

**生物特征识别课程报告**

**报告题目 噪声数据下的人脸属性编辑**

**主讲教师 孙哲南**

**学生姓名 王宇 学号 201928014628024**

**学位类别 工程学硕**

**学科专业 模式识别与人工智能**

**研究所（院系） 中国科学院自动化研究所**

**报告日期 2020.5.16**

**中国科学院大学人工智能技术学院制**

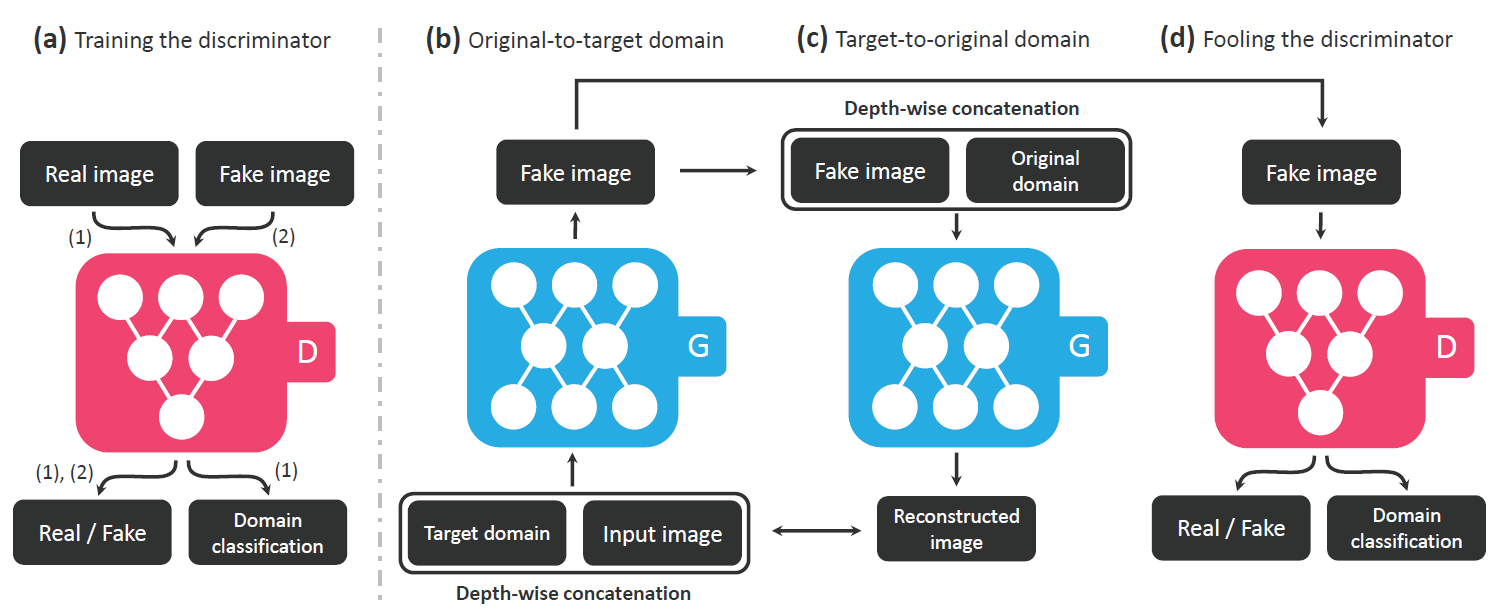
**噪声数据下的人脸属性编辑**

1. **问题背景**

在计算机视觉和机器学习中，生成模型得到了广泛的研究与发展，用以生成或重现与真实数据难于区分的样本。其中，通过对抗性训练学习生成性分布的生成对抗性网络(GANs)，由于其不需要显式密度估计就能学习任何数据分布，因此成为了使用最为广泛的模型。而所谓的人脸属性编辑任务实际上也是通过机器学习的方法对人脸的目标属性进行修改，因此也可以通过使用GANs网络来完成人脸属性编辑的任务，其中最常用的人脸属性编辑模型有AttGAN，StarGAN。通常在训练人脸属性编辑模型时需要大量的监督数据，而在制作数据集时需要大量的人力成本，且数据标记难免会出现标记错误的情况，因此我们需要提高模型的抗噪声能力，然而现有的模型在处理噪声数据时，效果并不是很好。在此基础之上，我们融合Stargan与rAC-GAN模型提出具有抗噪声能力的Rstargan模型。

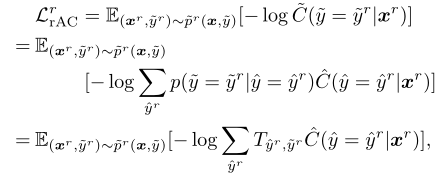
1. **模型方法**

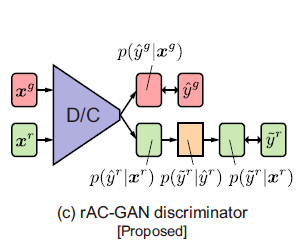
Rstargan主要基于stargan模型进行修改，stargan模型具体如图一示，图a为stargan的判别器，输入为图片，输出为真假以及图片的属性编辑的目标域，图b-d为整个stargan模型，将输入图片和目标生成域喂到生产网络中来合成假图片，然后假图片进入到判别器对判别器进行训练，而生成器本身还有一个对偶重构，通过将生成的假图片和原始图片的属性信息结合起来喂入到生成器G要求能重建出原始输入图片。



**图1 stargan模型框架**

本文在stargan模型的基础之上提出三个修改策略：(1). 将对偶重构修改为约束残差生成把G换成残差生成𝐺(𝑥) = 𝑥 + 𝑛𝑛(𝑥)，nn是神经网络输出。引入残差的l1范数进行限制；(2). 增大分类Loss的权重; (3). 使用rAC-GAN进行改进，命名改进后的模型为rstargan, 通过概率转移矩阵进行微调。改进方法是在判别器算Loss的时候要引入一个Label的马尔科夫转移矩阵T，元素T[i,j]表示把真实类别i变成j的概率，T矩阵在判别器中的具体作用如图2所示，用于对噪声label进行状态转移，判别器loss的数学表达式如公式1所示。

 (1)



**图2 rAC-GAN判别器模型**

1. **实验结果**

(1). 实验细节

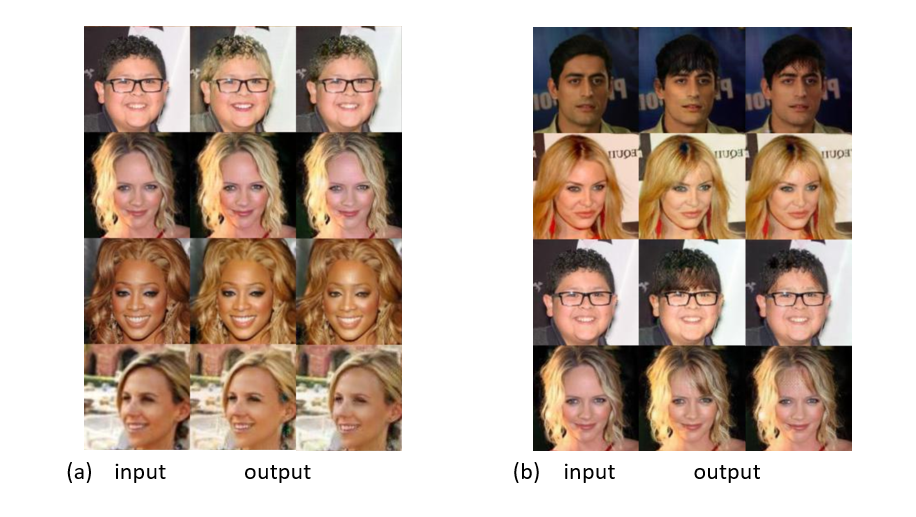
本次实验所使用到的数据集为celebA，环境为pytorch，分类权重lambda\_cls= 1.0, 对抗权重lambda\_adv= 0.3, res-loss权重lambda\_res = 0.2, 并且在训练的12W轮前用stargan的分类损失，12W轮后使用策略3进行fine-tune，loss的成分变为80%的rACGAN，20%的ACGAN，其他网络参数和stargan保持一致，目的是为了和stargan进行公平对比。

(2). 实验一

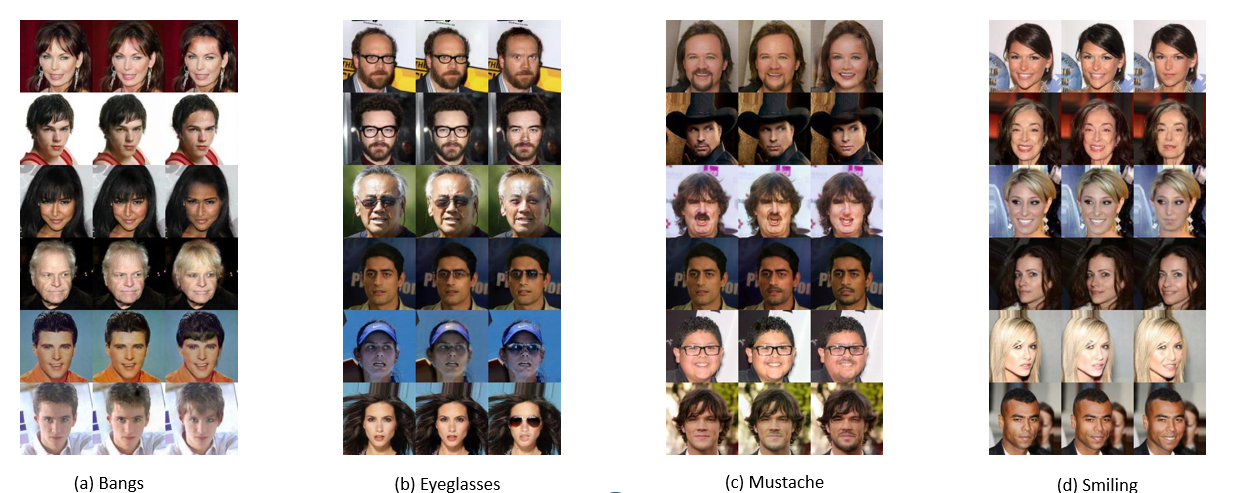
该实验为stargan在存在数据反转时生成图像的对比图。如图3所示，左图是Blond Hair，右图是Bangs，三列分别为原图，正常训练，0.3标注反转训练(即有百分之30的标错数据)。可以观察到当标注存在反转时，目标属性不够清晰，程度偏浅，stargan属性编辑能力变弱。且当存在较高数据反转的情况下，stargan的loss难以得到很好的下降。

(3). 实验二

实验二为rstargan与stargan模型在存在0.4标注反转下的生成图像的对比实验，实验结果如图4所示。第一列为原图，第二列为stargan，第三列为rstargan，观察不同属性下的人脸编辑效果，不难发现stargan在存在数据反转的情况下有时会出现无法对属性进行编辑的情况，而我们的rstargan模型无论是添加属性还是删除属性均能获得较为令人满意的结果。



**图3 实验一**



**图4 实验二**

(3). 实验三

该实验为rstargan的不同策略对比，实验结果如图5所示，数据反转比例为0.4，修改策略为本文2中所提及的三条，记StarGAN为Baseline，rStarGAN1仅使用策略1，rStarGAN2使用策略1+2，rStarGAN3使用策略1+2+3。可以观察到在仅使用策略1与策略2时便可以得到较好的人脸属性编辑效果，在使用策略三后属性编辑效果更为明显，抗干扰能力得到提升。

(4). 实验四

该实验为噪声比例测试，实验结果图如图6所示，本次实验编辑的属性为eyeglasses，四列分别为输入，30%噪声，40%噪声，45%噪声下的属性编辑图。可见在噪声为45%的时候，编辑非常困难。但当噪声低于40%可以很好的实现编辑效果，可见stargan模型具有相当高的抗噪声能力。



**图5 实验三**



**图6 实验四**

1. **总结**

对于本文所提出的三种stargan模型改进策略，在光使用策略1与策略2的情况下已经能够较好得到生成图像，使用策略3能够较好得提高模型抗噪声能力，但是所提出的模型仍然存在一些不足，rstargan模型仍然需要人为的去设计噪声的转移矩阵，转移矩阵设计的好坏会直接影响生成图像的好坏。