BeautyGAN+: 基于全新 PMT 数据集的美妆转移混合监督学习算法

白玉, 颜波*, 谭伟敏

复旦大学计算机科学技术学院上海市智能信息处理重点实验室, 上海 200433

* 通信作者. E-mail: byan@fudan.edu.cn

收稿日期: 2021-03-16; 修回日期: 2021-05-24; 接受日期: 2021-07-07; 网络出版日期: 2022-01-11

国家自然科学基金 (批准号: U2001209, 61902076) 资助项目

摘要 作为改变人脸图像外观的重要方式之一, 美妆转移任务在近年来获得了越来越多的关注. 美妆转移网络可以将任意参考图像的妆容转移到任意人脸图像上, 在节约时间和空间成本的同时帮助人们找到适合自己的妆容, 并获得妆容得到美化的图像. 美妆转移网络 BeautyGAN 成功展示了无监督方法妆容迁移的较好效果, 然而此类无监督学习策略缺乏可以提供良好妆容细节的监督标签, 导致实验结果中出现了背景变色、边缘不自然、容易受光照影响等问题. 用素颜-美妆配对标签进一步约束 BeautyGAN 模型训练可以解决上述问题, 但当前学术界缺乏此类精细标注的配对数据集. 针对以上问题, 本文提出了 BeautyGAN+ 方法, 贡献了一个用户详细标注的素颜-美妆配对数据集, 即 PMT 数据集, 在 BeautyGAN 的无监督基础上加入有监督损失进行训练, 最终通过实验发现结合无监督 (迁移性好) 和有监督 (细节修复好) 的训练策略可以提供更好的迁移效果. 实验结果中的图像结果和用户调研 (14 位) 均表明本文方法显著提升了美妆转移效果, 解决了美妆转移工作长期缺乏配对数据集及转移结果存在多种偏差的难题.

关键词 深度学习, 美妆转移技术, 生成式对抗网络, 有监督学习, 美妆数据集

1 引言

化妆 (makeup) 是改善人脸外观的重要方式之一. 对眼影、粉底等化妆品的使用能够修饰人的容貌, 从而获得扬长避短、提升人像美观程度的效果. 随着数字图像处理技术的发展, 人们对照片美颜的要求逐渐提高, 虚拟化妆这一领域逐渐兴起, 在试妆、人像美容等方面都有着广泛的应用. 然而, 目前的美妆应用提供的妆容模板较为单一, 无法帮助用户尝试任意参考图像的妆容. 在此背景下, 基于人脸图像的美妆转移尤为具有应用价值. 美妆转移能够在保留原有人像面部特征的基础上将参考图像的妆

引用格式: 白玉, 颜波, 谭伟敏. BeautyGAN+: 基于全新 PMT 数据集的美妆转移混合监督学习算法. 中国科学: 信息科学, 2022, doi: 10.1360/SSI-2021-0093
Bai Y, Yan B, Tan W M. BeautyGAN+: mixed-supervised makeup transfer learning algorithm based on new PMT dataset (in Chinese). Sci Sin Inform, 2022, doi: 10.1360/SSI-2021-0093

容转移到该人像上. 不同于传统的虚拟化妆任务, 美妆转移的参考图像是任意的, 提高了妆容的丰富性; 该技术也同时规避了在人像美颜时流行妆容变化、固定模板风格过时的问题. 同时, 美妆转移过程中不需要用户的手动交互, 大大提高了用户的便捷性.

近年来, 美妆转移模型的效果得到了不断的提高. 然而, 以 BeautyGAN 网络^[1]为代表的一系列美妆转移工作, 在许多背景、角度和遮挡条件下难以获得较好的转移效果. 基于传统方法的美妆转移工作往往缺乏足够多符合要求的训练数据且在实验结果上表现出较差的迁移性, 而基于深度学习的方法往往难以还原妆容的细节且在极端位姿和条件下难以正确进行转移. 差异较大图像间的妆容转移下背景和发色不必要的改变、淡妆转移能力较差、部分面部区域无法识别、腮红等部分化妆品不敏感、转移结果边缘粗糙等问题 (如图 1 所示), 都阻碍了美妆转移在实际生活中的应用.

针对上述问题, 本文首先按照 BeautyGAN 网络^[1]的文献内容及参数设置进行了复现, 获得了具有完整美妆转移功能的模型, 对网络性能进行了一系列测试, 并针对测试结果中暴露的问题提出了以结合有监督和无监督思想为主要思路的 BeautyGAN+ 方法. 目前的美妆转移网络很少使用有监督学习方法的原因主要在于缺乏人脸像素级对齐的、化妆与素颜成对的数据集. 针对这一现状, 本文提出了一个具有成对数据的数据集, 该数据集包含大量符合实际生活的淡妆数据且对腮红、修容等其他美妆数据集中缺乏关注的化妆品进行了特别的关注.

本文通过在不监督的 BeautyGAN 网络模型基础上加入有监督损失进行进一步训练来实现 BeautyGAN+, 帮助模型进一步收敛到更好的性能. 实验证明, 本文 BeautyGAN+ 方法的主要贡献可总结为以下几点:

(1) 本文贡献了第 1 个像素级对齐的美妆转移数据集, 简称为 PMT 数据集, 包含 1115 组素颜 - 美妆成对数据; 该数据集为美妆转移工作中的有监督学习创造了可能, 且对淡妆数据有更多的关注, 并为今后构造此类数据集提供了全新的思路.

(2) 方法上, 基于该 PMT 数据集, 本文提出了在目前无监督方法中加入有监督损失的 BeautyGAN+ 方法, 兼具了无监督学习的迁移性和有监督学习对细节的保持能力, 是目前美妆转移工作中第 1 个混合监督的方法.

(3) 训练方式上, 本文采取在不监督模型基础上继续加入有监督损失的训练方式, 用较少的训练次数克服了无监督模型对背景色彩处理不佳和边缘处理较差的问题, 使用较少的计算资源实现了模型性能较大的提升.

同时, 本文经过对原有损失函数的调整完成了对 BeautyGAN 网络生成器部分独立的有监督训练. 然而, 由于生成器网络本身具有缺乏在同类型其他数据上迁移能力的弊端, 即生成器模型的泛化性能较差, 该有监督生成器网络仅能在部分与训练集类似的妆容上表现出较好的效果; 而对于与训练集差异较大的妆容或面部, 实验结果存在妆容变色等多种不理想的情况. 两组对照实验证明了独立生成器网络在任务中存在的多种弊端, 说明了本方法中有监督和无监督结合的必要性. 综上所述, BeautyGAN+ 兼具了 BeautyGAN 网络本身的美妆转移性能和有监督网络修正美妆转移结果的优点, 在美妆转移问题中具有相对于同类型工作较好的实验效果和应用价值. 图 1 展示了 BeautyGAN+ 与无监督 BeautyGAN 方法实验结果的对比, 从细节转移或背景保持方面来说, BeautyGAN+ 都显示出了极大的优越性. 在后续实验中, 本文将 BeautyGAN+ 方法与目前其他美妆转移方法进行对比, 再次证明了本文方法效果的优越性.

本文在对上述工作进行描述和总结的同时, 结合生成式对抗网络和有监督学习的原理对每项实验设置及其结果进行了理论分析和相关论证, 通过不同美妆转移工作之间的比较对每项工作的优缺点进行了总结和评估, 并针对现有实验未能解决的问题对美妆转移工作提出了未来研究的展望.

图 1 (网络版彩图) BeautyGAN+ 与 BeautyGAN 网络实验结果对比图. 每列图像表示一组实验结果, 其中第 1 行为待转移的素颜图像, 第 2 行为带妆参考图像, 第 3 行为 BeautyGAN 网络实验结果, 第 4 行为 BeautyGAN+ 实验结果

Figure 1 (Color online) Qualitative comparisons between BeautyGAN+ and BeautyGAN. Each column represents a set of results, in which the first one is a non-makeup image to be transferred, the second is a reference image with makeup, the third one is the unsupervised BeautyGAN result, and the last one is the BeautyGAN+ result

2 相关工作

2.1 传统方法的美妆转移

早期的美妆转移任务主要使用传统的数字图像处理方法, 分为美妆提取、变形和转移 3 个步骤. 在基于实例的美妆转移方法^[2]中, 研究者利用经过三元图标注、皮肤特征提取等预处理的化妆前后对比图像进行妆容提取, 通过将图像变形为同一标准面部的方式进行转移, 并利用映射中的系数调整来控制妆容的浓淡. 然而, 该方法要求输入图像处于相似的光照条件和拍摄姿态, 妆容对比图也难以获得, 缺乏实用性.

此后提出的方法对此进行了改进. 研究者^[3]将美妆视为面部一个可分的颜色叠加层次, 通过将人像面部划分为脸部结构层、皮肤细节层和颜色层 3 个层次, 将经过叠加的皮肤细节层和 alpha blending 处理的颜色层叠加在待上妆图像的脸部结构层上, 实现了不需要成对参考图像的美妆转移方法. 该方法在当时获得了较好的美妆转移结果, 也可迁移应用在胡须、皱纹转移等问题中. 然而, 在人脸对齐过程中仍旧需要用户交互, 实际应用过程中用户亲和性较差.

2.2 机器学习方法的美妆转移

随着机器学习技术的发展, 美妆转移任务的研究进入下个阶段. 2016 年提出的深度化妆网络^[4]将美妆转移与风格推荐任务相结合, 利用人像欧氏距离推荐数据库中的参考图, 利用全卷积网络实现图像分割以完成给定人像的五官提取, 并最终实现了粉底、口红和眼影的妆容迁移.

BeautyGAN^[1] 采用双输入/输出的网络结构, 在参考图像非成对的条件下方较为精准地转移了妆容且保持了人像面部原有特征, 同时贡献了较大的美妆数据集. 下文对该方法有详细的分析与实践.

BeautyGlow^[5] 针对没有成对数据集的情形制定了全新的损失函数指导参考图像的分解, 将带妆人像分解为妆容和素颜两个向量, 并将妆容向量与待上妆图像相结合, 实现美妆转移的工作. 该方法结合 Glow 框架, 通过运用隐空间实现了结果较为精准、背景保持较好且能够调节妆容浓淡的美妆转移工作. 同时, 该方法可以应用在多种任务中, 在去雨、去雾等方面也有较好的应用价值.

PairedCycleGAN 网络研究者^[6] 提出, 美妆转移不应当仅仅看作是色彩的转移, 也包括睫毛等新轮廓的转移和嘴唇光泽的转移, 并针对该理论提出了新的前馈神经网络来实现这些数据的传递. 该方法基于生成式对抗网络, 联合前向美妆转移网络和后向卸妆网络, 实现了包括色彩和轮廓在内多种高频细节的传递, 完成了细节上更为精准的美妆转移.

2.3 风格转换

美妆转移问题和风格转换问题有许多类似的地方, 美妆转移可以看作人像的局部风格转换问题, 故风格转换方法思想有很多可借鉴之处. 风格转换模型可以根据每个模型对应的风格数分为: 一对一、一对多、一对全 3 种类型, 而美妆转移即可看作一个模型转换全部风格的问题. 不同的是, 美妆转移对于局部细节的精准性要求更高, 对于背景和面部特征保留的程度也更高^[5].

图像类比的传统处理方法在 2001 年就已经得到了实践. 该方法^[7] 需要源风格和目标风格成对的参考图像, 按照不同分辨率构造金字塔以学习图像间的映射关系, 并利用最邻近匹配方法获得目标结果. 在 2017 年该方法深度学习的版本^[8] 中, 研究者利用深度网络革除了原方法需要成对数据、处理速度较慢的弊端, 将图像划分为不同图像块进行最相似映射. 图像分部分处理、利用深度网络解决数据不成对问题的思想, 在美妆转移问题中也多有应用^[4].

3 数据集分析

3.1 MT 数据集介绍

MT 数据集是由 BeautyGAN 论文研究者^[1] 收集整理的、目前规模最大的化妆数据集. MT 数据集包含 3834 张经过筛选和裁剪的非成对 256×256 尺寸女性图像, 其中包括 1115 张素颜图像和 2719 张化妆图像以及每张图片对应的人脸结构位置标定文件. 基于数据集在美妆转移工作上的应用目的, 研究者将其命名为美妆转移 (makeup transfer, MT) 数据集.

MT 数据集具有规模最大 (同类型数据集一般少于 1000 张^[9~12])、美妆转移任务针对性强、妆容类型丰富和主题场景多样等诸多优点. 然而, 其存在的一些问题也导致模型在部分样例上表现不佳: MT 数据集缺乏对照组, 网络学习的是化妆和非化妆域间的关系而非图像本身之间的关系, 限制了对细节的处理性能; 浓妆数据远多于淡妆数据, 少数淡妆图像被分类为了素颜图像, 导致在该数据集上训练的模型对淡妆感知能力较差; 标定文件缺乏脸颊部位标定, 对脸颊处的化妆品感知能力差; 在性别上数据类型过分单一, 难以处理男性妆容转移任务等.

图 2 (网络版彩图) PMT 数据集成对图像示例

Figure 2 (Color online) Samples of paired images from PMT dataset

3.2 PMT 数据集

由 3.1 小节所述, 我们希望能够在 MT 数据集的基础上获得具有更多淡妆样本、像素级对齐成对图像的数据集. 然而, 考虑到直接拍摄化妆前后像素级对齐图像难度较大、同一主题下拍摄的成本也较高, 本文利用拥有众多用户和多种符合实际情况妆容模板的美图秀秀¹⁾工具对素颜人脸图像进行模板妆容嵌套, 以 MT 数据集非化妆部分 1115 张图片为基础获得了包含 1115 张素颜图像和对应 1115 张美妆图像的 PMT (paired makeup transfer) 数据集. 本文中使用的图像均来自目前已公开发表的 MT 数据集.

PMT 数据集在构造过程中将软件中的妆容强度参数设置为 100%, 为了避免过度美白失真, 将美白强度设置为 15%. 在后处理过程中, 对刘海和眼镜等遮挡下的图像进行了人工手动调整, 并进行了 JPG 到 PNG 格式的转换. 为了防止网络学习上的过拟合, 本文选取了 15 种风格各异的妆容进行数据集的重构, 并在滤镜分配过程中保持了不同人种图像比例一致.

PMT 数据集的成对数据使得美妆转移任务中的有监督学习获得了可能. 在使用过程中, 通过将素颜图像与经过相同美妆滤镜处理的参考图像相匹配, 即可将该滤镜中的妆容转移到素颜图像上, 并将结果与该素颜图像的预期结果图作比较形成有监督损失. 值得注意的是, PMT 数据集并非仅限于帮助网络学习给定十几种妆容的转移方法, 而是通过学习数据集中妆容迁移案例最终获得在任意妆容和面部条件下都具有迁移能力的网络. 该数据集下训练的网络具有对任意妆容转移的普适性能力.

经过处理获得的 PMT 数据集示例如图 2 所示. 相比于 MT 数据集, PMT 数据集在美妆转移任务中表现出了诸多优势: 具有成对数据, 有利于网络更好地处理妆容细节; 转移结果的参考图使得网络收敛更加具有方向性, 有助于提高收敛速度; 包含更多符合日常审美的淡妆图像, 更具有实际价值且有助于提高网络的淡妆感知能力等. 此外, PMT 数据集的构造思想使得美妆转移不再局限于现有的数据集, 任意包含素颜人脸的数据集都可以添加美妆滤镜成为具有成对数据的美妆数据集. 本文将 PMT 数据集公开在百度网盘链接²⁾中.

综上所述, 本文相关实验将以 PMT 数据集为主, MT 数据集为辅开展.

1) 手机版客户端: <https://xiuxiu.web.meitu.com/>.

2) <https://pan.baidu.com/s/1ikYihfdFWX325JG2SRIjzA> (提取码: 350u).

图 3 (网络版彩图) BeautyGAN 双输入 – 双输出结构示意图
Figure 3 (Color online) The dual input/output structure of BeautyGAN

4 方法

4.1 BeautyGAN

BeautyGAN^[1] 是发表在 2018 年国际多媒体会议 (ACM Multimedia, ACMMM) 上的一种基于生成式对抗网络的无监督美妆转移网络. 该网络将给定人脸的妆容迁移到另一人脸上, 获得美妆转移结果和参考图像的去美妆结果, 并相对较好地保持输入图像的面部和背景特征. 经过对各种色彩迁移方法的调查研究^[13], 该网络利用计算局部直方图损失方法实现了实例级 (instance-level) 的美妆迁移.

BeautyGAN 方法的双输入 – 双输出示意图如图 3 所示, I_{src} 和 I_{ref} 代表输入网络的素颜图像和美妆参考图, 中间输出结果 I_{src}^m 和 I_{ref}^{nonm} 代表美妆转移和去妆容的结果, 而输出部分的 I_{src}^{rec} 和 I_{ref}^{rec} 表示再次输入相同结构右侧生成器网络的重构结果. 网络采用双输入 – 双输出结构: 首先将素颜和参考美妆图像经过两组不同卷积进行特征提取, 将特征连接在一起, 输入残差块进行处理. 完成处理后, 利用两组反卷积对输出的特征层 (feature map) 进行上采样, 从而获得美妆转移和卸妆结果, 并将上妆结果、参考图和卸妆结果、素颜图像分别输入妆面和非妆面判别器中判断真实性.

训练过程中, BeautyGAN 通过规定下述 4 种损失实现转移的准确性和背景与面部特征的保留.

周期一致性损失 (cycle consistency loss). 用来保证两个对偶网络转换后的结果保持原有图像的背景信息. 将两次经过网络的输出和原始图像之间的距离作为周期一致性损失, 表示为

$$L_{cycle} = E_{I_{src}, I_{ref}} [\text{dist}(I_{src}^{rec}, I_{src}) + \text{dist}(I_{ref}^{rec}, I_{ref})], \quad (1)$$

其中, $(I_{src}^{rec}, I_{ref}^{rec}) = G(G(I_{src}, I_{ref}))$, 距离函数 $\text{dist}(\cdot)$ 可以选择 L1 范数、L2 范数或是其他指标.

感知损失 (perceptual loss). 通过计算深度卷积网络提取的高维特征差异保留图像面部信息, 使用 ImageNet 预处理的 16 层 VGG 网络按照式 (2) 进行损失的计算.

$$L_{perceptual} = \frac{1}{C_l \times H_l \times W_l} \sum_{ijk} E_l, \quad (2)$$

其中, $E_l = [F_l(I_{src}) - F_l(I_{src}^m)]_{ijk}^2 + [F_l(I_{ref}) - F_l(I_{ref}^{nonm})]_{ijk}^2$, F_{ijk}^l 为第 l 层位于 (j, k) 位置的第 i 个滤波器的激活值.

对抗损失 (adversarial loss). 在对抗学习中帮助生成器和判别器网络收敛, $D_{nonmakeup}$ 和 D_{makeup} 分别用于区分所给素颜和美妆图像的真伪并给出分数, 目标函数如式 (3) 所示, $L_{D_{makeup}}$ 同理:

$$L_{D_{nonmakeup}} = E_{I_{src}} [\log D_{nonmakeup}(I_{src})] + E_{I_{src}, I_{ref}} [\log(1 - D_{nonmakeup}(I_{ref}^{nonm}))]. \quad (3)$$

在每次迭代中通过分别更新判别器和生成器的参数实现网络的收敛. 表达式如下所示:

$$L_{\text{adversarial}} = L_{D_{\text{nonmakeup}}} + L_{D_{\text{makeup}}}. \quad (4)$$

最后, BeautyGAN 结合局部颜色迁移方法提出了利用人脸口红、眼妆和粉底引导掩码计算妆容直方图损失的化妆损失 (makeup loss) 以排除背景的影响, 如式 (5) 所示. 直方图损失如式 (6) 所示.

$$L_{\text{makeup}} = \lambda_l L_{\text{lips}} + \lambda_s L_{\text{shadow}} + \lambda_f L_{\text{face}}. \quad (5)$$

$$L_{\text{item}} = \|I_{\text{src}}^m \text{HM}(I_{\text{src}}^m \circ M_{\text{item}}^1, I_{\text{ref}} \circ M_{\text{item}}^2)\|_2, \quad (6)$$

其中 $M_{\text{item}}^1 = \text{FP}(I_{\text{src}}^m)$, $M_{\text{item}}^2 = \text{FP}(I_{\text{ref}})$. 函数 $\text{HM}(x, y)$ 表示颜色分布与 y 相同而保留内容信息为 x 的直方图损失, 函数 $\text{FP}(\cdot)$ 代表对参数的人脸解析过程, item 是集合 $\{\text{lips}, \text{shadow}, \text{face}\}$ 中的元素.

综上, 生成器的完整目标函数集合 4 类损失, 如式 (7); 而判别器目标函数仅由对抗损失构成.

$$L_G = \alpha L_{\text{adversarial}} + \beta L_{\text{cycle}} + \gamma L_{\text{perceptual}} + L_{\text{makeup}}. \quad (7)$$

4.2 有监督生成器网络

基于有监督学习本身有利于降低实验结果与输入待处理图像背景间差异、增强转移结果细节的特性, 本文利用具有成对数据的 PMT 数据集设计了以纯有监督方式完成美妆转移工作的技术路线, 与纯无监督的 BeautyGAN 网络实验共同构成 BeautyGAN+ 方法的两组对照实验.

BeautyGAN 的生成器网络 (generator) 具有独立的生成能力. 为了保证网络架构的一致性和周期一致性损失的作用, 本文的有监督生成器网络保留了原有的双输入-双输出、双生成器对称结构. 在原有损失的基础上, 本文加入了每次实验结果和 PMT 数据集中的目标化妆图作差得到的有监督损失 (supervised loss), 并排除了与判别器网络相关的对抗损失. 考虑到有监督损失同时具有感知损失保留面部信息的功能和化妆损失保持妆容的功能, 本文也排除了此两种损失. 本文最终采用了仅含有周期一致性损失和有监督损失的损失函数进行训练, 如下所示:

$$L_G = \lambda_{\text{cycle}} L_{\text{cycle}} + \lambda_{\text{supervised}} L_{\text{supervised}}. \quad (8)$$

4.3 BeautyGAN+

生成式对抗网络具有较强的学习能力, 其能力源于对抗学习过程中生成器生成的大量样本, 该过程中产生的样本几乎可以覆盖问题域所涉及的全部高维特征. 同时, 该网络对于问题具有较强的迁移能力, 可以生成两个样本间的一系列插值结果: 经过充分训练的网络在具有处理第 n 维高维特征为 a 的样本 A 和该特征为 b 的样本 B 生成能力的同时, 对于第 n 维高维特征为 $\alpha a - (1 - \alpha)b$ 的输入能够获得 A 和 B 之间 α 位置的、符合输出条件的输出, 而非获得 A 和 B 对应输出的半透明叠加.

生成式对抗网络的迁移能力来自于判别器网络的质疑能力和生成器网络对于质疑的调整能力. 网络的性质决定了训练时判别器网络“着重考察高分的子问题域”的工作方式: 考虑到在判别器中得分较低子问题域往往是生成效果较差的, 但得分较高的不一定是因为生成效果好而可能源自该部分训练样本的缺失. 基于这一原理, 生成式对抗网络以对得分较高的子问题域进行生成和打分为主, 避免样本缺失导致的虚假高分, 经过不断训练使得生成器网络获得迁移能力.

然而, 生成式对抗网络的不足之处在于, 判别器所拥有的知识是域级的, 对于涉及细节的问题仅能给出一个宽泛而非准确的答案. 于是, 判别器的指导下便出现了类似于 BeautyGAN 存在的“生成的结

图 4 (网络版彩图) BeautyGAN+ 网络结构图

Figure 4 (Color online) The framework of BeautyGAN+

果符合目标域条件而附带了许多不必要的背景迁移”的问题. 基于这一问题, 本文设计了在 BeautyGAN 基础上加入有监督损失的 BeautyGAN+ 方法, 以优化生成结果的细节. 本方法网络结构图如图 4 所示.

BeautyGAN+ 的主要思路为结合有监督和无监督学习两种学习方式, 使得网络在训练中获得更好的性能. 目前美妆转移工作的方法多为纯有监督或无监督学习, 二者各有利弊: 有监督方法缺乏迁移性, 而无监督方法在细节处理上有所欠缺. 使用混合监督主要有两个条件: 拥有相关数据集且训练效果能够在单种监督方法上得到改善. 在包括美妆转移在内的多种任务上, 这种混合监督方式的学习不但能最大限度地利用现有的数据^[14], 还能够优化模型的性能.

网络结构方面, BeautyGAN+ 保留了 BeautyGAN 网络所具有的双输入 – 双输出及双生成器 – 双判别器结构, 与前文中的有监督生成器网络构成是否存在判别器这一条件上的对照. 训练方式上, 本方法将训练分解为无监督 BeautyGAN 网络训练和有监督训练两个步骤: 前半段的训练直接复用 BeautyGAN 网络训练过程中的损失和参数设置, 如式 (7) 所示; 而有监督部分的损失则和有监督生成器网络相同, 且仅对生成器网络进行进一步训练, 如式 (8) 所示.

5 实验

5.1 网络结构和参数设置

无监督 BeautyGAN 网络依据生成式对抗网络原理^[15~17]、基础生成式对抗网络中的 CycleGAN

网络^[18]原理及相关网络代码进行实现.

网络设置上, 判别器两个输入输出分支在层内不共享参数; 而对于妆面判别器和非妆面判别器, 则均利用 70×70 的 PatchGANs^[19] 对局部重叠的图像块进行真实性判别. 网络示意图如图 3 所示.

损失设置上, 在训练过程中对抗损失在生成器和判别器的训练中使用方法不尽相同: 损失值与对输入待处理图像判断分数正相关且与对生成图判断分数负相关, 代表了网络的判别能力; 在生成器网络的训练中目的为将其最小化, 而在判别器网络中目的为将其最大化. 故而两组对抗损失实则为同一个损失对生成器和判别器的不同作用. 训练过程中构建感知损失时提取图像特征使用的 VGG-16 (visual geometry group, VGG) 网络直接复用了文献 [20] 中在 ImageNet 上预训练完成的模型, 应用过程中使用 VGG 网络的 relu4_1 层; 而在对抗损失训练的过程中, 为了便于训练的进行, 利用最小平方误差代替损失计算公式中所有负对数似然目标, 以 $D_{\text{nonmakeup}}$ 为例 (D_{makeup} 同理): 将原损失式 (3) 在训练过程中修改为式 (9):

$$L_{D_{\text{nonmakeup}}} = E_{I_{\text{src}}}[(D_{\text{nonmakeup}}(I_{\text{src}}) - 1)^2] + E_{I_{\text{src}}, I_{\text{ref}}}[D_{\text{nonmakeup}}(I_{\text{ref}}^{\text{nonm}})^2]. \quad (9)$$

修改公式的原理为: 将 $D_{\text{nonmakeup}}(I_{\text{src}})$ 的期望值理解为一个小于 1 的分值, 非妆面判别器给定自然素颜图像的分值越高说明判别器性能越好, 判定能力越强. 当 $D_{\text{nonmakeup}}(I_{\text{src}})$ 的期望从 0 趋近于 1 时, $\log D_{\text{nonmakeup}}(I_{\text{src}})$ 函数的变化趋势与 $(D_{\text{nonmakeup}}(I_{\text{src}}) - 1)^2$ 函数的变化趋势非常相似, 即可近似利用最小平方误差代替负对数似然目标. 同时, 初始情况下 $D_{\text{nonmakeup}}(I_{\text{src}})$ 的期望与 1 的距离较大, 随机初始化可能导致其值非常接近于 0, 从而使 $\log D_{\text{nonmakeup}}(I_{\text{src}})$ 值过小, 训练稳定性较差. 上述修改可以稳定训练过程并提高生成图像质量.

无监督 BeautyGAN 网络的生成器的完整目标函数集合了前述 4 类损失, 表示为式 (7). 其中, 设置 $\alpha = 1$, $\beta = 10$, $\gamma = 0.0001$, 如式 (10) 所示. 判别器则直接由对抗损失构成损失.

$$L_G = L_{\text{adversarial}} + 10 \times L_{\text{cycle}} + 0.0001 \times L_{\text{perceptual}} + L_{\text{makeup}}, \quad (10)$$

其中, 为了获得更好的实验结果, 经过实验对比, 将文献 [1] 提出的 $\gamma = 0.005$ 的取值调整为 $\gamma = 0.0001$ 的取值, 并沿用到此后的实验中. γ 参数值下调的原因为调整前感知损失比重过高, 结果中对于人脸形状拟合过度, 严重影响了美妆转移本身的效果.

在化妆损失中, 针对式 (5) 设置 $\lambda_l = 1$, $\lambda_s = 1$, $\lambda_f = 0.1$, 获得如式 (11) 所示的化妆损失:

$$L_{\text{makeup}} = L_{\text{lips}} + L_{\text{shadow}} + 0.1 \times L_{\text{face}}. \quad (11)$$

训练过程中, 设置学习率 (learning rate) 为 0.0002, 批量大小 (batch size) 为 1.

有监督生成器网络中, $\lambda_{\text{cycle}} = 10$, $\lambda_{\text{supervised}} = 10$, 具体损失表示如式 (12) 所示:

$$L_G = 10 \times L_{\text{cycle}} + 10 \times L_{\text{supervised}}. \quad (12)$$

BeautyGAN+ 在前半段的训练中, 复用了 BeautyGAN 网络训练过程中使用的参数, 即式 (10) 所示的损失函数; 在完成 BeautyGAN 网络训练后, 再在训练完成模型的基础上加入式 (12) 的有监督损失进行训练. 其中, 后半段有监督训练仅作用于生成器网络, 此处停止对判别器网络的训练.

5.2 实验环境和训练细节

本文进行 3 组训练的实验环境为 2 张如下版本型号的 GPU 卡: Intel(R) Core(TM) i7-9700K 3.60 GHz CPU, 32 GB RAM, NAVID RTX 1080 GPU.

图 5 (网络版彩图) BeautyGAN 实验结果

Figure 5 (Color online) Results of BeautyGAN

图 6 (网络版彩图) BeautyGAN 转移效果较差结果

Figure 6 (Color online) Failure cases of BeautyGAN

第 1 组 BeautyGAN 网络训练共计训练 250 次迭代, 每次迭代平均耗时 10 min (相当于每张 GPU 在实验过程中独立计算需耗时 20 min). 在实验停止时网络已达到收敛状态.

第 2 组有监督生成器网络与前述实验形成对照, 进行了共计 250 次迭代的训练 (事实证明, 该设置从生成器网络的收敛时间来考虑也是较为合理的). 同上所述, 每次迭代的平均耗时也为 10 min;

第 3 组 BeautyGAN+ 实验在第 1 组实验训练好的 BeautyGAN 网络基础上增加了针对生成器网络的 10 次迭代的有监督训练, 从而获得了 BeautyGAN+ 模型. 这里 10 次迭代的选取源自网络的收敛时间, 本实验仅进行了 10 次迭代的训练就实现了网络在有监督训练下的收敛.

此外, 为了防止成对数据在训练过程中导致过拟合的现象, 本文所有有监督实验均有意识地加入了少量 ground truth 图像与待转移图像妆容不匹配的噪声, 以获得更好的实验结果.

由 BeautyGAN+ 实验数据可知, 有监督训练本身并不需要消耗过多的计算资源即可帮助网络得到较大的改变. 由此可见本文提出的方法从效率的角度来讲也具有较大的应用价值, 以较为精巧的设置通过对计算资源要求较低的有监督修正获得了实验结果上较大的改善.

6 讨论

6.1 BeautyGAN 结果分析

无监督 BeautyGAN 实验结果在部分测试图像上展现了眼影、口红等化妆品的较好转移效果, 具有自然的边缘和较好的美学效果. 图 5 第 2 行展示了左侧素颜图以第 1 行美妆图为参考转移的结果.

然而, BeautyGAN 会不可避免地导致图像背景、发色上不必要的改变, 以及肤色变暗的情况 (如图 5 第 2 行第 2 列). 此外, 实验中出现了单侧光照下, 仅获得半边妆容转移的结果, 如图 6 左侧结果; 而由于缺乏对鼻影等化妆品的学习, 部分结果将鼻翼阴影错误转移到了人中位置, 如图 6 右侧结果.

结合上述实验结果, 本文对无监督 BeautyGAN 的弊端总结为: (1) 在参考图和输入素颜图像色调差异较大的情况下, 网络在转移妆容的同时会改变输入素颜图像的背景色和发色; (2) 由于 MT 数据集缺乏淡妆图像, 网络对于淡妆的敏感度及转移性能表现较差, 在日常淡妆中缺乏应用价值; (3) 面部光照不均匀的情况下, 受光照干扰会出现仅转移部分化妆品的情况; (4) 对于腮红、鼻影等多种修饰类

图 7 (网络版彩图) 有监督生成器网络实验结果细节图展示. 红色框图展示了转移结果和素颜图眼妆对比, 橙色为眉妆, 黄色为唇妆, 蓝色为肤色, 绿色为腮红, 紫色为发色对比; 同时, 还放大展示了参考图眼妆和唇妆的细节部分

Figure 7 (Color online) Detailed results of the supervised generator network. The red block shows the comparison between the eye makeup result and the non-makeup image. The orange, yellow, blue, green, and purple blocks show that by the eyebrow, lip, skin, cheek, and hair makeup, respectively. We show the zoom-in details of eye and lip makeup for reference

图 8 (网络版彩图) 有监督生成器网络实验结果图. 每列图像中, 第 1 行为带妆参考图, 第 2 行为待转移的素颜图像, 第 3 行为转移结果图

Figure 8 (Color online) Results of the supervised generator network. In each column, the first image is a reference with makeup, the second image is a non-makeup image to be transferred, and the third image is the transfer result

化妆品的处理能力较差; (5) 妆容边缘转移效果较差, 存在眼妆边缘参差不齐等美观度不及参考图妆容的情况.

此外, 由于数据集本身主题的局限性, BeautyGAN 具有图像面部大小和性别上的局限性. 然而解决该问题对数据集要求较高, 本文着重解决前述 5 种能够使用我们提出的方法在 PMT 数据集上得到解决的问题.

6.2 有监督生成器网络

有监督训练下的生成器网络在部分测试图像上获得了较好的效果. 如图 7 所示, 该网络的转移结果较好地保持了多种化妆品的色彩、质感和透明度: 腮红高透明度叠加、眼影边缘自然晕染、口红质地光泽且没有影响到牙齿或其他部位, 而肤色上保持了肩部、面部及素颜图像的一致性. 背景和面部形状上, 较好地保留了人物双眼皮、金黄发色的特点而未受参考图内双、黑发的影响. 该示例中输入素颜图像和参考图像来自不同人种, 表明该网络在跨人种转移任务中表现较好. 有监督生成器网络在遮挡图像上也表现出了较好的效果. 如图 8 所示, 在存在眼镜、刘海等遮挡的情况下没有机械地进行妆容叠加, 成功地使妆容避开了遮挡, 并进行了较好的边缘处理. 此外, 本方法在卸妆输出中也获得了较好的结果.

然而, 有监督生成器网络在部分图像中却有着较差的表现. 如图 9 所示, 部分结果出现了化妆品色彩改变的现象: 第 3 行的转移结果中 3 幅图像均出现了将深红色等暗色调口红转移为橙红色等色彩鲜明的唇妆; 在眼妆的转移中, 也存在着夸大了眼下卧蚕妆容或将眼妆改用鲜艳色彩的情况.

图 9 (网络版彩图) 有监督生成器网络部分失败结果展示

Figure 9 (Color online) Failure cases of the supervised generator network

图 10 (网络版彩图) BeautyGAN+ 网络去美妆实验结果. 图中一列为一组实验结果, 每列中第 1 行为参考素颜图像, 第 2 行为待去美妆图像, 第 3 行为去美妆结果

Figure 10 (Color online) De-makeup results of BeautyGAN+. Each column represents a set of results, in which the first one is a non-makeup reference, the second one is a makeup image to be dealt with and the third one is the de-makeup result

上述实验结果可以用有监督学习的特性进行解释. 由于有监督训练更具有方向性, 对实例的细节能够进行更精准的处理; 而泛化能力较差, 训练的网络生成能力有限, 对与训练集中图像差异较大的图像处理能力较差. 这里的相似性指的是网络提取高维特征的相似性, 输入素颜图像或参考图像与训练集的差异都可能导致转移效果的降低. 这是有监督学习的系统性弊端, 必须制定从原理上改进的技术路线来克服该问题. 综上, 本组实验证明了 BeautyGAN+ 方法中无监督部分和判别器部分的必要性.

6.3 BeautyGAN+ 结果分析

由图 1 展示的实验结果可知, BeautyGAN+ 方法解决了无监督和有监督网络存在的一系列问题: 首先, 从结果图第 1 列可知, BeautyGAN+ 较好地纠正了结果色调的改变, 保留背景深蓝色的同时避免了暗沉肤色过度迁移, 有利于在不同肤色人像上的迁移. 其次, BeautyGAN+ 较好地排除了光照的影响, 第 2 列展示了参考图像左侧 (观察者的左侧, 下同) 脸部阴影较重时, BeautyGAN 对阴影的错误转移和 BeautyGAN+ 在光影上较为自然的结果. 此外, 第 3 列光照差异较大的图像中, 本文方法基于较差光照转移了自然的深色口红而 BeautyGAN 错误地提取了裸色口红. 最后, 第 4 列显示本方法在腮红等化妆品上也作出了较好的表现, 而 BeautyGAN 的转移未涉及到此类化妆品. 综上, 本次实验结果通过增加有监督训练, 解决了 BeautyGAN 的多种缺陷, 获得了自然美观的美妆转移结果, 大大提高了网络的应用价值.

如图 10 所示, 本方法在与美妆转移对称的去美妆问题上获得了较好的效果. 图 11 从细节上展示了 BeautyGAN+ 方法的优越性, 第 1~4 列分别展示了 BeautyGAN+ 方法在极端表情、背景色调、头

图 11 (网络版彩图) BeautyGAN+ 实验结果对比图. 每列从上到下分别为参考图、素颜图、DMT 结果、PSGAN 结果、DFME 结果、BeautyGAN 结果和 BeautyGAN+ 实验结果, 方框中为细节展示

Figure 11 (Color online) Qualitative comparisons for BeautyGAN+. Each column from top to bottom shows the reference image, the non-makeup image, and the results of BeautyGAN, DMT, PSGAN, DFME, and BeautyGAN+, respectively. The zoom-in blocks show the details

发遮挡和腮红化妆品上相对于无监督方法的良好表现. 图中的细节框对两种方法在细节上的处理进行了重点展示.

本文还与其他的美妆转移方法进行了对比. 从转移方法的角度来看, 部分美妆转移方法对包括背景和发色在内的色彩均进行了转移^[21], 与本文保留背景特征的目的不符; 另一些美妆转移方法仅能对给定参数的特定化妆品进行生成^[22]. 故本文实验结果无法与上述方法进行比较. 如图 11 所示, 本文与目前效果较好的传统方法 DFME^[3] 及机器学习方法的 DMT^[23]、PSGAN^[24] 进行对比. 考虑到本文的混合监督方法与非有监督方法进行数值对比有失公平性, 故仅从视觉效果上进行定性对比说明, 而不进行定量对比. 由图可知, DFME 在一定程度上保持了妆容和面部特征, 但由于受刘海影响严重, 面部边缘具有较多伪影, 眼球也受参考图影响较为严重; 同时, 该方法需要用户交互调整人脸关键点, 便捷性较差. PSGAN 方法 (图中部分图像受模型尺度影响而放大, 不影响美妆转移的效果对比) 转移效果和 BeautyGAN 类似, 在转移妆容的同时未能较好地规避光照影响和背景变色等问题. DMT 方法没有体现出明显的伪影和变色, 但在口红、腮红等部分化妆品上转移不彻底, 结果图上妆容色彩相较

图 12 (网络版彩图) BeautyGAN+ 网络边缘生成较差的实验结果
Figure 12 (Color online) A failure case of BeautyGAN+

表 1 用户调研结果
Table 1 Result of user study

Method	Group 1 (chosen from supervised training dataset)			Group 2 (chosen randomly)		
	Rank1	Rank2	Rank3	Rank1	Rank2	Rank3
BeautyGAN+	81.905%	10.952%	7.143%	81.714%	8.095%	6.191%
Unsupervised	3.809%	62.857%	33.334%	8.095%	83.334%	8.571%
Supervised	14.286%	26.191%	59.523%	6.191%	8.571%	85.238%

于参考图更为寡淡. 由此可见, 本文提出的 BeautyGAN+ 方法不但比 BeautyGAN 方法有了较大提升, 在与其它先进方法的对比过程中也展现了其妆容提取、背景保持和边缘处理上的优越性.

此外, 考虑到多种工业应用都具有人脸美妆的功能, 本文对美图秀秀 (9.0.3.0 正式版)、抖音 (15.9.0)、B612 (9.7.11)、轻颜相机 (3.0.8) 和醒图 (1.8.1) 进行了功能调研. 结果显示, 上述应用均具有固定模板的人脸美妆功能, 部分应用包含一些基于风格迁移方法的特效. 然而, 上述应用均没有成熟的美妆转移功能, 基本局限于给定妆容模板的嵌套. 即便在未调研的应用中存在该功能, 美妆转移工作至今仍未作为一项成熟的功能得到广泛应用, 说明美妆转移目前仍旧处于研究阶段而未达到应用水平. 上述调研更进一步说明了本文方法的前瞻性和在未来工作中的应用价值.

然而, 结果显示该方法也存在一定的固有问题: 网络认为美妆转移是对颜色的转移, 未考虑睫毛等妆容轮廓的生成. 如图 12 所示, 测试结果未能生成参考妆容中的白色假睫毛. 此外, 由于数据集本身的限制, 本方法在远距离或小脸图像上不具备美妆转移能力.

6.4 用户调研

考虑到美妆转移任务具有的主观性, 本文组织 14 名志愿者进行了用户调研. 调研过程中, 要求每名志愿者进行两组实验: 每组实验均为让志愿者对 15 组图像分别进行评价排序, 每组图像均包含素颜图、参考图, 以及随机顺序出现、未标定所属实验组 (无监督、有监督或 BeautyGAN+) 的实验结果, 要求志愿者在不知晓每张结果所对应实验组的情况下从图像的真实度、对素颜图像的保持度和美妆转移的效果三方面进行排序, 效果最好的为第 1 位, 其次为第 2 位, 最差为第 3 位. 其中, 第 1 组使用的测试数据均使用了有监督学习训练集中出现过的妆容模板进行构造, 而第 2 组使用的测试数据则完全不考虑其与训练数据的关系. 在每组实验中选择 3 种方法为第 1, 2, 3 位的结果比例如表 1 所示.

由表 1 所示的实验结果可知, 在两组实验中, BeautyGAN+ 都在超过 80% 的测试中优于有监督和无监督方法; 当测试集中的妆容模板和有监督学习的训练集相似时, 有监督方法在部分案例上表现非常优秀, 但在其他案例中表现较差, 分析认为, 这是因为有监督方法仅学习过这些妆容而未在训练集中学习过这些人像; 当随机选择测试集时, 相比于有监督, 无监督方法则更有优势. 然而, 无论测试集

的选择方式如何, 本文提出的 BeautyGAN+ 方法在主观判断上始终优于其他两种方法.

7 总结与未来工作

本文提出了 BeautyGAN+ 方法, 解决了有监督学习泛化能力差和无监督学习细节处理能力差、结果粗糙等问题, 具有淡妆敏感、化妆品种类丰富、遮挡鲁棒、背景保持等多项优势, 并在对比实验和相关用户调研中显示了其优越性. 此外, 本文还提出了一个具有成对美妆数据的 PMT 数据集.

然而, 本方法在许多细节上仍旧存在可提升空间. 本方法仅实现了妆容色彩的迁移, 而没有考虑睫毛等轮廓的生成; 由于数据集本身的限制, 本方法在小脸图像上难以获得理想的效果. 另外, 由于数据集的训练集和测试集中均为女性图像, 我们没有进行男性美妆相关实验. 在未来工作中, 本文考虑在现有方法的基础上对包括面部妆容物理性分离^[22]、针对不同化妆品设计不同转移方法^[4]在内的多种其他美妆迁移方法进行尝试, 以获得美妆转移能力更强的模型.

参考文献

- 1 Li T, Qian R, Dong C, et al. BeautyGAN: instance-level facial makeup transfer with deep generative adversarial network. In: Proceedings of the 26th ACM International Conference on Multimedia, 2018. 645–653
- 2 Tong W, Tang C, Brown M, et al. Example-based cosmetic transfer. In: Proceedings of IEEE International Conference on Automatic Face and Gesture Recognition, 2007. 211–218
- 3 Guo D, Sim T. Digital face makeup by example. In: Proceedings of IEEE Conference on Computer Vision and Pattern Recognition, 2009. 73–79
- 4 Liu S, Ou X, Qian R, et al. Makeup like a superstar: deep localized makeup transfer network. In: Proceedings of International Joint Conference on Artificial Intelligence, 2016. 2568–2575
- 5 Chen H, Hui K, Wang S, et al. BeautyGlow: on-demand makeup transfer framework with reversible generative network. In: Proceedings of IEEE Conference on Computer Vision and Pattern Recognition, 2019. 10042–10050
- 6 Chang H, Lu J, Yu F, et al. PairedCycleGAN: asymmetric style transfer for applying and removing makeup. In: Proceedings of IEEE Conference on Computer Vision and Pattern Recognition, 2018. 40–48
- 7 Hertzmann A, Jacobs C E, Oliver N, et al. Image analogies. In: Proceedings of the 28th Annual Conference on Computer Graphics and Interactive Techniques, 2001. 327–340
- 8 Liao J, Yao Y, Yuan L, et al. Visual attribute transfer through deep image analogy. ACM Trans Graph, 2017, 36: 1–15
- 9 Chen C, Dantcheva A, Ross A. An ensemble of patch-based subspaces for makeup-robust face recognition. Inf Fusion, 2016, 32: 80–92
- 10 Dantcheva A, Chen C, Ross A. Can facial cosmetics affect the matching accuracy of face recognition systems? In: Proceedings of IEEE 5th International Conference on Biometrics: Theory, Applications and Systems (BTAS), 2012. 391–398
- 11 Chen C, Dantcheva A, Ross A. Automatic facial makeup detection with application in face recognition. In: Proceedings of International Conference on Biometrics (ICB), 2013. 1–8
- 12 Chen C, Dantcheva A, Swearingen T, et al. Spoofing faces using makeup: an investigative study. In: Proceedings of IEEE International Conference on Identity, Security and Behavior Analysis (ISBA), 2017. 1–8
- 13 Faridul H S, Pouli T, Chamaret C, et al. A survey of color mapping and its applications. In: Proceedings of Eurographics, 2014. 43–67
- 14 Sahasrabudhe M, Shu Z, Bartrum E, et al. Lifting AutoEncoders: unsupervised learning of a fully-disentangled 3D morphable model using deep non-rigid structure from motion. In: Proceedings of IEEE International Conference on Computer Vision, 2019. 4054–4064
- 15 Goodfellow I J, Pouget-Abadie J, Mirza M, et al. Generative adversarial nets. In: Proceedings of the 27th International Conference on Neural Information Processing Systems, 2014. 2672–2680

- 16 Rosca M, Lakshminarayanan B, Warde-Farley D, et al. Variational approaches for auto-encoding generative adversarial networks. 2017. ArXiv:1706.04987
- 17 Mirza M, Osindero S. Conditional generative adversarial nets. 2014. ArXiv:1411.1784
- 18 Zhu J, Park T, Isola P, et al. Unpaired image-to-image translation using cycle-consistent adversarial networks. In: Proceedings of IEEE International Conference on Computer Vision, 2017. 2242–2251
- 19 Li C, Wand M. Precomputed real-time texture synthesis with markovian generative adversarial networks. In: Proceedings of the European Conference on Computer Vision, 2016. 702–716
- 20 Simonyan K, Zisserman A. Very deep convolutional networks for large-scale image recognition. In: Proceedings of International Conference on Learning Representations, 2015
- 21 Yang Y, Zhao H L, You L H, et al. Semantic portrait color transfer with Internet images. *Multimed Tools Appl*, 2017, 76: 523–541
- 22 Li C, Zhou K, Lin S, et al. Simulating makeup through physics-based manipulation of intrinsic image layers. In: Proceedings of IEEE Conference on Computer Vision and Pattern Recognition, 2015. 4621–4629
- 23 Zhang H L, Chen W Q, He H, et al. Disentangled makeup transfer with generative adversarial network. 2019. ArXiv:1907.01144
- 24 Jiang W T, Liu S, Gao C, et al. PSGAN: pose and expression robust spatial-aware GAN for customizable makeup transfer. In: Proceedings of IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2020. 5193–5201

BeautyGAN+: mixed-supervised makeup transfer learning algorithm based on new PMT dataset

Yu BAI, Bo YAN* & Weimin TAN

Shanghai Key Laboratory of Intelligent Information Processing, School of Computer Science, Fudan University, Shanghai 200433, China

* Corresponding author. E-mail: byan@fudan.edu.cn

Abstract As one of the important ways to change the appearance of face image, makeup transfer has received more and more attention in recent years. Makeup transfer networks can translate the makeup style of a reference image to any other non-makeup one while preserving face identity, helping people find the most suitable makeup for them and get the beautified image. BeautyGAN, a makeup transfer network, has successfully demonstrated the results of unsupervised makeup transfer. However, unsupervised learning strategies lack supervised labels that can provide makeup details, leading to problems such as background discoloring, artifacts on edges and illumination disturbance in the experimental results. Constraint of BeautyGAN with makeup and non-makeup pairwise images can solve this problem, but academia lacks such fine-labeled paired dataset. To solve this problem, we propose BeautyGAN+, which adds supervised loss to BeautyGAN to train, and build up a new makeup dataset, the PMT dataset, that consists of makeup and non-makeup pairwise images. Finally, we find that combining unsupervised (migration) and supervised (details) loss can get better results through experiments. The experimental results and user study (14 participants) indicate that BeautyGAN+ significantly improves the results of makeup transfer, solving the problem of lack of paired dataset and various deviations in the transfer results.

Keywords deep learning, makeup transfer, generative adversarial network (GAN), supervised learning, makeup dataset