

基于生成对抗网络的人脸年龄渐进合成算法

杨晓雨, 王爱侠, 杨钢, 李晶皎
(东北大学 信息科学与工程学院, 辽宁 沈阳 110819)

摘要: 人脸年龄合成(face age synthesis, FAS)的目标是根据源人脸图像合成指定年龄人脸图像,同时保留人脸的个人特征和身份信息. 针对年龄变换时无关特征容易改变和产生伪影鬼影的问题,提出一种基于生成对抗网络的人脸年龄渐进合成算法. 采用基于门控循环单元的年龄编辑模块自适应地过滤或加入特征,并使用属性解耦模块在潜在空间进行对抗学习,通过生成器和判别器的对抗策略保证了真实自然的人脸合成,使用年龄分类约束拟合特定年龄分布,为了保证年龄无关属性的保留,还在生成对抗网络中引入了重建学习. 在跨年名人数据集(cross-age celebrity dataset, CACD)下的实验结果表明,对比其他基于条件生成对抗网络的算法,提出的算法生成的人脸图像伪影失真有所减少,年龄显著性增强,具有较好的年龄准确性和较高的身份一致性.

关键词: 人脸年龄合成;生成对抗网络;属性解耦;潜在空间;门控循环单元;重建学习

中图分类号: TP 391 **文献标志码:** A **文章编号:** 1005-3026(2024)07-0944-09

Progressive Face Age Synthesis Algorithm Based on Generative Adversarial Network

YANG Xiao-yu, WANG Ai-xia, YANG Gang, LI Jing-jiao

(School of Information Science & Engineering, Northeastern University, Shenyang 110819, China. Corresponding author: YANG Xiao-yu, E-mail: sophiayxyqq@163.com)

Abstract: The goal of face age synthesis (FAS) is to synthesize face images of specified ages based on the source face image, while preserving personal characteristics and identity information of the face. To solve the problem that irrelevant features are easy to change and artifact ghosting occurs when age is changed, a progressive face age synthesis algorithm based on generative adversarial network is proposed. The age editing module based on gate recurrent unit is used to filter or add features adaptively, and attribute decoupling module is used for adversarial learning in the latent space. Through the adversarial strategy of generator and discriminator, the real and natural face synthesis is guaranteed. The age classification constraint is used to fit the specific age distribution. In order to preserve age-independent properties, reconstruction learning is also introduced into generative adversarial network. Experimental results on CACD dataset show that, compared with other algorithms based on conditional generative adversarial network, the proposed algorithm has reduced artifacts and distortions, enhanced age significance, and has better age accuracy and higher identity consistency.

Key words: face age synthesis; generative adversarial network; attribute decoupling; latent space; gate recurrent unit; reconstruction learning

面部衰老是一个贯穿人们一生的复杂过程. 人脸年龄变换是指以给定人脸图像为素材,依靠计算机技术和图像处理算法,合成出目标年龄人脸图像的技术. 计算机辅助的人脸年龄合成技术

的应用范围包括社会安全、数字娱乐等诸多领域,有助于警方在执法时寻找和辨认嫌犯,协助弱势群体防范连环罪犯,以及帮助寻找走失人口. 随着人工智能和深度学习的发展,在过去的

二十年中,该领域的研究活动有所增加.但是由于物理衰老过程本身的复杂性和个体间的差异性,以及人脸数据集图像受到环境因素(面部姿态、光照等)干扰,人脸年龄合成是一个充满挑战的课题.

传统的人脸年龄变换方法利用建模的方式模拟面部衰老过程,根据是否采用参数建模,可以大致将这些方法划分为采用参数化模型的物理模型方法^[1-6]和采用非参数化模型的原型(prototype)方法^[7-8].物理模型方法通过建模肌肉、皮肤或头骨的衰老机制来模拟面部衰老,原型方法主要是将每个年龄组的平均脸间的差异作为老化模式.

随着深度学习技术的极大发展,基于深度学习的人脸年龄合成方法取得了令人瞩目的成果.Wang等^[9]提出基于循环神经网络(recurrent neural network, RNN)的循环面部老化模型,首次将深度学习用于人脸年龄变换.此外,其他基于时间结构的深度生成模型^[10-11]也被用于老化人脸图像的合成,但是这类方法大多需要大量连续的序列样本,因此实际应用潜力有限.

近年来,以生成对抗网络GAN(generative adversarial network)^[12]为代表的生成模型在图像翻译任务上取得了卓越的成果^[13-17],许多研究人员尝试将GAN及其变体CGAN(conditional GAN)^[18]应用于无监督人脸老化模型的训练.Antipov等^[19]率先将CGAN用于年龄渐进合成任务.Zhang等^[20]提出了一种条件对抗性自编码器,通过遍历低维流形实现人脸年龄的渐进和回归合成.Wang等^[21]采用年龄估计损失和身份损失对生成图像进行约束,达到良好的人脸老化效果.Yang等^[22-23]设计了一个金字塔结构模型来增强老化细节.Liu等^[24]和Shao等^[25]采用小波包变化模块在频率空间的多个尺度捕捉年龄相关的纹理细节.Li等^[26]设计了一个全局生成器和三个局部生成器,用于捕捉全局和局部特征.Huang等^[27]利用多个生成网络来渐进模拟老化过程.然而上述方法主要是为了提高图像的视觉质量,由潜在表示中属性纠缠引起的视觉伪影和不必要的属性编辑仍然是限制人脸年龄合成性能提升的重要原因.此外,采用多尺度网络或多个GANs也会限制模型的效率,为稳定的训练带来挑战.鉴于此,目前一些研究将重点转向年龄特征的提取和编辑.Li等^[28]提出了一种基于注意力机制的方法,利用注意力机制将图像修改范围限制在与老化有关的区域.Huang等^[29]采用多任务训练的

方式来解决关联身份和年龄特征,从而保留了人脸的个性化模式.

针对上述观察,本文提出了一个基于单GAN模型的人脸年龄渐进合成算法框架.首先使用编码器将人脸特征映射到高维语义空间中,然后对值得关注的特征进行提取和变换,利用年龄编辑模块对人脸特征进行自适应特征编辑,在保留年龄无关信息的同时提高了年龄准确性;在编码器和解码器之间引入了一个属性解耦模块,实现年龄特征和个人特征的充分解耦;在判别器末端引入年龄分类器分支,保证生成图像年龄在指定范围内;在身份确认模块和重建损失共同作用下保障生成人脸图像的身份和背景信息不丢失.

1 算法框架

1.1 整体网络模型

本文提出的人脸年龄渐进合成算法框架如图1所示.主要包括1个编码器 G_{enc} ,2个年龄编辑模块,1个解码器 G_{dec} ,1个属性解耦模块 D_f ,1个判别器 D .

模型的输入为源人脸图像 x_s ,目标年龄段真实人脸图像为 x_t .定义源图像所属的年龄段为 α_s ,目标年龄段为 α_t . x_s 和 x_t 服从的分布分别表示为 $p_{data}(x_s)$ 和 $p_{data}(x_t)$.由于简单的一维向量空间难以表示与年龄有关的特征(胡子、皱纹等),这些特征通常出现在更高级的语义空间中.因此本文采用特征图 $f \in \mathbf{R}^{H \times W \times C}$ 形式表示提取的人脸特征,其中 C 为通道数, H 和 W 分别为特征图的高和宽.用 f_s, f_t 分别表示经过年龄编辑模块变换前、后的人脸特征潜在表示.把经过修改的潜在表示输入到解码器中生成目标年龄人脸图像过程表示为 $G_{dec}(f_t)$,合成人脸需要与源人脸具有相同的身份特征,并且被年龄分类约束认为尽可能符合目标年龄标签的输入.每次生成目标年龄段人脸图像的同时,再次调用解码器从变换前的潜在表示中重建原始人脸图像,将这个过程表示为 $G_{dec}(f_s)$.为确保合成图像的真实性,将合成人脸图像和真实人脸图像同时送入判别器,生成器和判别器在不断的对抗博弈中逐渐使图像趋于逼真的效果.

1.2 各个网络结构设置

1.2.1 编码器-解码器架构

编码器 G_{enc} 采用了连续的5个步长为2的 4×4 卷积核执行下采样,通道数分别为64,128,256,512,1024.并且在每个卷积核操作后均采用批归

一化 (batch normalization, BN), 并采用 Leaky ReLU 激活函数进行激活.

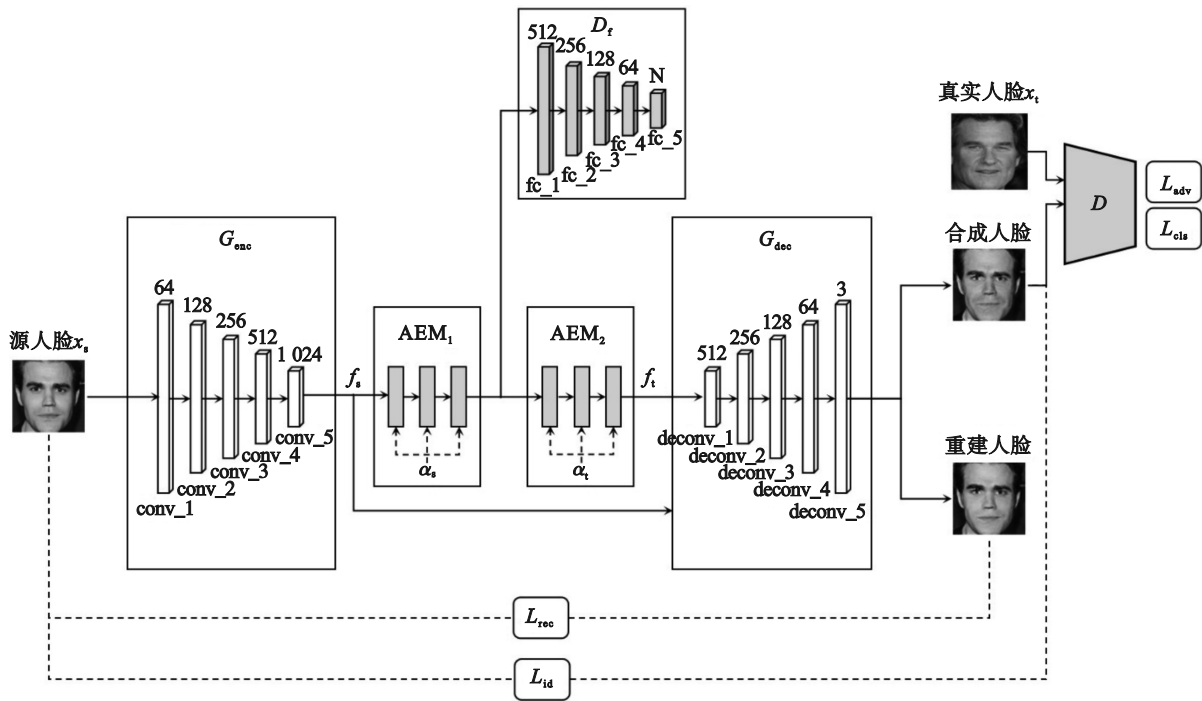


图 1 人脸年龄渐进合成算法框架

Fig. 1 Framework of face progressive synthesis algorithm

解码器 G_{dec} 与编码器 G_{enc} 互为对称结构, 采用 5 个步长为 2 的 4×4 反卷积进行上采样操作, 通道数分别为 512, 256, 128, 64, 3. 在前 4 次反卷积后同样加入 BN 层, 并使用 ReLU 函数激活. 最后一次反卷积不进行标准化操作, 仅采用 tanh 激活函数对合成张量的数值尺度进行归一化, 从而得到生成器经过解码后的输出图像.

1.2.2 属性解耦模块

基于编码器-解码器结构的人脸年龄合成方法实现年龄映射的关键问题是如何对年龄属性和人脸图像的潜在表示之间进行建模. 理想情况下, 希望生成图像中的年龄信息完全取决于输入的年龄属性, 并且加入的年龄属性仅仅影响图像特征中与年龄相关的部分, 但是实际情况中由编码器映射得到的隐藏空间中往往包含着大量未解耦的人脸特征. Fader Networks^[30] 在潜在空间中采取了一种对抗策略来学习属性不变表示. 受此启发, 本文设计了属性解耦模块 D_f . 在编码器和属性解耦模块之间进行对抗博弈, 根据输入的潜在表示判断年龄信息是否存在, 而编码器则要试图欺骗属性解耦模块, 使得输入到属性解耦模块中的潜在表示中不包含年龄相关信息. D_f 由 5 个全连接层构成, 在前 4 个全连接层后均使用 Leaky ReLU 函数进行激活.

1.2.3 年龄编辑模块

与一般的生成网络结构有所不同, 本文在编码器和解码器之间的潜在空间中加入了设计年龄编辑模块 (age editing module, AEM). 模块采用了类似门控循环单元的结构, 通过门控制自适应地根据输入的年龄属性调整潜在表示特征图的信息. 每个年龄编辑模块 $AEM_i (i=1, 2)$ 由串行连接的卷积门控循环单元组成, 卷积门控循环单元内部结构如图 2 所示.

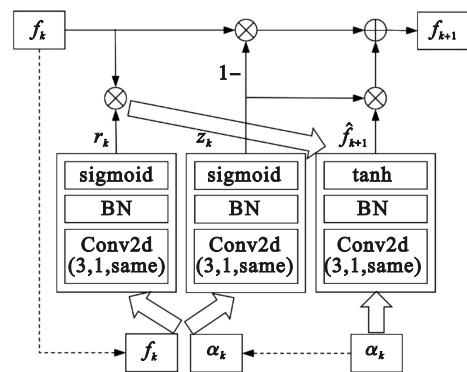


图 2 卷积门控循环单元内部结构图

Fig. 2 Diagram of convolution gate recurrent unit internal structure

设输入卷积门控循环单元第 k 个状态的特征图为 f_k , 年龄属性为 α_k , 则输出特征图可以用式(1)~式(4)表示:

$$r_k = \sigma(W_r * [f_k, \alpha_k]), \quad (1)$$

$$z_k = \sigma(W_z * [f_k, \alpha_k]), \quad (2)$$

$$\hat{f}_{k+1} = \tanh(W_h * [r_k \circ f_k, \alpha_k]), \quad (3)$$

$$f_{k+1} = (1 - z_k) \circ f_k + z_k \circ \hat{f}_{k+1}. \quad (4)$$

其中: $*$ 表示卷积核为3,步长为操作1的卷积操作; \circ 表示逐元素相乘; $[\]$ 表示张量在通道维度的拼接操作; $\sigma(\cdot)$ 表示sigmoid函数.

1.2.4 判别器架构

判别器 D 的输入为合成人脸图像和目标年龄段真实人脸图像,它在网络中的作用是促使生成器生成被判别器预测为真的目标年龄域人脸图像.判别器网络的主体结构由5个步长为2的 4×4 卷积核构成,通道数分别为64,128,256,512,1024.在每一层后面都进行实例归一化(instance normalization, IN)和采用了Leaky ReLU激活.判别器末端由两个全连接层分支组成,一个分支输出图像的真伪,另一个分支为年龄分类器,其输出为一个 n 维(n 为年龄组的个数)的one-hot向量,表示输入的合成人脸图像的预测年龄组.

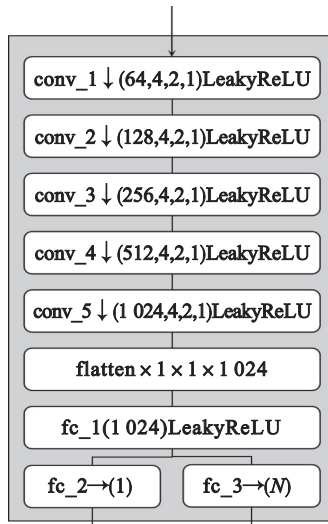


图3 判别器结构图

Fig. 3 Diagram of discriminator structure

1.3 损失函数

1.3.1 对抗损失

训练不稳定及其导致的生成图像失真原始GAN面临的一大问题.原始GAN学习的不稳定是由于目标函数引起的,在训练过程中使用非饱和损失和极小极大损失都会导致更新生成器时梯度消失,最终的结果是生成器难以更新而判别器很容易就能区分真伪图像.为了克服原始GAN训练不稳定和模式崩塌的问题,一种Wasserstein GAN(WGAN)的变体被提出^[31].本文

遵循WGAN策略定义判别器和生成器的对抗损失函数如下:

$$\min_{\|D\|_L \leq 1} L_{adv_d} = -E_{x_t \sim p_{data}(x_t)} D(x_t) + E_{x_s \sim p_{ann}(x_s)} D(G(x_s, \alpha_t)), \quad (5)$$

$$\min_G L_{adv_g} = -E_{x_s \sim p_{ann}(x_s)} D(G(x_s, \alpha_t)). \quad (6)$$

生成器和判别器的对抗损失采用WGAN-GP^[32]策略进行优化.在训练过程中,交替训练生成器和判别器,迭代地最小化生成器损失和判别器损失,直到生成图像具有令人满意的质量为止.

根据1.2.2节所述,属性解耦模块 D_f 对潜在表示施加一定约束,在与编码器博弈的过程中使得对于年龄信息的修改尽可能来自于年龄编辑模块的输入,损失函数定义如下:

$$L_{adv_f} = -E_{x_s \sim p_{ann}(x_s)} [\log(D_f(AEM_1(G_{enc}(x_s))))], \quad (7)$$

$$L_{adv_c} = -E_{x_s \sim p_{ann}(x_s)} [\log(1 - D_f(AEM_1(G_{enc}(x_s))))]. \quad (8)$$

1.3.2 分类损失

在人脸年龄编辑这一特殊任务中,判别器除了能够准确预测输入图像的真伪外,还需要具备判断输入图像所属年龄组的能力.因此,为确保合成人脸图像的年龄在目标年龄域内,本文在传统的判别器中增加了一个年龄分类器分支 C ,用目标年龄标签来指导分类器的学习,约束生成的人脸图像在一定的年龄范围内.定义用于训练判别器和生成器的年龄分类损失如下:

$$L_{cls_d} = \sum_{i=1}^n -\alpha_i \log C_i(x_t) - (1 - \alpha_i) \log(1 - C_i(x_t)), \quad (9)$$

$$L_{cls_g} = \sum_{i=1}^n -\alpha_i \log C_i(G(x_s, \alpha_t)) - (1 - \alpha_i) \log(1 - C_i(G(x_s, \alpha_t))). \quad (10)$$

其中: C_i 表示第 i 个年龄组的预测值; α_i 表示第 i 个年龄组的真实标签.

1.3.3 重建损失与身份损失

如图1所示,编码器输出的潜在表示 $f_s = G_{enc}(x_s)$ 直接被送入解码器中进行重建学习,目的是使保留图像的属性无关的细节特征,如人脸身份、光照和背景等.重建得到的图像要求与源图像尽可能接近,将重建学习的损失函数表述如下:

$$L_{rec} = \|x_s - G_{dec}(f_s)\|_1, \quad (11)$$

其中, $\|\cdot\|_1$ 表示 \mathcal{L}_1 范数.重建损失采用 \mathcal{L}_1 范数可以抑制生成图像的模糊性,并且有助于保留图像的边缘信息.

本文采用一个预先训练好的身份保留模块 ϕ_{id} ,在网络的最后加入带有softmax激活函数的全连接层实现分类.将源人脸图像和合成人脸图

像分别输入到 ϕ_{id} 中,如果合成人脸图像与源人脸图像的身份一致,就给出一个小的惩罚.相反,如果合成人脸图像的身份被认为发生了改变,则给出一个较大的惩罚.本文将身份损失 L_{id} 引入目标函数,定义如下:

$$L_{id} = \text{MSE}(\phi_{id}(x_s), \phi_{id}(G(x_s, \alpha_t))). \quad (12)$$

其中, $\text{MSE}(\cdot)$ 表示均方误差损失.

将上述损失函数结合到一起,可以分别得到整体的判别器目标函数和生成器目标函数,其公式归纳如下:

$$\min_{D,C} L_{dis_cls} = L_{adv_g} + \lambda_1 L_{cls_g}, \quad (13)$$

$$\min_{G_{enc}, G_{dec}, AEM} L_g = L_{adv_g} + \lambda_2 L_{cls_g} + \lambda_3 L_{rec} + \lambda_4 L_{id} + \lambda_5 L_{adv_c}. \quad (14)$$

其中, $\lambda_i, i=1, \dots, 5$ 为超参数,用于协调损失函数的各个部分的权重,调整这些超参数可以平衡人脸年龄合成任务的各项约束条件,从而使生成图像在保留相应的人脸属性的同时达到年龄准确性.

2 实验结果与分析

2.1 实验数据集

本文在 CACD^[33] 上进行实验,验证本文提出算法的年龄合成性能. CACD 是通过 Google 图像搜索收集的公共数据集,包含 10 年间 2 000 位名人的 16 334 张彩色图片,年龄标注从 14 岁到 62 岁,数据集中的图像在姿态、照明、表情等方面均显示了不同程度的变化.在训练之前,对 CACD 中的数据进行对齐、裁剪为 $128 \times 128 \times 3$ 尺寸,并过滤掉数据集中人脸数超过 1 个的图像.由于数据集中不包含儿童图像,并且 20 岁以下人脸图像数目较少而且年龄分布不均,本文只采

用 20 岁以上人脸图像进行训练测试.将数据集分为 4 组,分别为 21~30, 31~40, 41~50, 51~62.将处理后的数据集中的 95% 作为训练集,余下 5% 作为测试集. CACD 训练集和测试集各个年龄段的数据信息统计如表 1 所示.

表 1 CACD 数据信息统计
Table 1 CACD data information statistics

年龄组	21~30	31~40	41~50	51~62
训练集	38 620	41 024	37 805	26 534
测试集	2 085	2 111	1 955	1 465

2.2 实验参数设计

在训练阶段,设置超参数 $\lambda_1 = 10, \lambda_2 = 10, \lambda_3 = 100, \lambda_4 = 5 \times 10^{-5}, \lambda_5 = 10$. 使用 Adam 优化器 ($\beta_1 = 0.9, \beta_2 = 0.999$) 对生成器、判别器、属性解耦模块进行优化,学习率设置为 5×10^{-4} . 网络以 16 张图片为一个批次训练 40 万次迭代.

2.3 定性分析

生成对抗网络的模型效果描述体现在网络生成的直观图像上,本实验通过视觉上的定性分析提供关于面部年龄变化的评估. CACD 上的年龄渐进合成结果的例子如图 4 所示.可以观察到,这些例子涵盖了性别、姿势、妆容、表情等特征丰富的人脸面孔,合成的结果显示,随着年龄的增长,人脸图像面部逐渐出现自然的皱纹、眼袋、法令纹和额头的纹理加深,男性的胡须逐渐变多、头发呈现稀疏的趋势,并且头发和胡须颜色变浅甚至变白.从直观的角度看,人脸图像变换前后的身份特征没有受到显著影响,说明合成结果在视觉上取得了合理和令人信服的渐进衰老效果.

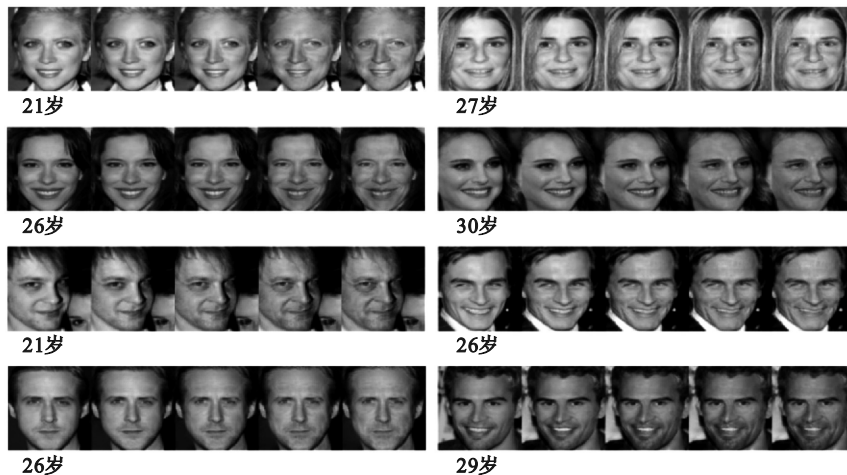


图 4 人脸年龄渐进合成结果

Fig. 4 Progressive face age synthesis results

将本文算法与 CAAE(conditional adversarial autoencoder)^[20], IPCGAN(identity-preserved conditional GAN)^[21], SAM(spatial attention modules)-GAN^[28]和 PFA(progressive face aging)-GAN^[27]人脸年龄合成算法进行对比.图 5 显示了其他人脸年龄渐进合成算法和本文算法在 CACD 上的人脸老化结果.通过观察可以发现,本文算法在图像质量、身份保留、老化效果方面均体现出了优势:CAAE 合成出来的图像降低了输出

图像的质量,对于背景信息的保存能力较弱,衰老效应相对不明显. IPCGAN 合成的图像有时呈现出伪影以及与年龄无关属性的改变,比如图 5 中左侧第 3 行第 5 列的面部出现了源图像中没有的眼镜.此外,通过 IPCGAN 合成的人脸在肤色上呈现出明显的变化. SAM-GAN 和 PFA-GAN 可以合成自然的衰老效果,但是它们仅仅关注与纹理有关的老化细节并且面部属性未充分解耦,与本文算法相比,面部肌肉和毛发的老化效应不显著.

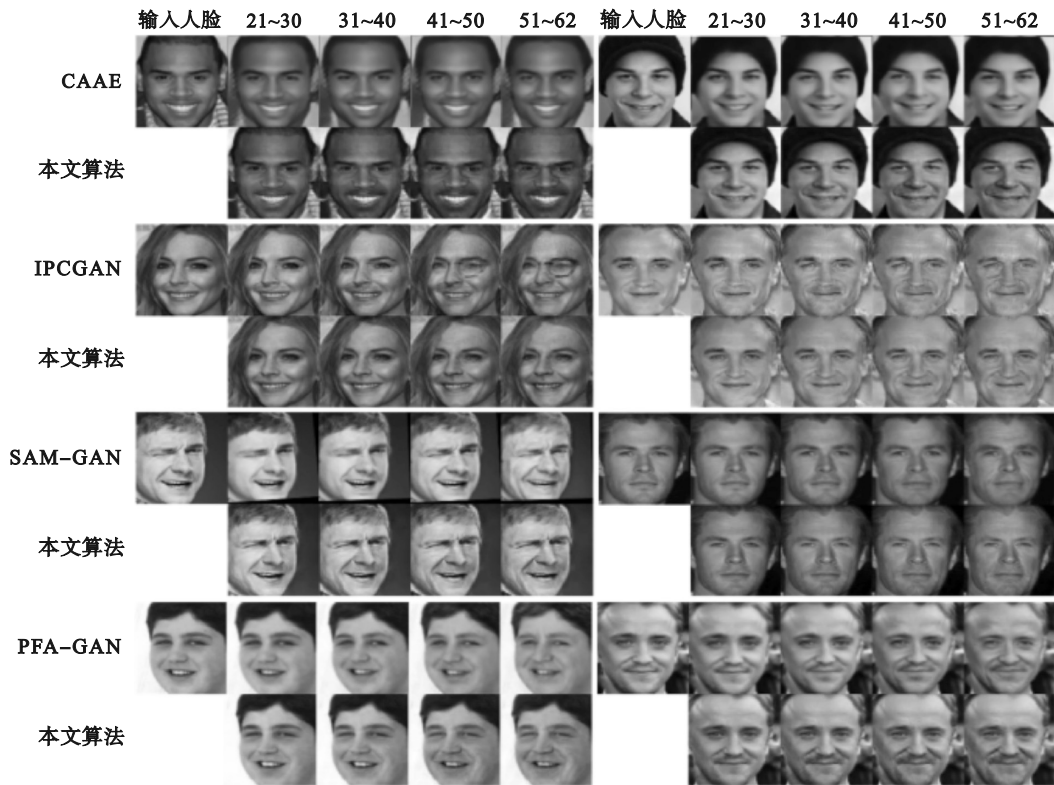


图 5 与其他人脸年龄渐进合成算法对比

Fig. 5 Comparison with other progressive face age synthesis algorithms

2.4 定量分析

在视觉效果上,本文算法呈现出真实自然的人脸年龄渐进合成效果.为了更加全面地评估年龄变换结果,本实验用定量分析的方法评价人脸年龄合成性能.本文采用一个在线人脸分析工具 Face++ (<http://www.faceplusplus.com>.)客观分析了所提出的人脸年龄渐进合成算法的性能,并与其他算法进行了各项性能指标的对比.

由于个体之间存在差异性,为了客观评价不同人脸图像的衰老情况,通过 Face++ 统一衡量测

试集中人脸图像老化的精度.尽管个体的衰老速度受到外界和自身的影响而有所不同,导致估计年龄和实际年龄的偏差,但是数据集整体的变化趋势具有一定的鲁棒性.分别计算合成人脸和真实人脸各个年龄段的平均年龄及标准差,以此作为描述数据集年龄分布情况的参考标准.从表 2 的年龄分布评估结果中可以看到,合成人脸数据和测试集真实人脸数据之间的均值和标准差在各个年龄段都较为接近,说明合成人脸数据分布与数据集真实数据分布拟合状况较好.

表 2 年龄分布评估结果
Table 2 Age distribution assessment results

年龄组	21~30	31~40	41~50	51~62
真实人脸	28.43±7.28	35.31±8.83	43.00±11.25	49.88±12.20
合成人脸	29.63±7.94	36.34±9.42	44.43±9.87	48.39±11.95

为了更加清晰体现本文算法合成目标年龄人脸的能力,表3对比了各种方法的平均年龄误差,误差越小说明合成人脸数据的年龄越接近真实人脸数据的年龄.从对比结果可以看到,本文方法在31~40,41~50和51~62这三个年龄段的平均年龄误差均为最小,表明本文算法具有良好的年龄准确性.

表3 平均年龄误差比较
Table 3 Comparison of mean age errors

年龄组	21~30	31~40	41~50	51~62
IPCGAN	0.74	1.16	1.55	3.09
SAM-GAN	0.67	1.29	1.46	2.06
PFA-GAN	0.81	1.05	1.51	2.73
本文算法	1.20	1.03	1.43	1.49

使用Face++进行客观人脸验证,表4中列出了在置信阈值为76.5和0.001%错误率下的验证

表4 客观人脸验证结果
Table 4 Objective face verification results

年龄组	21~30	31~40	41~50	51~62
真实人脸	93.75±1.25	92.58±2.19	89.87±4.17	88.74±4.66
21~30	—	96.20±0.99	93.91±3.12	93.00±3.63
31~40	—	—	—	95.54±1.77
41~50	—	—	—	96.92±0.20

表5 人脸验证率比较
Table 5 Comparison of face verification rate

年龄组	21~30	31~40	41~50	51~62	%
IPCGAN	99.99	98.50	97.82	95.94	
SAM-GAN	100	99.87	98.39	96.20	
PFA-GAN	100	99.99	98.47	96.97	
本文算法	100	99.86	98.65	97.12	

2.5 消融实验

为了更加清晰显示年龄编辑模块和属性解耦模块的作用,本文采用消融实验进行对比分析.在实验中,分别研究包含与不包含年龄编辑模块(AEM)、属性解耦模块(D_f)对于人脸年龄变换结果的影响,其中不包含年龄编辑模块的身份信息输入采用将年龄向量与潜在表示拼接的方式,训练相同的批次后将人脸合成结果展示在图6中.图6分别列出了相同测试样本在本文算法、不包含年龄编辑模块的方法、不包含属性解耦模

置信度.置信度表示两张人脸的相似度,置信度越高,两张脸属于同一个人的可能性就越大.如果置信度超过“ $1e-5$ ”的阈值,那么它们很有可能来自同一个人.虽然老化过程难以避免地造成了年龄渐进合成验证置信度随着源图像年龄和目标年龄之间间隔的增大而降低,但是人脸衰老效应最明显的年龄组的验证置信度仍然远远超过了置信阈值.

表5中比较了各方法的人脸验证率.人脸验证率表示各年龄组中身份相同的人脸对数占各年龄组验证总人脸对数的百分比,验证率越高说明被认为是身份相同的人脸对越多,身份一致性越高.根据表5的结果显示,本文算法在21~30,41~50和51~62三个年龄组均取得了最高的人脸验证率,仅在31~40年龄组略低于PFA-GAN的验证率.实验结果总体上证明了本文算法的身份保存能力.

块的方法以及两者都不包含4种方法下的年龄合成结果.通过将第1行分别和第2,3行对比可以发现,加入AEM或 D_f 使得由老化引起的图像局部阴影减少,图像清晰度有所提升.与其他模型相比,本文算法合成老年人脸图像的眼袋和皱纹更加明显,说明对于年龄的编辑效果更加显著.此外,本文算法的老化人脸与输入人脸之间皮肤颜色变化较小,而其他方法则会产生皮肤颜色变暗的现象,表明本文算法的各个模块之间取得了较好的平衡.

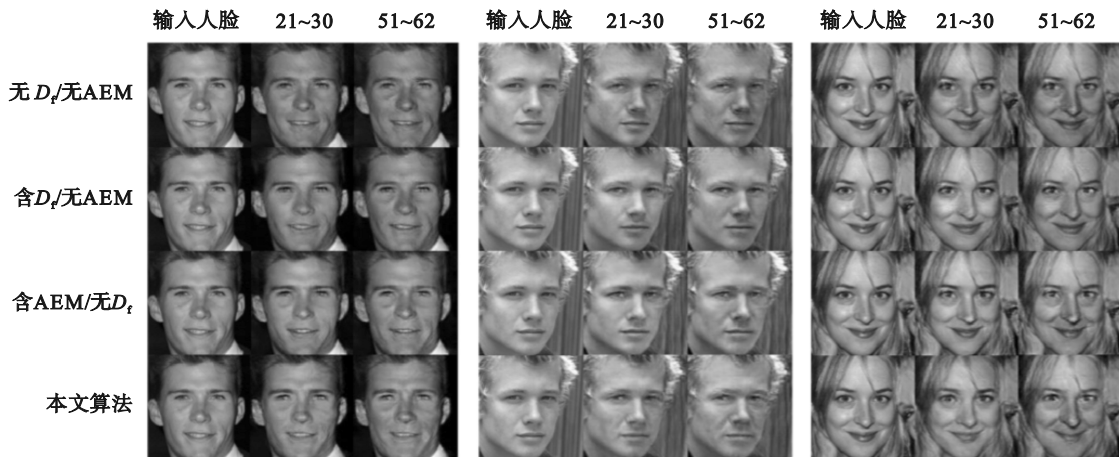


图6 消融实验结果

Fig. 6 Ablation results

3 结 论

本文针对人脸年龄合成中伪影和属性特征改变的问题,提出了一种基于生成对抗网络的人脸年龄渐进合成算法.通过使用卷积门控循环单元构成的年龄编辑模块,自适应地加入年龄特征,提出了通过对抗策略使潜在空间年龄特征与个人特征解耦,并在传统的判别器中添加了年龄分类器,对于合成人脸图像的年龄分类进行约束,以及通过身份确认和重建损失共同实现身份一致性.通过WGAN-GP对网络进行训练,避免了GAN的模式崩塌,使得训练结果更加鲁棒.定性和定量实验结果表明,本文算法可以得到真实自然的人脸渐进老化效果,且在很大程度上保留了人脸身份特征,与以往算法的对比结果进一步证明了提出算法的有效性.

参考文献:

- [1] Suo J L, Zhu S C, Shan S G, et al. A compositional and dynamic model for face aging [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2010, 32 (3): 385-401.
- [2] Tazoe Y, Gohara H, Maejima A, et al. Facial aging simulator considering geometry and patch-tiled texture [M]. New York: ACM, 2012.
- [3] Tsai M H, Liao Y K, Lin I C. Human face aging with guided prediction and detail synthesis [J]. *Multimedia Tools and Applications*, 2014, 72(1): 801-824.
- [4] Todd J T, Mark L S, Shaw R E, et al. The perception of human growth [J]. *Scientific American*, 1980, 242(2): 132-145.
- [5] Suo J L, Chen X L, Shan S G, et al. A concatenational graph evolution aging model [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2012, 34 (11): 2083-2096.
- [6] Ramanathan N, Chellappa R. Modeling age progression in young faces [C]//2006 IEEE Computer Society Conference on Computer Vision and Pattern Recognition. New York, 2006: 387-394.
- [7] Kemelmacher-Shlizerman I, Suwajanakorn S, Seitz S M. Illumination-aware age progression [C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Columbus, 2014: 3334-3341.
- [8] Tiddeman B, Burt M, Perrett D. Prototyping and transforming facial textures for perception research [J]. *IEEE Computer Graphics and Applications*, 2001, 21(5): 42-50.
- [9] Wang W, Cui Z, Yan Y, et al. Recurrent face aging [C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Las Vegas, 2016: 2378-2386.
- [10] Duong C N, Luu K, Quach K G, et al. Longitudinal face modeling via temporal deep restricted boltzmann machines [C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Las Vegas, 2016: 5772-5780.
- [11] Duong C N, Quac K G, Luu K, et al. Temporal non-volume preserving approach to facial age-progression and age-invariant face recognition [C]//Proceedings of the IEEE International Conference on Computer Vision. Venice, 2017: 3735-3743.
- [12] Goodfellow I, Pouget-Abadie J, Mirza M, et al. Generative adversarial nets [J]. *Advances in Neural Information Processing Systems*, 2014, 27(2): 2672-2680.
- [13] Zhu J Y, Park T, Isola P, et al. Unpaired image-to-image translation using cycle consistent adversarial networks [C]//Proceedings of the IEEE International Conference on Computer Vision. Venice, 2017: 2223-2232.
- [14] Isola P, Zhu J Y, Zhou T H, et al. Image-to-image translation with conditional adversarial networks [C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Honolulu, 2017: 1125-1134.
- [15] Choi Y, Uh Y, Yoo J, et al. Stargan v2: diverse image synthesis for multiple domains [C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. Seattle, 2020: 8188-8197.
- [16] Park T, Liu M Y, Wang T C, et al. GauGAN: semantic image synthesis with spatially adaptive normalization [M]. New York: ACM, 2019.
- [17] Karras T, Aittala M, Laine S, et al. Alias-free generative adversarial networks [J]. *Advances in Neural Information Processing Systems*, 2021, 34(1): 852-863.
- [18] Mirza M, Osindero S. Conditional generative adversarial nets [EB/OL]. (2014-10-06) [2023-03-04]. <https://arxiv.org/pdf/1411.1784.pdf>.

- [19] Antipov G, Baccouche M, Dugelay J L. Face aging with conditional generative adversarial networks [C]//2017 IEEE International Conference on Image Processing. Beijing, 2017:2089-2093.
- [20] Zhang Z F, Song Y, Qi H R. Age progression/regression by conditional adversarial autoencoder [C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Honolulu, 2017:5810-5818.
- [21] Wang Z W, Tang X, Luo W X, et al. Face aging with identity-preserved conditional generative adversarial networks [C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Salt Lake City, 2018:7939-7947.
- [22] Yang H Y, Huang D, Wang Y H, et al. Learning face age progression: a pyramid architecture of gans [C]//IEEE Conference on Computer Vision and Pattern Recognition. Salt Lake City, 2018:31-39.
- [23] Yang H Y, Huang D, Wang Y H, et al. Learning continuous face age progression: a pyramid of gans [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2021, 43(2):499-515.
- [24] Liu Y F, Li Q, Sun Z N. Attribute-aware face aging with wavelet-based generative adversarial networks [C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. Long Beach, 2019: 11877-11886.
- [25] Shao J, Bui T D. Wavelet-based multi-level generative adversarial networks for face aging [C]//The 32nd British Machine Vision Conference. Virtual: British Machine Vision Association, 2021: 1388-1399.
- [26] Li P P, Hu Y B, Li Q, et al. Global and local consistent age generative adversarial networks [C]//International Conference on Pattern Recognition. Beijing, 2018: 1073-1078.
- [27] Huang Z Z, Chen S Z, Zhang J P, et al. PFA-GAN: progressive face aging with generative adversarial network [J]. *IEEE Transactions on Information Forensics and Security*, 2020, 15(16):2031-2045.
- [28] Li Q, Liu Y F, Sun Z N. Age progression and regression with spatial attention modules [C]//Proceedings of the AAAI Conference on Artificial Intelligence. New York: AAAI, 2020:11378-11385.
- [29] Huang Z Z, Zhang J P, Shan H M. When age-invariant face recognition meets face age synthesis: a multi-task learning framework [C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. Virtual, 2021: 7282-7291.
- [30] Lample G, Zeghidour N, Usunier N, et al. Fader networks: manipulating images by sliding attributes [C]//Annual Conference on Neural Information Processing Systems. Long Beach: MIT Press, 2017:5963-5972.
- [31] Arjovsky M, Chintala S, Bottou L. Wasserstein GAN [EB/OL]. (2017-01-26) [2023-03-04]. <https://arxiv.org/pdf/1701.07875.pdf>.
- [32] Gulrajani I, Ahmed F, Arjovsky M, et al. Improved training of Wasserstein gans [C]//Annual Conference on Neural Information Processing Systems. Long Beach: MIT Press, 2017:5769-5779.
- [33] Chen B C, Chen C S, Hsu W H. Face recognition and retrieval using cross-age reference coding with cross-age celebrity dataset [J]. *IEEE Transactions on Multimedia*, 2015, 17(6):804-815.