据集数据分布中间的间隔补上，如图所示。



而有了这些生成图片，在优化识别模型的时候就会更容易拟合真实数据分布，作为一种离散数据的原数据随着图片的增多而分布地更加稠密，这使得无论是过拟合还是欠拟合都更容易避免。我想这就是生成模型能够在一定程度上做数据增强的原因，而使用pose作为指导信息的一个重要原因可能正是为了避免类别偏差随着无条件GAN的学习而增大（使新数据并不过紧地挨着原有数据）。

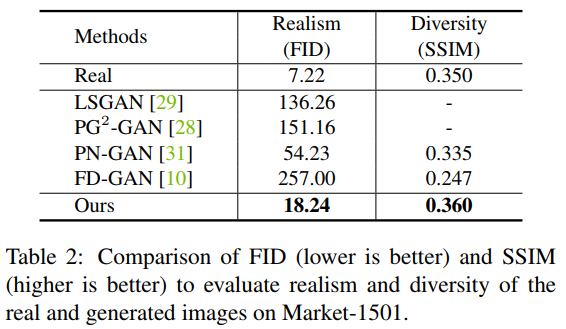
周六去自所听SFFAI分享的时候，师兄师姐的图像翻译工作中都提到了使用知识蒸馏来优化性能的技巧，我在请教马聪学长后有了一些粗浅的想法——我觉得这个方法能work在某种程度上和生成模型数据增强能work的道理是相通的，他们都是一种对标签数据分布的优化（分类模型🡪数字标签 | 生成模型🡪图像标签），目的都是改善模型的拟合情况。而质量够高的生成模型做数据增强或许也可以这样一层层地往外拓展（迭代）来逼近真实分布。

本周其实还看了看今年的CVPR oral论文《Joint Discriminative and Generative Learning for Person Re-identification》，尚且只细读了首尾，但是其效果却十分令人震撼。他们的方法更像是“数据增强”而没有那么注重“图像生成”，但是对行人重识别的提升确实巨大的。其网络结构如下图所示。

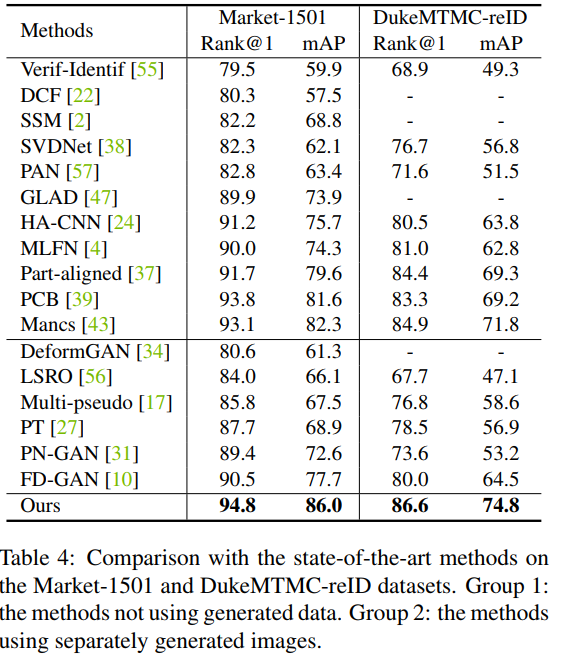


我觉得这个提升可能要从两方面来说。

一方面，他们降低了需求，只需要更换表层特征（服饰、颜色），使得生成图片更加容易，图像质量更高更真实，所以其数据的效果更好（见Table2）。

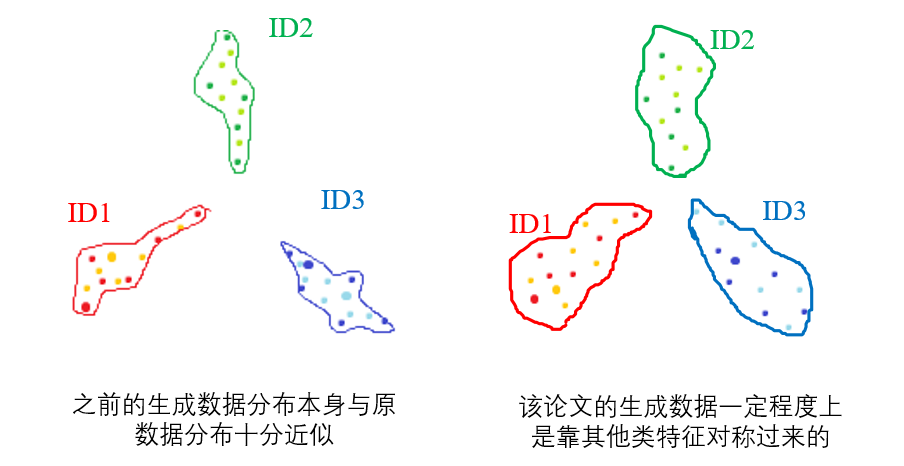


另一方，他们是在将其他ID类中的一些特征信息“对称”到另一ID类里，增加了表层特征的多样性，从而引导识别模型关注表层颜色之外的信息，进而在一定程度上提高了泛化性能，提升了效果。（见Table4）



问题分析：

为什么原来的增强方法无法对ReID带来这么大的提升呢？我猜测一方面在于图像质量的问题（说起来简单但很关键），另一方面可能是由于生成数据本身也是对原数据分布轻微过拟合的，所以其插值并未突破原有数据分布模式，只能让数据类内分布更加平滑而很难增强泛化性。而这篇论文的目的更加“单纯”，就是数据增强，可以稳定地增加类内多样性，突破了原数据的固有分布模式，所以能增强泛化性能提高识别准确率。



这就是我这周的思考。

冯汉森

2019.5.26