

Learning Face Age Progression: A Pyramid Architecture of GANs

2018/8/10 王诺

由年轻的面孔图像生成老化面孔图像有两个基本要求，即老化准确和身份特征保持，但在以往的相关工作并没有将这两点做的很好。在本文中，作者提出了一种新的基于生成对抗网络的方法。它同时学习面孔的内在特征和年龄造成的面部变化，并使用时间信息约束，确保生成的面部呈现特定年龄的老化效果，同时保持个性化的特征。此外，为了生成更逼真的面部细节，由金字塔形鉴别器对生成器的输出图像的多规格的高级特征进行评估，以更精细的方式模拟老化效应。本文所提出的方法适用于在姿势，表情，化妆等变化的情况下的各种面部样本，并且实现了非常生动的老化效果。视觉效果和定量评估都表明该方法推进了最先进的技术。

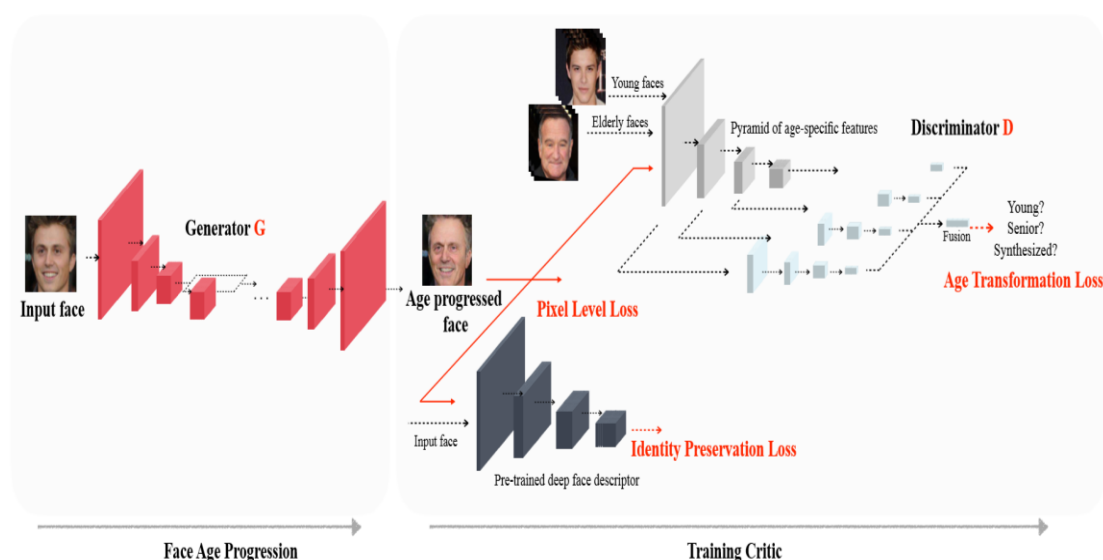


图 1

图 1 展示了本文所提出的算法的整体架构。生成器的结构是典型的编码解码结构，只是将一般的最大值池化（上采样）操作换成了卷积操作(3*3 的卷积核，步长为 2)。作者在鉴别器和生成图像质量评估上做了较多的更新。首先，鉴别器的输入不是生成器输出图像，而是将生成的图像经过训练好的年龄特征提取网络，将年龄辨别网络第 2、4、7、10 层输出的特征图交由鉴别器进行评估。并且考虑到训练过程中训练器的收敛速度总是快于生成器，鉴别器的参数初值取自于部分训练好的 VGG16 网络。同时鉴别器的输出结果不再是一个标量数字，而是一个

12*3 的矩阵 (图中四个分支的输出大小均为 3*3, 合并得到 12*3 总输出)。其次, 为了保持图像的身份特征, 作者使用了一个去除最后一层的预先训练好人脸识别网络, 将生成器的输入图像和输出图像分别输入该网络得到输出的特征图, 使用两特征图之间的平方欧式距离评定输出图像是否良好地保持了身份特征。

$$\begin{aligned}\mathcal{L}_{GAN.D} = & \frac{1}{2} \mathbb{E}_{x \sim P_{old}(x)} [(D_{\omega}(\phi_{age}(x)) - 1)^2] \\ & + \frac{1}{2} \mathbb{E}_{x \sim P_{young}(x)} [D_{\omega}(\phi_{age}(G(x)))^2 + D_{\omega}(\phi_{age}(x))^2]\end{aligned}\quad (3)$$

$$\mathcal{L}_{GAN.G} = \mathbb{E}_{x \sim P_{young}(x)} [(D_{\omega}(\phi_{age}(G(x))) - 1)^2] \quad (4)$$

$$\mathcal{L}_{identity} = \mathbb{E}_{x \in P_{young}(x)} d(\phi_{id}(x), \phi_{id}(G(x))) \quad (5)$$

$$\mathcal{L}_{pixel} = \frac{1}{W \times H \times C} \|G(x) - x\|_2^2 \quad (6)$$

$$\mathcal{L}_G = \lambda_a \mathcal{L}_{GAN.G} + \lambda_p \mathcal{L}_{pixel} + \lambda_i \mathcal{L}_{identity} \quad (7)$$

$$\mathcal{L}_D = \mathcal{L}_{GAN.D} \quad (8)$$

图 2

图 2 展示了损失函数的定义。为了生成含有老化细节和保持身份特征的图片, 生成器的损失函数由鉴别器相关损失、L2 范数、身份信息损失三部分组成。其中 ϕ_{age} 代表图 1 中四个分支网络对输入的特征图的处理, D_{ω} 代表对 ϕ_{age} 产生的 12*3 的矩阵的相关处理。

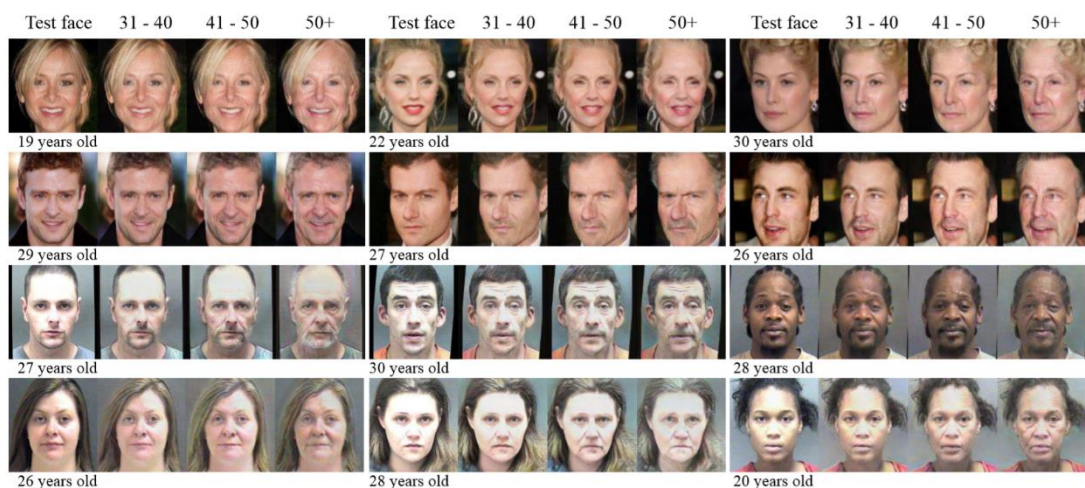


图 3

图 3 展示了使用本文提出的算法对 CACD (前两行) and MORPH (后两行) 数据集中的数据进行处理后的部分结果。(生成不同年龄阶段的图像需要训练不同的网络。)

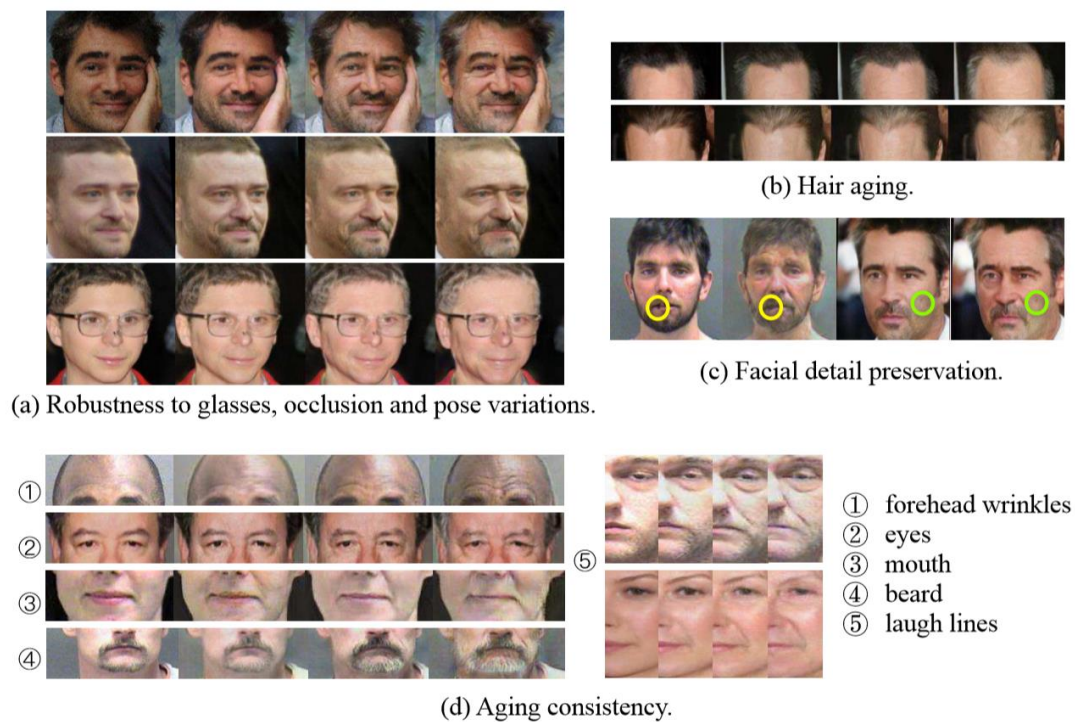


图 4

图 4 展示了使用本文提出算法得到的细节处理的相关结果，可以看出本文提出的算法在额头皱纹、眼睛、嘴巴、胡子和嘴型等细节处理方面由良好的效果。

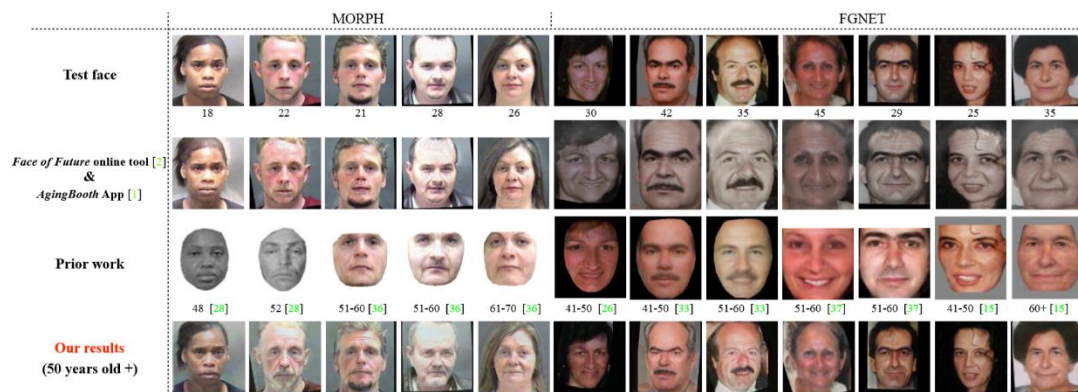


Figure 9. Performance comparison with prior work (zoom in for a better view of the aging details).

图 5

图 5 展示了使用不同算法对 MORPH 和 FGNET 数据集中的数据进行处理后的结果。可以看出本文提出的算法产生的图像效果更好。