



面向人脸识别的口罩区域修复算法

李悦¹, 钱亚冠¹, 关晓惠², 李蔚¹, 王滨³, 顾钊铨⁴

- (1. 浙江科技学院大数据学院, 浙江 杭州 310023;
2. 浙江水利水电学院信息工程与艺术设计学院, 浙江 杭州 310027;
3. 杭州海康威视网络与信息安全实验室, 浙江 杭州 310051;
4. 广州大学网络空间先进技术研究院, 广东 广州 510006)

摘要: 遮挡下的人脸识别一直是现实场景中的一个难题。特别是新冠肺炎疫情爆发后, 在机场、车站等需要鉴别入场人员身份信息的场所, 口罩遮挡使得可供识别的面部特征大幅减少, 原有的人脸识别算法准确率随之下降。对去除口罩遮挡进行了研究, 提出了一个新的框架修复人脸, 利用边缘生成网络还原遮挡区域的边缘, 在此基础上再利用区域填充网络恢复被遮挡的人脸, 同时保留身份信息。为提升模型的性能, 提出空间加权对抗损失和身份一致性损失训练上述网络, 并利用关键点信息, 构建了两个戴口罩的人脸数据集。实验结果表明, 恢复被口罩遮挡的人脸的图像使人脸识别算法 ArcFace 的准确率达到 98.39%, 比直接采用 ArcFace 识别遮挡人脸提升了 4.13% 的准确率。

关键词: 口罩遮挡; 人脸识别; 边缘生成网络; 区域填充网络

中图分类号: TP393

文献标识码: A

doi: 10.11959/j.issn.1000-0801.2021193

Region inpainting algorithm of mouth-muffles for facial recognition

LI Yue¹, QIAN Yaguan¹, GUAN Xiaohui², LI Wei¹, WANG Bin³, GU Zhaoquan⁴

1. School of Big Data Science, Zhejiang University of Science and Technology, Hangzhou 310023, China
2. College of Information Engineering & Art Design, Zhejiang University of Water Resources and Electric Power, Hangzhou 310027, China
3. Network and Information Security Laboratory of Hangzhou Hikvision Digital Technology Co., Ltd., Hangzhou 310051, China
4. Cyberspace Institute of Advanced Technology, Guangzhou University, Guangzhou 510006, China

Abstract: Facial recognition under occlusion is a well-known difficult problem in real scenes. Especially after the outbreak of COVID-19, in airports, stations and other places that need to verify the identity of visitors, the mouth-muffle occlusion greatly reduces the facial features that can be important for identification, and the accuracy of face recognition algorithm decreases. Face de-occlusion was studied, and a novel framework was proposed to restore face, which used the edge generation network to generate edge maps. On this basis, the occluded region was

收稿日期: 2021-05-27; 修回日期: 2021-08-11

基金项目: 国家重点研发计划项目 (No.2018YFB2100400); 国家自然科学基金资助项目 (No.61902082); 浙江省公益技术应用研究项目 (No. LGF20F020007)

Foundation Items: The National Key R&D Program of China (No.2018YFB2100400), The National Natural Science Foundation of China (No.61902082), Applied Research Project of Public Welfare Technology of Zhejiang Province of China (No. LGF20F020007)

restored by the region completion network while preserving identity information. In order to improve the performance, the spatial weighted adversarial loss and identity-preserving loss were introduced to train the above two networks. Then two face datasets with mouth-muffles were constructed by face landmarks. The experimental results show that the accuracy of facial recognition algorithm ArcFace on the face dataset restored by the proposed model was 98.39%, which was 4.13% higher than that of directly using ArcFace.

Key words: mouth-muffle occlusions, facial recognition, edge generation network, region completion network

1 引言

基于深度学习的人脸识别技术已被广泛应用于金融、公安、教育等领域, 其中 DeepFace^[1]是应用深度学习的经典人脸识别算法, 在 LFW (labeled faces in the wild)^[2]数据集上取得了 97.35% 的识别率, 已经接近人类的水平。最近提出的 ArcFace^[3]算法, 则达到了 99.8% 的识别率。但目前的人脸识别算法在实际应用中易受一些外部因素 (如人脸遮挡、光照、面部表情、面部朝向等) 的影响。其中, 遮挡的影响最为显著, 可导致识别率降低 10% 左右^[4]。

2020 年年初, 新型冠状病毒肺炎疫情突然爆发, 要求人们必须佩戴口罩进入公共场所。尤其在机场、火车站等需要计算机人脸识别的场所, 口罩对人脸的遮挡必然会影响识别的效率和准确率, 而摘下口罩又会增加交叉感染的风险。因此, 如何在戴口罩的情况下实现高精度的人脸识别成为当前急迫需要解决的问题。

事实上, 遮挡条件下的人脸识别一直是该领域的研究热点^[5-6]。Wang 等^[7]提出了两阶段的重建方法: 先检测遮挡再恢复遮挡部分, 这种联合遮挡检测和恢复方法可以产生有利于分类的良好全局特征。Zhang 等^[8]用去遮挡的自编码网络恢复被遮挡区域。Yuan 等^[9]利用 3D 重建修复人脸遮挡部位, 但该方法需要与目标身份相同的参考数据。Zhao 等^[6]以相同身份的人脸图片为基础, 使用生成对抗网络 (generative adversarial network, GAN)^[10]生成无遮挡的人脸。Duan 等^[4]提出了 BoostGAN (boosting generative adversarial network) 先粗略

后精细地恢复人脸。He 等^[11]提出一种人脸图片编辑方法, 通过属性分类约束更加精确地修改人脸属性。Iizuka 等^[12]和 Li 等^[13]利用 GAN, 结合全局判别器与局部判别器, 从整体以及局部区域判断修复图片的语义一致性。以全局判别器与局部判别器同时使用为基础, ID-GAN (identity-diversity generative adversarial network)^[14]将基于深度学习的识别器和基于 GAN 的判别器相结合, 修复真实且保留身份的人脸。这类方法的缺陷是在生成遮挡区域边缘时会失真。

但是这些方法并没有专门针对口罩遮挡, 因而去除口罩遮挡的效果不佳。口罩遮挡面积较大, 鼻子以下的面部特征被掩盖, 可利用的特征关键点减少。考虑人脸边缘包含了大量的人脸结构信息, 如果以边缘为基础修复人脸, 有可能产生更好的效果。因此, 本文提出新的人脸口罩区域修复算法 ID-EFCGAN (ID-edge to face conditional generative adversarial network), 利用条件生成对抗网络 (conditional generative adversarial network, CGAN)^[15]生成边缘图, 在此基础上再利用 CGAN 恢复被遮挡区域的人脸。为提升 ID-EFCGAN 的性能, 提出空间加权对抗损失和身份一致性损失训练上述网络, 其中身份一致性损失通过约束生成人脸与真实人脸有相似的特征表示, 使生成人脸与真实人脸的身份信息一致, 从而保证人脸识别效果。利用关键点信息构建了两个佩戴口罩的人脸数据集进行实验, 结果表明, 佩戴口罩的人脸经过本文的去遮挡处理后, 可使 ArcFace 的识别准确率达到 98.39%。



2 口罩遮挡下的人脸识别

针对佩戴口罩的人脸, 本文不改变现有人脸识别算法, 采取修复口罩遮挡部分的思路, 提高人脸识别准确率, 具体过程如图 1 所示。本文提出端到端的人脸修复模型 ID-EFCGAN (以下简称 ID-EFC), 由边缘生成网络和区域填充网络构成, 分为两步去除口罩遮挡: (1) 恢复人脸边缘。对于戴口罩的人脸图片, 利用人脸关键点信息定位口罩, 并生成口罩掩码。借助口罩掩码获得口罩区域缺失的人脸图像。利用边缘生成网络恢复口罩缺失区域的边缘。(2) 口罩缺失区域填充。基于恢复的边缘图, 区域填充网络还原遮挡区域的人脸。最后, 利用已有的人脸识别算法对去除口罩遮挡的人脸进行识别。

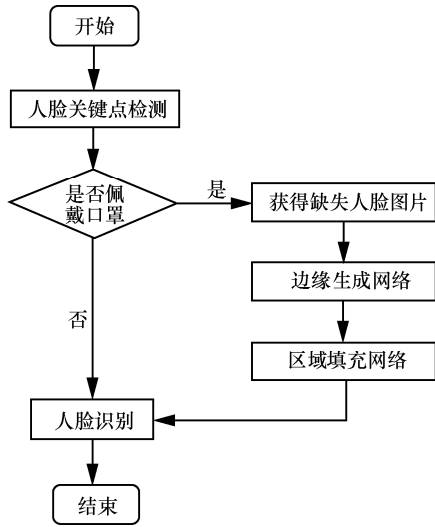


图 1 去除口罩遮挡实现人脸识别的流程

边缘生成网络和区域填充网络均采用 CGAN 的结构, 由生成器和判别器组成。为方便后续描述, 将边缘生成网络的生成器和判别器分别记为 G_1 和 D_1 , 区域填充网络的生成器和判别器分别记为 G_2 和 D_2 。用 C 表示人脸边缘图, X 表示人脸图, 将戴口罩的人脸图片记为 X_{wear} , 借助口罩掩码获得的口罩区域缺失的人脸图记为 X_{mask} , 真实的无遮挡人脸图记为 X_{gt} , 其灰度图和边缘图分别记为 X_{gray} 和 C_{gt} 。

2.1 边缘生成网络

为了生成口罩区域的人脸边缘, 需先定位口罩并获得口罩区域缺失的人脸图。利用人脸关键点信息, 结合先验知识初步获取口罩的位置和形状。由于每张图片中的头部姿势和脸部大小不同, 在算法中对口罩形状的大小以及角度进行自动调整。根据获取的口罩位置和形状信息, 生成与人脸图片同样大小的口罩掩码, 记为 M 。该掩码口罩区域像素值为 1, 其余区域像素值为 0。相比用矩形框定位口罩, 本文方法更加精准, 不会过度扩大遮挡面积。利用式 (3) 获得口罩区域缺失的人脸图 X_{mask} 。

$$X_{mask} = X_{wear} \odot (1 - M) \quad (3)$$

其中, $\odot$ 为 Hadamard 乘积。

在获得 X_{mask} 后, 进一步利用边缘生成网络还原缺失区域的人脸边缘, 通过生成器 G_1 与判别器 D_1 的动态博弈, 生成尽可能真实的无遮挡边缘图 (如图 2 所示)。具体步骤为:

(1) 获取 X_{mask} 的灰度图和边缘图, 分别记为 X_{gray_mask} 和 C_{mask} , 其中, 边缘图利用 Canny 边缘检测器^[16]获取;

(2) 将获得的灰度图 X_{gray_mask} 输入生成器 G_1 , 边缘图 C_{mask} 作为生成器 G_1 的条件, 生成无遮挡的边缘图 C_{pred} ;

(3) 获得无遮挡边缘图 C_{pred} 后, 判别器 D_1 不仅判断 C_{pred} 是来自真实数据还是生成器 G_1 , 而且判断生成的 C_{pred} 与灰度图 X_{gray} 是否匹配;

(4) 利用式 (4) 获得原始的无遮挡边缘图 C_{gt} 和 C_{pred} 的组合, 记为 C_{comp} :

$$C_{comp} = C_{gt} \odot (1 - M) + C_{pred} \odot M \quad (4)$$

C_{comp} 中的无遮挡区域的边缘来自 C_{gt} , 遮挡区域的边缘来自生成器的预测 C_{pred} 。

2.2 区域填充网络

在获得组合边缘图 C_{comp} 后, 进一步利用区域填充网络填充 C_{comp} 生成人脸, 通过生成器 G_2 和判

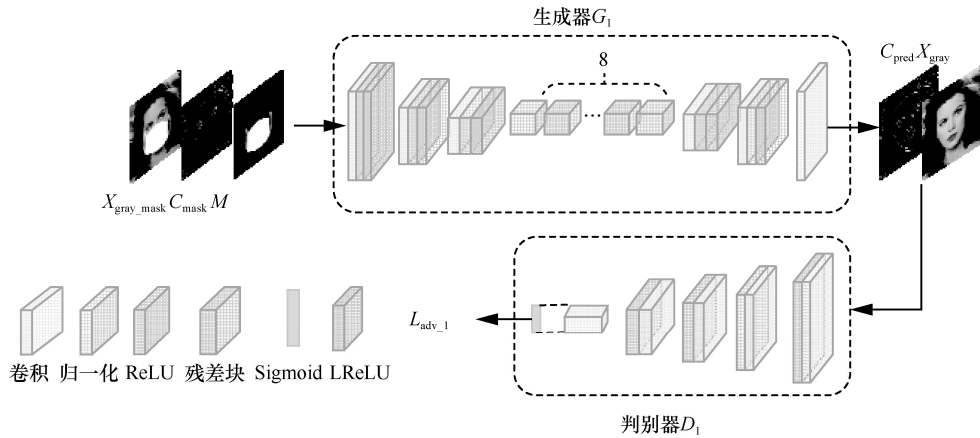


图2 边缘生成网络的结构

别器 D_2 的动态博弈，生成尽可能真实的无遮挡人脸图（如图3所示）。具体过程为：

(1) 将缺失的人脸图 X_{mask} 和组合边缘图 C_{comp} 输入生成器 G_2 ，其中 C_{comp} 作为生成器 G_2 的条件，生成无遮挡的人脸图片 X_{pred} ；

(2) 获得无遮挡人脸图 X_{pred} 后，判别器 D_2 不仅辨别出生成器 G_2 “伪造”的 X_{pred} ，而且判断生成的 X_{pred} 与边缘图 C_{comp} 是否匹配；

(3) 借助式(5)获得原始的无遮挡人脸图 X_{gt} 和 X_{pred} 的组合，记为 X_{comp} ：

$$X_{\text{comp}} = X_{\text{gt}} \odot (1 - M) + X_{\text{pred}} \odot M \quad (5)$$

X_{comp} 中的无遮挡区域来自 X_{gt} ，遮挡区域来自 G_2 生成的 X_{pred} 。在获得无遮挡人脸图片 X_{comp}

后即可进行识别。

2.3 空间加权对抗损失

图4(a)展示了遮挡区域与未遮挡区域对抗损失函数的分布，可以看到两个区域的分布存在差异，无差别地对待这两个区域可能会影响生成图片的质量。为此，本文引入空间加权对抗损失。考虑遮挡可看作大面积空间连续噪声^[5]，本文对口罩遮挡区域和未遮挡区域的对抗损失赋予不同的权重，定义空间加权对抗损失为：

$$L_{\text{adv}_1} = \mathbf{W} \cdot E_{(C_{\text{gt}}, X_{\text{gray}})} [\lg D_1(C_{\text{gt}}, X_{\text{gray}})] + \mathbf{W} \cdot E_{X_{\text{gray}}} \lg [1 - D_1(C_{\text{pred}}, X_{\text{gray}})] \quad (6)$$

其中， $\mathbf{W}$ 称为空间权重，大小为 $N \times N$ 的矩阵，其中每个元素表示相应区域损失的权重。 $\mathbf{W}$ 中的

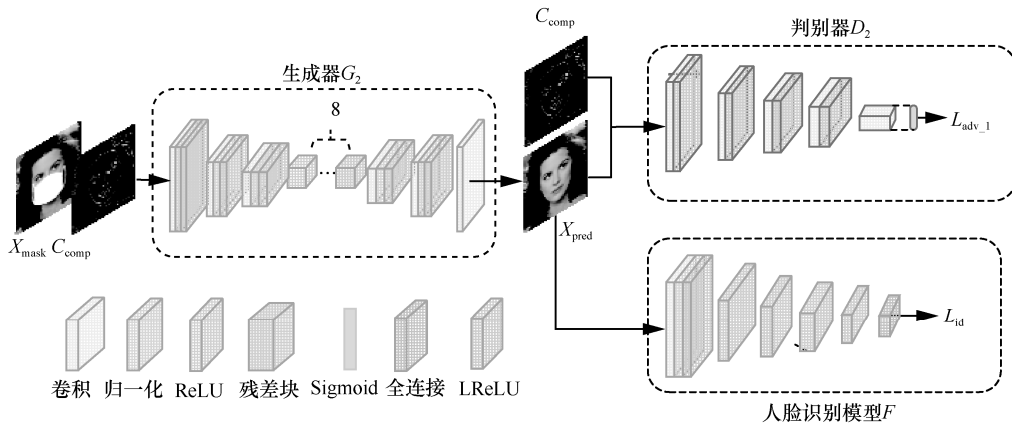
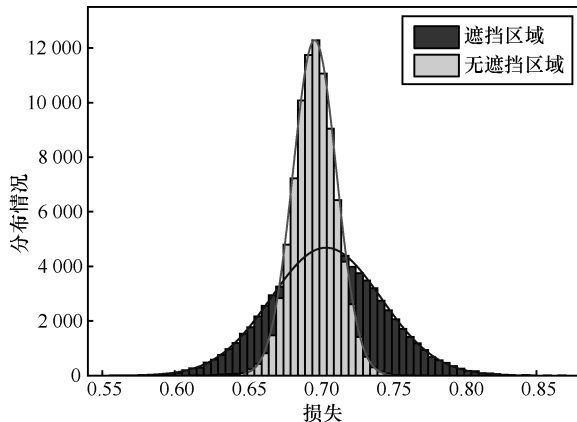


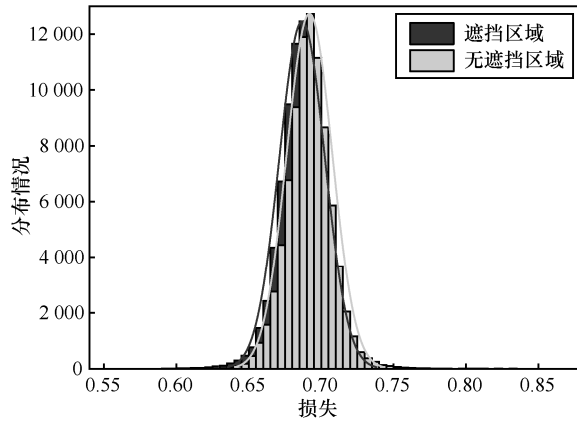
图3 区域填充网络的结构



元素仅有两种取值, 分别为口罩区域对应的权重取值和未遮挡区域对应的权重取值 (如口罩区域对应的权重均为 0.75, 未遮挡区域对应的权重均为 1), 该部分的具体取值将在实验部分讨论。引入空间加权对抗损失后, 两个区域的分布更加接近, 如图 4(b)所示, 这表明空间加权对抗损失更加合理地对待这两个区域。



(a) 引入空间权重前的对抗损失分布



(b) 引入空间权重后的对抗损失分布

图4 被口罩遮挡区域与无遮挡区域对抗损失函数的分布

2.4 身份一致性损失

去除人脸口罩遮挡后的人脸识别准确率取决于:

- (1) 生成的人脸要尽可能真实;
- (2) 生成人脸 X_{pred} 与真实人脸 X_{gt} 的身份要一致。

为此, 本文在 GAN 的结构中引入一个训练好的识别网络, 用于计算身份一致性损失, 度量生成人脸与真实人脸的身份特征之间的距离, 使生

成人脸与真实人脸在语义上更加相似:

$$L_{\text{id}} = \|F(X'_{\text{gt}}) - F(X'_{\text{pred}})\|_2^2 \quad (7)$$

其中, F 为预训练好的人脸识别模型, 提取身份判别所需的人脸特征; X'_{gt} 和 X'_{pred} 分别是经过人脸检测器 MTCNN (multi-task convolutional neural network) [17] 裁剪对齐后的真实人脸和生成人脸。通过最小化身份一致性损失, 真实人脸的身份特征引导着人脸的生成, 使得生成人脸在身份特征空间中逼近真实人脸, 从而保留大量身份信息。

为了实现边缘图到人脸图这两种不同风格图像之间的转换, 本文引入感知损失 [18] 与风格损失 [19], 与本文设计的身份一致性损失、空间加权对抗损失加权求和作为总的损失函数训练区域填充网络。风格损失通过计算预训练网络不同层特征图之间的距离, 衡量两张图片之间高级感知及语义差别:

$$L_{\text{perc}} = E \left[\sum_i \frac{1}{N_i} \|\phi_i(X_{\text{gt}}) - \phi_i(X_{\text{pred}})\|_1 \right] \quad (8)$$

其中, N_i 是第 i 层激活映射的元素个数, ϕ_i 是 VGG-19 网络的第 i 层激活映射。风格损失利用表示风格特征的 Gram 矩阵 [20], 计算风格特征之间的绝对误差, 衡量两张图片的风格差异。风格损失可以有效避免棋盘效应, 对于大小为 $C_i \times H_j \times W_j$ 的特征映射, 风格损失函数定义为:

$$L_{\text{style}} = E \left[\left\| \phi_j^{\phi}(X_{\text{pred}}) - \phi_j^{\phi}(X_{\text{gt}}) \right\|_1 \right] \quad (9)$$

其中, ϕ_j^{ϕ} 是由激活映射 ϕ_i 构造的大小为 $C_j \times C_j$ 的 Gram 矩阵。

为了提高生成图片的质量, 在区域填充网络的损失函数中引入本文提出的空间加权对抗损失函数, 对口罩区域的损失赋予与其他位置不同的权重。区域填充网络的空间加权对抗损失:

$$L_{\text{adv}_2} = W \cdot E_{(X_{\text{gt}}, C_{\text{comp}})} [\lg D_2(X_{\text{gt}}, C_{\text{comp}})] + W \cdot E_{C_{\text{comp}}} [\lg [1 - D_2(X_{\text{pred}}, C_{\text{comp}})]] \quad (10)$$

最终得到区域填充网络生成器 G_2 和判别器 D_2 的损失分别如下:

$$L_{G_2} = \lambda_{adv,2} L_{adv,2} + \lambda_p L_{perc} + \lambda_s L_{style} + \lambda_{id} L_{id} \quad (11)$$

$$L_{D_2} = -L_{adv,2} \quad (12)$$

利用 L_{G_2} 训练区域填充网络生成器 G_2 , 利用 L_{D_2} 训练区域填充网络判别器 D_2 。

3 实验设置和效果评估

首先, 定性和定量地分析生成图片的质量; 然后, 分析空间加权对抗损失的有效性; 最后, 从人脸识别的准确率角度评估本文模型是否更好地保留了身份信息。

3.1 实验设置

目前用于人脸识别的公共数据集主要有 CASIA-WebFace^[21] 和 LFW^[2]。CASIA-WebFace 用于身份鉴定和人脸识别模型的训练和测试, 包含 10 575 个身份的 494 414 张图片, 每张图片的大小为 250 dpi×250 dpi。LFW 用于人脸识别模型的测试, 包含 5 749 个身份的 13 233 张人脸图像, 绝大部分人仅有一张图片, 每张图片的大小为 250 dpi×250 dpi。但以上两个数据集中的人脸均未戴口罩, 为此本文借助关键点信息, 将口罩添加到人脸图片上, 合成了两个戴口罩的数据集 CASIA-WebFace mask 和 LFW mask。CASIA-WebFace mask 包含约 42 万张图片, 其中 34 万张用于训练, 4 万张用于验证, 其余用于测试; LFW mask 包含 13 170 张人脸图片, 用于测试。测试集中的人脸身份均没有在训练集中出现。

本文所有实验均在 4×GeForce RTX 2 080ti 服务器上, 生成器以参考文献[18]为基础构建, 判别器使用 PatchGAN^[22], 网络的所有层都使用实例归一化 (instance normalization, IN)。实验使用的人脸识别模型为 ArcFace。

设置 Candy 边缘检测器的参数 σ 为 2, 损失函数的权重分别为 $\lambda_{adv,1} = 1$, $\lambda_{adv,2} = \lambda_p = 0.1$,

$\lambda_s = 250$, $\lambda_{id} = 1$ 。分两个阶段训练: 第一阶段, 分别训练生成器 G_1 和 G_2 直到收敛, 设置生成器 G_1 和判别器 D_1 的学习率分别为 10^{-4} 和 10^{-5} , 使用 Adam 优化器优化, 设置参数 $\beta_1 = 0$, $\beta_2 = 0.9$ 。第二阶段, 联合训练生成器 G_1 和 G_2 , 设置生成器和判别器的学习率分别为 10^{-5} 和 10^{-6} 。

3.2 生成图片的质量

本文分别展示边缘生成网络和区域填充网络修复的边缘图 (如图 5 所示) 和人脸图 (如图 6 所示)。图 5 中从左往右各列依次为真实的人脸图 X_{gt} 、戴口罩的人脸图 X_{wear} 、口罩区域缺失的人脸灰度图 X_{gray_mask} 、口罩区域缺失的人脸边缘图 C_{mask} 、预测的人脸边缘图 C_{pred} 、组合边缘图 C_{comp} 和真实的人脸边缘图 C_{gt} 。图 6 中从左往右各列依次为真实的人脸图 X_{gt} 、戴口罩的人脸图 X_{wear} 、口罩区域缺失的人脸图 X_{mask} 、组合边缘图 C_{comp} 、预测的人脸图 X_{pred} 和组合人脸图 X_{comp} 。从图 5 和图 6 可以看出, 基于人脸边缘图恢复了被口罩遮挡的区域, 并保留了原图中未被遮挡的区域。

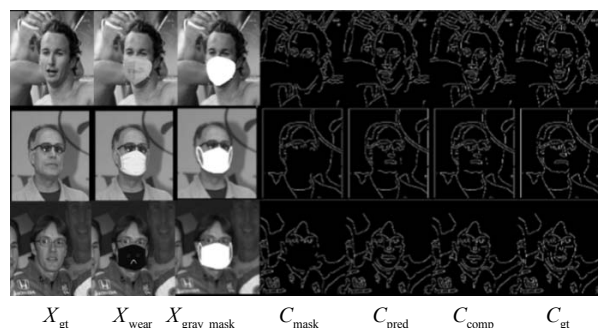


图 5 边缘生成网络生成的人脸边缘图示例

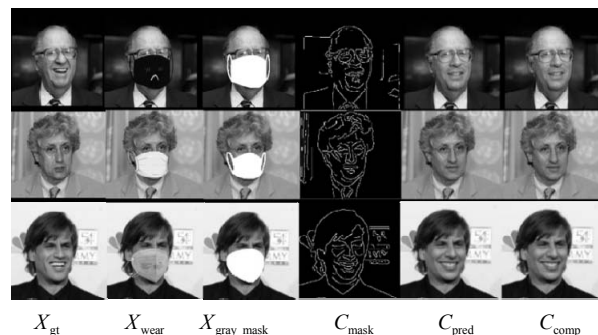


图 6 区域填充网络生成的人脸图示例



本文从视觉效果定性分析去除口罩后的图片质量。图 7 和图 8 展示了人脸头部姿势、光照和口罩样式不同的人脸图片。第一行是姿势和光线不同的戴口罩人脸图片，第二行是未戴口罩的真实人脸，第三行是 ID-EFC 去除口罩后的人脸图片，接近第二行的真实人脸。实验结果表明，ID-EFC 的有效性不受头部姿势、光照和口罩样式的影响。尽管测试数据集和训练数据集中的人物身份无一相同，包括头部姿势、脸部大小和口罩的遮挡形状也不相同，但 ID-EFC 仍能生成没有虚假边缘和区域的人脸图片，恢复大部分细节，取得良好的视觉效果。



图 7 ID-EFC 在 CASIA-WebFace mask 数据集上去口罩遮挡人脸示例



图 8 ID-EFC 在 LFW mask 数据集上去口罩遮挡的人脸示例

本文以未引入空间加权对抗损失和身份一致性损失的模型为基准模型，命名为 EFCGAN(edge to face conditional generative adversarial network)，简称为 EFC，将其与 ID-EFC 进行实验对比。图 9

给出了各种模型在 CASIA-WebFace mask 数据集上去除口罩遮挡后的部分结果。图 9 中第一行和最后一行分别为戴口罩的人脸 X_{wear} 和真实人脸 X_{gt} ，中间三行中从上往下依次为使用矩形掩码训练的模型 EFC_{roc} 、使用本文口罩掩码训练的模型 EFC_{mask} 和 ID-EFC 去除口罩后的人脸。可以看出，使用 EFC_{roc} 得到的人脸不够真实，容易产生虚假区域。使用 EFC_{mask} 去遮挡的效果明显提升，人脸变得真实，表明本文生成口罩掩码的有效性。但 EFC_{mask} 修复的人脸在嘴巴和鼻子等关键部位有些失真，因为该模型未充分考虑身份信息。相比之下，ID-EFC 修复的鼻子和嘴巴部位更真实、自然，这是因为引入了身份一致性损失，有利于身份细节的恢复。



图 9 不同模型在 CASIA-WebFace mask 数据集上去除口罩的人脸示例

本文进一步用 PSNR 和 SSIM 这两个常用的图像质量评价指标定量分析修复图片的质量。各模型去除口罩遮挡后的图片对应的 PSNR 和 SSIM 见表 1。从表 1 中的结果看， EFC_{roc} 的两项指标均为最低，ID-EFC 的两项指标最高，而 EFC_{mask} 的两项指标居中，说明 ID-EFC 生成的人脸图片质量最好，这与之前的定性分析结论是一致的。

表 1 不同模型生成图片的 PSNR 和 SSIM

对比项	CASIA-WebFace mask			LFW mask		
	EFC_{roc}	EFC_{mask}	ID-EFC	EFC_{roc}	EFC_{mask}	ID-EFC
SSIM	89.79%	94.38%	95.43%	94.41%	96.12%	97.57%
PSNR	26.58%	29.03%	30.20%	29.05%	30.90%	31.98%

3.3 空间加权对抗损失的效果

下面将进一步验证空间加权对抗损失对提升图片质量的效果,对比引入空间加权对抗损失 L_{adv} (包括 L_{adv_1} 和 L_{adv_2})前后EFC_{mask}模型在CASIA-WebFace mask和LFW mask上的PSNR和SSIM值(见表2)。表2中引入 L_{adv} 的EFC_{mask}模型修复图片的PSNR和SSIM值均较高,说明 L_{adv} 对提升图片质量有较好的效果。为了进一步研究口罩区域损失的权重取值对生成图片质量的影响,固定无遮挡区域的权重均为1,口罩区域的权值 w_{mask} 分别取0.5、0.75、1、1.25和1.5,实验结果如图10所示。权重 w_{mask} 取值为1时,表示模型EFC_{mask}的对抗损失没有引入空间权重。从图10可以看出,权重小于1时的PSNR高于权重大于1时的PSNR,并且在权重为0.75时,PSNR最高。因此,本文选取口罩区域对抗损失的权重为0.75,无遮挡区域对抗损失的权重为1。

3.4 人脸识别效果

通过对比添加身份一致性损失 L_{id} 和空间加权对抗损失 L_{adv} (包括 L_{adv_1} 和 L_{adv_2})前后模型对遮挡下人脸识别的影响,证明ID-EFC能更好地保留身份信息。首先将人脸对齐裁剪成尺寸为112 dpi×112 dpi的图片,用ArcFace提取裁剪对齐后的人脸特征,然后验证成对人脸图片是否属于同一

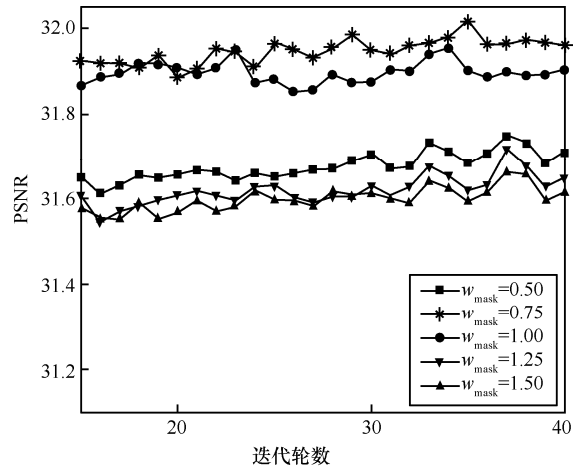


图10 口罩区域对抗损失权重取值对PSNR的影响

身份。成对图片中一个是图库中未戴口罩的人脸,另一个是戴口罩的人脸或去除口罩遮挡的人脸。使用模型去除口罩后的人脸进行识别的准确率见表3。

表3中最后一列,ArcFace在LFW数据集上的准确率高达99.82%,但是当人脸存在口罩遮挡,对遮挡人脸不进行任何处理,直接基于整张遮挡人脸进行识别时,其在LFW mask数据集上的准确率只有94.26%,下降了5.56%,说明口罩遮挡的确会影响人脸识别准确率。以ArcFace在LFW mask数据集上的准确率94.26%为比较基准,从左往右对比:ArcFace在EFC_{roc}修复的人脸数据上的准确率只有93.87%,低于直接识别整张遮挡人脸的准确率(94.26%);ArcFace在EFC_{mask}修复的

表2 空间加权对抗损失使用前后的结果对比

对比项	CASIA-WebFace mask		LFW mask	
	EFC _{mask}	EFC _{mask} + L_{adv}	EFC _{mask}	EFC _{mask} + L_{adv}
SSIM	94.38%	95.20%	96.12%	97.43%
PSNR	29.03%	29.98%	30.90%	31.94%

表3 不同模型在LFW mask数据集上的人脸识别准确率

数据集	EFC _{roc}	EFC _{mask}	EFC _{mask} + L_{adv}	EFC _{mask} + L_{id}	EFC _{mask} + L_{adv} + L_{id}	ArcFace
LFW	—	—	—	—	—	99.82%
LFW mask	93.87%	97.98%	98.06%	98.14%	98.39%	94.26%



人脸数据上的准确率为 97.98%，较遮挡下人脸识别准确率提高了 3.72%； EFC_{mask} 加入 L_{adv} 后，准确率提高了 3.8%；加入 L_{id} 后，准确率提高了 3.88%。而同时引入 L_{adv} 和 L_{id} 后，ID-EFC 将遮挡下人脸识别准确率提升了 4.13%，最终准确率为 98.39%，说明 ID-EFC 的模型能更好地保留身份信息。

图 11 更直观地展示了 ArcFace 在各个模型去除口罩后的人脸数据上进行人脸识别的 ROC 曲线。GT 和 OCC 分别表示无口罩遮挡和口罩遮挡下人脸识别的 ROC 曲线。随着空间加权损失与身份一致性损失的加入，AUC 值不断增大，说明人脸识别的效果变好。通过 AUC 值可以发现，遮挡条件下，ID-EFC ($EFC_{mask} + L_{adv} + L_{id}$) 对提高识别准确率的效果最好。

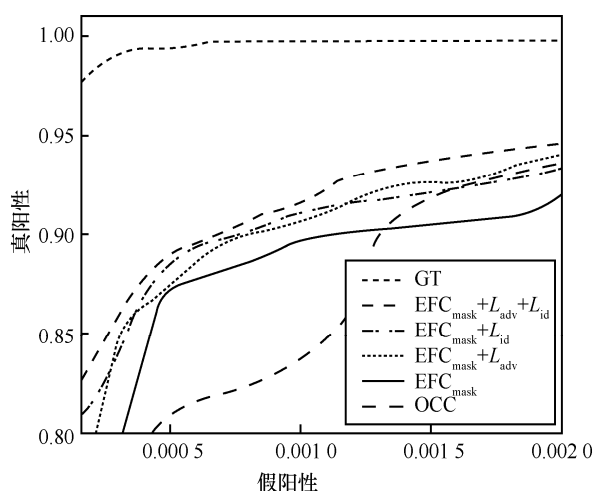


图 11 在 LFW mask 数据集上人脸识别 ROC 曲线

本文通过比较不同模型修复的人脸的身份特征，进一步分析 ID-EFC 取得较好效果的原因。利用 t-SNE^[23]将 512 维的身份特征降至 2 维，观察属于同一身份特征的聚集性与可分性，可视化结果如图 12 所示。图 12(a)(b)(c)分别表示口罩遮挡下的人脸、 EFC_{mask} 修复的人脸和 ID-EFC 修复的人脸的身份特征。图中同一形状的点表示属于同一个身份的特征。由图 12(a)可发现，口罩遮挡下不同身份的人脸在身份特征上存在重叠，因此即

使性能优良的 ArcFace 也无法在遮挡的情况下保持高精度的准确率。用 EFC_{mask} (如图 12(b)所示) 去除遮挡后的人脸特征也仍然存在部分重叠。而 ID-EFC (如图 12(c)所示) 修复的人脸特征实现了最大的聚集度，容易根据身份特征对其进行分组，显示了身份一致损失对提高口罩遮挡下人脸识别准确率的有效性。

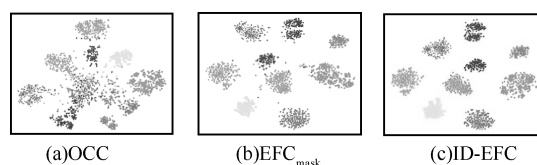


图 12 身份特征可视化

3.5 口罩类型对识别效果的影响

为了进一步研究不同类型的口罩遮挡对人脸识别的影响，本文分别测试不同类型口罩遮挡下的识别准确率，结果见表 4。本文使用黑、白和蓝 3 种口罩，对人脸的遮挡面积分别约为 40%、39% 和 37%，在这 3 种口罩遮挡下，ArcFace 的识别准确率分别为 91.23%、95.77% 和 93.18%，见表 4。其中遮挡面积为 40% 的黑色口罩遮挡下的识别准确率最低，说明该种口罩对人脸识别效果的影响较大。

表 4 不同类型口罩遮挡下的人脸识别准确率

口罩类型(遮挡面积)	黑色(40%)	白色(39%)	蓝色(37%)
人脸识别准确率	91.23%	95.77%	93.18%

为了进一步分析遮挡面积和口罩颜色对识别效果的影响，本文分别测试不同颜色 and 不同面积的口罩遮挡下的识别准确率，结果见表 5。

表 5 不同面积和颜色的口罩遮挡下人脸识别准确率

遮挡面积	口罩颜色		
	黑色	白色	蓝色
30%	94.21%	97.13%	95.44%
35%	92.61%	96.24%	93.68%
40%	91.23%	95.56%	92.72%

可以发现,在口罩颜色相同的条件下,识别准确率随着遮挡面积增加而逐渐下降,这说明遮挡面积对识别效果有一定的影响。除此之外,在遮挡面积相同的情况下,黑、白、蓝 3 种口罩遮挡会对识别效果产生不同程度的影响,以 35%遮挡面积为例,黑、白、蓝 3 种口罩遮挡下的识别准确率分别为 92.61%、96.24%和 93.68%,说明黑色口罩遮挡对人脸识别模型性能的影响最为严重,蓝色次之,白色最小。

4 结束语

本文提出了一个人脸口罩区域修复深度学习模型 ID-EFC,并设计了空间加权对抗损失和身份一致性损失对其训练。实验表明,空间加权对抗损失的引入能有效提升生成图片的质量;身份一致性损失的引入能更好地保留生成图片的身份信息,这两种损失在口罩区域的修复中起着重要作用。ID-EFC 的应用,可提高口罩遮挡下的人脸识别准确率。此外,ID-EFC 能够处理不同头部姿势、光照和口罩样式的人脸图片,具有良好的泛化性能。下一步的工作是优化现有模型,推广到多种遮挡类型的人脸识别任务中。

参考文献:

- [1] TAIGMAN Y, YANG M, RANZATO M, et al. DeepFace: closing the gap to human-level performance in face verification[C]//Proceedings of 2014 IEEE Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE Press, 2014: 1701-1708.
- [2] HUANG G B, MATTAR M, BERG T, et al. Labeled faces in the wild: a database for studying face recognition in unconstrained environments[EB]. 2008.
- [3] DENG J K, GUO J, XUE N N, et al. ArcFace: additive angular margin loss for deep face recognition[C]//Proceedings of 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Piscataway: IEEE Press, 2019: 4685-4694.
- [4] DUAN Q Y, ZHANG L. BoostGAN for occlusive profile face frontalization and recognition[EB]. 2019.
- [5] ZHAO F, FENG J S, ZHAO J, et al. Robust LSTM- autoencoders for face de-occlusion in the wild[J]. IEEE Transactions on Image Processing, 2018, 27(2): 778-790.
- [6] ZHAO Y J, CHEN W K, XING J, et al. Identity preserving face completion for large ocular region occlusion[EB]. 2018.
- [7] WANG M, HU Z P, SUN Z, et al. Varying face occlusion detection and iterative recovery for face recognition[J]. Journal of Electronic Imaging, 2017(26): 033009.
- [8] ZHANG J, KAN M N, SHAN S G, et al. Occlusion-free face alignment: deep regression networks coupled with de-corrupt AutoEncoders[C]//Proceedings of 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Piscataway: IEEE Press, 2016: 3428-3437.
- [9] YUAN X W, PARK I K. Face de-occlusion using 3D morphable model and generative adversarial network[C]//Proceedings of 2019 IEEE/CVF International Conference on Computer Vision (ICCV). Piscataway: IEEE Press, 2019: 10061-10070.
- [10] GOODFELLOW I, POUGET-ABADIE J, MIRZA M, et al. Generative adversarial nets[C]//Proceedings of the Advances in Neural Information Processing Systems. New York: IncPress, 2014: 2672-2680.
- [11] HE Z L, ZUO W M, KAN M N, et al. AttGAN: facial attribute editing by only changing what you want[J]. IEEE Transactions on Image Processing: a Publication of the IEEE Signal Processing Society, 2019, 28(11): 5464-5478.
- [12] IIZUKA S, SIMO-SERRA E, ISHIKAWA H. Globally and locally consistent image completion[J]. ACM Transactions on Graphics, 2017, 36(4): 1-14.
- [13] LI Y J, LIU S F, YANG J M, et al. Generative face completion[C]//Proceedings of 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Piscataway: IEEE Press, 2017: 5892-5900.
- [14] GE S, LI C, ZHAO S, et al. Occluded face recognition in the wild by identity-diversity inpainting[J]. IEEE Transactions on Circuits and Systems for Video Technology, 2020, 30(10): 3387-3397.
- [15] MIRZA M, OSINDERO S. Conditional generative adversarial nets[J]. arXiv:1411.1784, 2014.
- [16] CANNY J. A computational approach to edge detection[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 1986, 8 (6): 679-698.
- [17] ZHANG K P, ZHANG Z P, LI Z F, et al. Joint face detection and alignment using multitask cascaded convolutional networks[J]. IEEE Signal Processing Letters, 2016, 23(10): 1499-1503.
- [18] JOHNSON J, ALAHI A, LI F F. Perceptual losses for real-time style transfer and super-resolution[M]//Computer Vision – ECCV 2016. Cham: Springer International Publishing, 2016: 694-711.
- [19] GATYS L A, ECKER A S, BETHGE M. Image style transfer

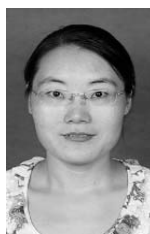


- using convolutional neural networks[C]//Proceedings of 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Piscataway: IEEE Press, 2016: 2414-2423.
- [20] BARTH N. The gramian and k-volume in n-space: some classical results in linear algebra[J]. Journal of Young Investigators, 1999, 2(1): 1-4 .
- [21] YI D, LEI Z, LIAO S, ET AL. Learning face representation from scratch[J]. arXiv:1411.7923, 2014 .
- [22] ISOLA P, ZHU J Y, ZHOU T H, et al. Image-to-image translation with conditional adversarial networks[C]//Proceedings of 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Piscataway: IEEE Press, 2017: 5967-5976.
- [23] MAATEN L, HINTON G. Visualizing data using t-SNE[J]. Journal of Machine Learning Research, 2008, 9(3):2579-2605.

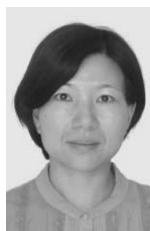
[作者简介]



李悦(1996-), 女, 浙江科技学院硕士生, 主要研究方向为深度学习、计算机视觉。



关晓惠(1977-), 女, 浙江水利水电学院副教授, 主要研究方向为数字图像处理与模式识别。



李蔚(1978-), 女, 浙江科技学院副教授, 主要研究方向为科学计算、模式识别和计算机视觉中的机器学习问题。



王滨(1977-), 男, 杭州海康威视网络与信息安全实验室研究员, 主要研究方向为物联网安全、人工智能安全和网络安全。



钱亚冠(1976-), 男, 浙江科技学院副教授, 主要研究方向为深度学习、人工智能安全和大数据处理。



顾钊铨(1989-), 男, 广州大学网络空间先进技术研究院教授, 主要研究方向为无线网络、分布式计算、大数据分析和人工智能安全。