

## 基于循环生成对抗网络的含遮挡人脸识别

徐润昊, 程吉祥, 李志丹, 付小龙

(西南石油大学 电气信息学院, 成都 610500)

**摘要:** 人脸图像的遮挡会严重影响人脸识别准确率, 当前处理带遮挡人脸识别的方法主要有丢弃法和修复法两种。丢弃法因忽略或丢弃大量遮挡区域的有效特征易造成识别准确率不高, 而当前大多数修复法需要原图的相关信息, 限制了其应用。针对现有含遮挡人脸识别方法存在的问题, 提出一种基于循环生成对抗网络的人脸识别算法, 通过利用2对生成器和判别器进行循环训练, 实现遮挡人脸图像的盲修复, 此过程不需要除遮挡区域外的原图信息。在此基础上, 采用ResNet-50网络对修复后人脸进行识别, 该网络通过跳跃连接消除深层卷积神经网络中的网络退化问题, 能够降低网络训练的难度, 且不会增加额外的参数和计算量。针对盲修复后人脸特征存在类内差异大和类间差异小的特性, 引入一种能够量化类间距离的分类损失函数RegularFace作为识别网络损失函数。实验结果表明, 与DCGAN+CNN算法相比, 所提算法对不同遮挡类型和遮挡区域的人脸图像识别准确率均有所提高, 当线性遮挡面积为40%时, 所提算法的识别准确率提高了14.4个百分点。

**关键词:** 含遮挡人脸; 盲修复; 人脸识别; 循环生成对抗网络; RegularFace损失函数

开放科学(资源服务)标志码(OSID):



**中文引用格式:** 徐润昊, 程吉祥, 李志丹, 等. 基于循环生成对抗网络的含遮挡人脸识别[J]. 计算机工程, 2022, 48(5): 289-296, 305.

**英文引用格式:** XU R H, CHENG J X, LI Z D, et al. Face recognition with occlusion based on cyclic generative adversarial networks[J]. Computer Engineering, 2022, 48(5): 289-296, 305.

## Face Recognition with Occlusion Based on Cyclic Generative Adversarial Networks

XU Runhao, CHENG Jixiang, LI Zhidan, FU Xiaolong

(School of Electrical Information, Southwest Petroleum University, Chengdu 610500, China)

**[Abstract]** Face image occlusion has an important impact on the accuracy of face recognition. Currently, two primary methods are used to process face recognition with occlusion: discard and repair. The discarding method ignores (discards) the effective features of several occluded areas, which easily leads to a low recognition accuracy. However, most of the current restoration methods require a known original image with related information for restoration, which limits its application. Aiming at the problems existing in the existing face recognition methods with occlusion, this paper proposes a face recognition method based on cyclic Generative Adversarial Networks (GAN). This method first uses two pairs of generators and discriminators to cyclically train to realize the blind restoration of the occluded face image. This process does not require information of the original image, except the occluded area. On this basis, a ResNet-50 network recognizes the repaired face after, eliminating the network degradation problem in deep convolutional neural networks using jump connections, which can reduce the difficulty of network training without additional parameters and computation. Considering the large and small inter-class differences in face features after blind reparation, a classification loss function that can quantify the distance between classes is introduced as the recognition network loss function. Compared with the DCGAN+CNN algorithm, the results show that, the recognition accuracy of the proposed algorithm improved for face images with different occlusion types and occlusion areas, and, when the linear occlusion area is 40%, the recognition accuracy of the proposed algorithm improved by 14.4 percentage points.

**[Key words]** face with occlusion; blind restoration; face recognition; cyclic Generative Adversarial Network (GAN); RegularFace loss function

**DOI:** 10.19678/j.issn.1000-3428.0061124

**基金项目:** 国家自然科学基金(61603319, 61601385); 西南石油大学智能控制与图像处理青年科技创新培育团队资助项目(2017CXTD010)。

**作者简介:** 徐润昊(1997—), 男, 硕士研究生, 主研方向为计算机视觉、人脸识别; 程吉祥、李志丹, 副教授、博士; 付小龙, 硕士研究生。

**收稿日期:** 2021-03-15 **修回日期:** 2021-05-08 **E-mail:** 877219132@qq.com

## 0 概述

运用图像处理技术进行人脸识别以获取身份信息在实际生活中具有重要意义。该技术已被广泛应用于公共安全监控、移动支付、警方取证等领域。人脸识别的发展主要分为3个阶段<sup>[1]</sup>:第1阶段主要研究人脸识别所需要的面部几何特征;第2阶段是人机交互阶段,主要将人脸面部特征用多维特征矢量表示;第3阶段是机器自动学习阶段,推动人脸识别进行实用化。目前,正常情境下的完整人脸识别准确率已经满足一定场景需要,然而实际人脸表情和姿态具有复杂性和易变性,且帽子、口罩、围巾等物体易对人脸产生遮挡,检测到的人脸不一定完好无缺,这易造成人脸识别准确率急剧下降,限制该技术在复杂场景下的应用<sup>[2]</sup>。因此,研究有遮挡的人脸识别对于拓展人脸识别应用具有重要意义。

本文提出一种基于循环生成对抗网络的含遮挡人脸识别算法,通过构建一种改进的循环生成对抗网络对遮挡人脸进行盲修复,该算法不需要使用遮挡区域的信息,仅通过2个生成器和2个判别器循环训练,就能提高生成人脸与原遮挡人脸的特征重合度及丰富有效人脸特征。在此基础上,构建基于ResNet-50<sup>[3]</sup>的卷积网络,并对修复后的人脸图像进行再识别。通过令RegularFace损失函数<sup>[4]</sup>与softmax损失函数配合使用来处理修复后人脸可能存在的特征融合问题。

## 1 相关工作

研究有遮挡的人脸识别一般有2个思路。一种是对遮挡人脸进行有效的人脸提取,然后再根据提取的有效人脸特征进行人脸识别。鲁棒的特征提取方法是对人脸图像中颜色、亮度等低阶特征和姿态、表情等高阶特征进行分解<sup>[5]</sup>。传统的鲁棒特征提取方法如梯度脸<sup>[6]</sup>,通过将人脸图像处理成梯度直方图,去除不必要的信息,并保留关键信息。FPH框架<sup>[7]</sup>则通过将卷积神经网络引入经典的“特征图-模式图-柱状图”模式来提取特征。文献<sup>[8]</sup>提出子空间回归方法,将不同类别的人脸划分为不同的子空间,令有遮挡人脸和无遮挡人脸各自回归到相应的子空间,并对无遮挡的子空间进一步处理。CEN等<sup>[9]</sup>提出一种基于深度字典表示的人脸图像分类方法,其使用卷积神经网络对遮挡人脸图像进行深层特征提取,然后使用字典对提取的深层特征进行线性编码。这类方法所提取的有效特征忽略甚至直接丢弃了许多其他特征,容易造成识别准确率过低。另一种思路是先对遮挡部分的人脸进行修复,然后再经过恰当的识别网络对修复后的人脸进行分类,这种方法通过恢复原来缺失的特征信息,使待识别的人脸拥有接近原图的丰富特征,采用的方法包括自动编码器、生成对抗网络(Generative Adversarial Networks, GAN)等。文献<sup>[10]</sup>使用自动编码器作为

生成器,采用全局和局部2个判别器使生成图像的语义更加丰富。文献<sup>[11]</sup>针对图像修复可能出现不够清晰的情况,采用图像解耦后融合的方法来修复人脸。文献<sup>[12]</sup>提出一种Multi-Loss模型,将注意力机制和GAN结合来提升修复效果。文献<sup>[13]</sup>针对网格遮挡设计一种加权的FCN修复模型,但不适用于其他类型的遮挡。文献<sup>[14]</sup>提出一种级联生成对抗网络的修复方法,生成器采用由粗到精的级联模式,判别器采用将局部和全局相融合的双重判别式模型。

现有大多数修复方法均需要遮挡区域信息来保证复原的真实性,但是这对于一定规模的人脸识别来说不切合实际。对修复后人脸进行分类的方法有SVM<sup>[15]</sup>、贝叶斯分类器<sup>[16]</sup>等,但这些分类方法效果有偏差并且很难与特征提取网络相结合,实际使用中有诸多不便。基于深度学习的方法是一种端到端的方法,其将特征提取和特征分类融合在一个模型中,是当前研究热点。

## 2 卷积神经网络和生成对抗网络

### 2.1 卷积神经网络

卷积神经网络是深度学习人脸识别的重要组成部分,其作用是从人脸中提取更丰富、更深层次的特征,主要由卷积层、池化层和全连接层组成。卷积层的核心为卷积核,输入的特征通过卷积核进行一系列卷积操作后就能得到更深层次的特征。池化层连接在卷积层之后,主要的作用是减少通过卷积层得到的特征图大小,从而减少参数计算量,一般常用的池化层有全局平均池化和最大池化2种。全连接层通常使用在网络的末端,把二维的特征图转换成一维特征用于识别分类。设计恰当的卷积网络可以有效提取人脸特征,对于最终识别准确率也有较大影响。自从卷积神经网络AlexNet<sup>[17]</sup>问世,研究人员设计了诸多网络结构以更深层次地提取图像特征,如VGGNet、GoogleNet、ResNet等网络。

### 2.2 生成对抗网络

生成对抗网络<sup>[18]</sup>是当下人工智能研究的热点,经典的GAN模型如图1所示。

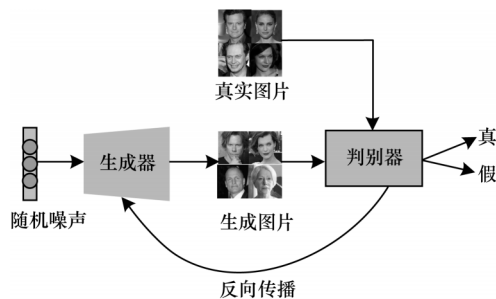


图1 经典的GAN模型

Fig.1 Classical GAN model

如图1所示,经典的GAN模型由生成器和判别器2部分构成,生成器能够学习真实图像分布,判别

器通过将生成器得到的图像与原图像进行对比来判断生成图片的真实性。GAN训练的目标函数为:

$$\min_G \max_D V(D, G) = E_{x \sim p_{\text{data}}(x)} [\log_a(D(x))] + E_{z \sim p(z)} [\log_a(1 - D(G(z)))] \quad (1)$$

其中:  $x$  表示真的样本;  $P(z)$  表示假的样本分布;  $z$  为输入  $G$  的随机噪声分布;  $G(z)$  表示生成器  $G$  生成的图片;  $D(x)$  表示判别器  $D$  判断真实的图片是否真实的概率;  $D(G(z))$  则表示判别器  $D$  判断  $G$  网络生成的图片是否真实的概率。

GAN的优化目标函数为最大最小优化问题,其可以采用博弈思想交叉训练,具体训练过程可以拆解为如下2步:

1) 优化判别器  $D$ , 即将真实图像  $X$  和生成图像  $G(z)$  放入判别器  $D$  中训练, 使  $D(X)$  越大越好、 $D(G(z))$  越小越好, 相应优化目标如式(2)所示:

$$\max_D V(D, G) = E_{x \sim p_{\text{data}}(x)} [\log_a(D(x))] + E_{z \sim p(z)} [\log_a(1 - D(G(z)))] \quad (2)$$

2) 优化生成器  $G$ , 将上述判别器  $D$  得到的结果再反馈给  $G$  网络, 然后  $G$  网络调整自身的参数生成更加真实的图片再去“欺骗”判别器, 使得  $D(G(z))$  越大越好, 相应优化目标如式(3)所示:

$$\min_G V(D, G) = E_{z \sim p(z)} [\log_a(1 - D(G(z)))] \quad (3)$$

经过生成器和判别器交替训练一段时间后, 生成器  $G$  通过大量地学习同类图像数据中的特征以生

成更加真实的图像, 当判别器无法准确判断出图片到底是真实的还是由  $G$  网络生成的时, 即实现了生成目的。

### 3 本文方法

本文提出一种基于循环生成对抗网络的遮挡人脸图像识别算法, 其结构如图2所示, 该模型分为遮挡人脸修复和人脸识别2个模块。不同于只对遮挡部分进行修复的方法, 本文人脸修复方法是对整个图像进行重构且不需要遮挡部位信息, 因此是一种盲修复方式。修复模块的输入是一张带遮挡的人脸, 通过模型直接输出一张还原好原图特征的图片。修复模块采用了2对生成器和判别器循环训练, 生成器1将遮挡人脸尽可能地还原成无遮挡的正常人脸, 判别器2则判别修复后人脸与正常无遮挡人脸的差异, 即把带遮挡的域映射到无遮挡正常人脸的域。考虑到生成器生成过程具有随机性, 导致还原的图片可能带有大面积其他人脸的特征, 因此通过生成器2将已修复图片再生成回原始遮挡图片, 判别器2则判别该图片与原遮挡图片的差异, 通过第2对生成器和判别器的操作使得由生成器1修复好的图片拥有更多的原始人脸特征。识别模块通过Resnet-50来提取人脸特征, 同时针对修复后的人脸可能存在一些特征融合的问题, 引入一种能够将不同类的类间距离量化的损失函数 RegularFace, 以处理类内距离过大和类间距离过小对分类的影响。

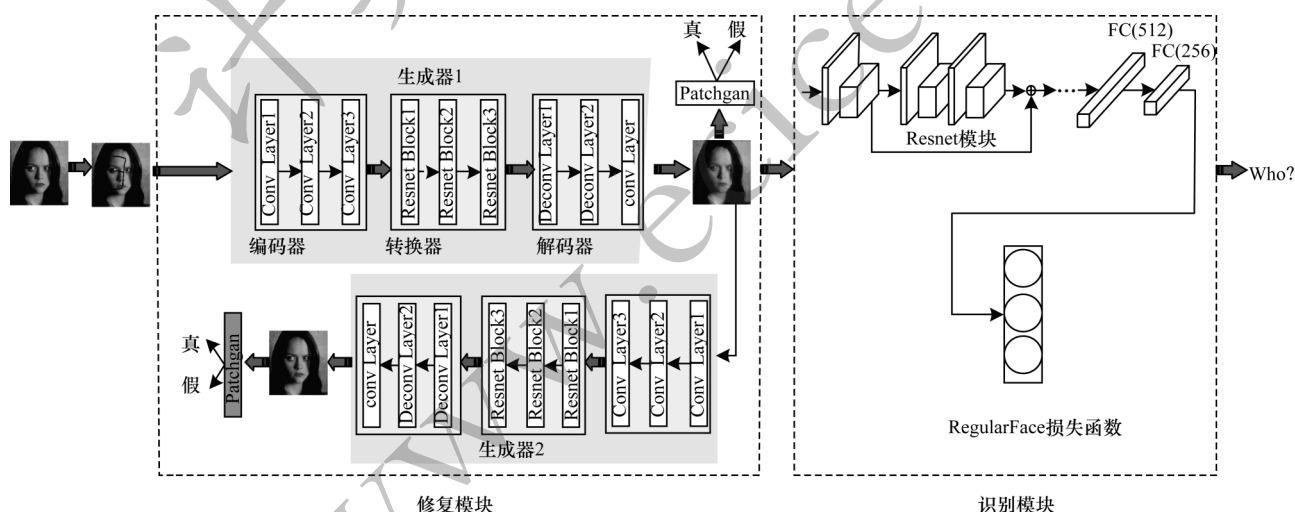


图2 基于循环生成对抗网络的遮挡人脸识别模型

Fig.2 Face recognition with occlusion model based on cyclic generative adversarial networks

#### 3.1 基于循环生成对抗网络的人脸修复

人脸修复模块由2对生成器和判别器的循环生成对抗网络构成, 其中: 生成器由编码器、转换器和解码器构成。编码器由3层卷积层组成, 依次逐渐提取更高级的特征, 特征的数量也即滤波器数量, 本文实验输入图片大小为  $256 \times 256 \times 3$  像素, 通过编码器后, 输入变成  $64 \times 64 \times 256$  像素。转换器由6个残差块组成, 由编码器输出图像的不同通道组合了图像的不同特征,

根据这些特征将图像的特征向量即编码由带遮挡转换到正常无遮挡。每个残差块由2个卷积层组成, 为了确保先前图层的输入属性也可以用于后面的图层, 保证其输出与原始输入不会有太大偏差, 将输入残差添加到输出中。解码器的作用是从特征向量重新构建低级特征, 这里用了3层的转置卷积, 将低级特征转换为正常无遮挡的图像。

普通判别器的基本构成是卷积层、全连接层、密集



连接层和全连接层,然后将判别的结果即输入样本为真样本的概率作为一个实数输出,其输入一般是真实图像和生成图像的整张图像,但这类判别器接收的图像范围过大,可能造成图像生成时分辨率较低、图片不清晰等情况。本文修复模块中的判别器采用 PatchGAN<sup>[19]</sup>,其原理如图3所示。该判别器由5个卷积层构成,输出为一个 $n \times n$ 的矩阵,输出矩阵的均值作为 True 或 False 输出。输出矩阵中的每一个输出均代表原图的一个感受野,对应了原图的一部分,相比一般的 GAN 网络判别器输出为一个数,该判别器考虑到了图像不同部分的影响,且对图像的一小块进行判别输出,这样训练时能够使模型更能关注图像细节,生成的图片将更加清晰。本文将输入图像处理成大小为 $8 \times 8 \times 1$ 像素的不同块,判别器的输入大小为 $256 \times 256 \times 3$ 像素。

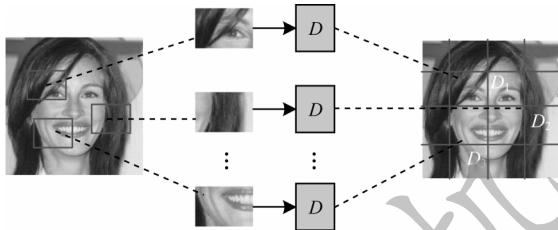


图3 PatchGAN 判别器的原理

Fig.3 Principle of PatchGAN discriminator

修复模块损失函数包括判别器损失和循环一致性损失。判别器损失的目的是使生成目标与原始图片尽可能相似,采用均方误差形式,表达式如式(4)所示:

$$\text{MSE} = \sum_{i=1}^n (y_i - y_i^p)^2 / n \quad (4)$$

其中: $y_i$ 代表实际值; $y_i^p$ 代表估计值。在此基础上可以定义2个判别器的损失,判别器1使第1个生成器生成的无遮挡图片和原无遮挡图片相似,损失函数表达式如式(5)所示:

$$G_1 = \text{MSE}(D_B(f_B), r_B) \quad (5)$$

判别器2使第2个生成器生成的有遮挡图片跟原有遮挡图片相似,损失函数表达式如式(6)所示:

$$G_2 = \text{MSE}(D_A(f_A), r_A) \quad (6)$$

其中: $A$ 表示有遮挡的人脸; $B$ 表示无遮挡的人脸; $f_A$ 表示当输入为原图时生成器2生成的图片, $f_B$ 表示当输入为原图时生成器1生成的图片, $r_A$ 和 $r_B$ 分别表示原有遮挡的图片和无遮挡的图片。

循环一致性损失使得原图经过1个循环训练后得到的图片与原图相似,其主要采用 $L_1$ 损失函数,如式(7)所示:

$$L_1(\hat{y}, y) = \sum_{i=0}^m |y^{(i)} - \hat{y}^{(i)}| \quad (7)$$

其中: $y^{(i)}$ 代表实际值; $\hat{y}^{(i)}$ 代表估计值。

正循环的损失表达式如式(8)所示:

$$c_1 = L_1(s_A, r_A) \quad (8)$$

逆循环的损失表达式如式(9)所示:

$$c_2 = L_1(s_B, r_B) \quad (9)$$

其中: $s_A$ 表示当输入为生成图片时生成器2生成的图

片; $s_B$ 表示当输入为生成图片时生成器2生成的图片。

在上述损失基础上,可定义修复模块总的损失函数表达式如式(10)所示:

$$L_{\text{总}} = G_1 + G_2 + c_1 + c_2 \quad (10)$$

### 3.2 人脸识别模块

由于人脸特征非常丰富,浅层的卷积神经网络结构无法很好地提取到深层的特征,但一味地通过堆叠卷积层和池化层来加深网络深度会导致梯度消失和网络退化,影响网络优化性能。本文所提识别网络采用残差网络 ResNet-50 结构,该网络通过跳跃连接消除深层卷积神经网络中的网络退化问题,降低了网络训练的难度,且不会增加额外的参数和计算量。

损失函数用于度量本文网络输出预测值和实际值的差距,它是一个非负函数。一般来说,损失函数越小,模型的鲁棒性越好。损失函数的宗旨是增大不同类别的类间距离和缩小同一类别的类内距离,常见的损失函数如 CenterLoss<sup>[20]</sup>、Softmax<sup>[21]</sup> 及其变体<sup>[22-24]</sup>,这些损失函数能够较好地缩小同类别特征的内间距离,但由于人脸很多特征具有相似性,不同类别间的距离也可能较小,因此上述损失函数在人脸识别中可能会导致识别准确率较低。本文识别模块采用 RegularFace<sup>[4]</sup> 损失函数,其使用不同类别中心点的角距离来量化不同类别的类间距离,类间距离的表达式如式(11)所示:

$$p_i = \max_{j \neq i} \cos(\phi_{i,j}) = \max_{j \neq i} \frac{\omega_i \times \omega_j}{\|\omega_i\| \times \|\omega_j\|} \quad (11)$$

其损失函数为:

$$L_r(\omega) = \frac{1}{c} \max_{j \neq i} \frac{\omega_i \times \omega_j}{\|\omega_i\| \times \|\omega_j\|} \quad (12)$$

其中: $\omega$ 代表类别的权重向量; $\omega_i$ 表示 $\omega$ 的第 $i$ 列, $j$ 表示距离 $i$ 最近的聚类中心; $\phi_{(i,j)}$ 表示 $\omega_i$ 和 $\omega_j$ 之间的夹角,识别的目的是希望 $\phi_{(i,j)}$ 值越大越好,即 $p_i$ 的值越小越好。本文识别模块中,将该损失函数与 Softmax 函数搭配使用,这样既能有效地缩小同一类的类内距离,又能增大不同类的类间距离。Softmax 函数的表达式如式(13)所示:

$$L_1 = -\frac{1}{N} \sum_{i=1}^N \log_a \frac{e^{\omega_{y_i}^T x_i + b_{y_i}}}{\sum_{j=1}^n e^{\omega_j^T x_i + b_j}} \quad (13)$$

因此,识别模块损失函数如式(14)所示:

$$L_{\text{总}} = L_1 + \lambda L_r(\omega) \quad (14)$$

其中: $\lambda$ 用于平衡2个损失函数,本文取值为1。

## 4 仿真实验

为验证本文算法的有效性,本文对盲修复结果及识别准确率进行了多轮实验验证。实验环境为 Windows 10, 采用编程环境为 pytorch 和 python3.6.8, 硬件为显存 11 GB 的 RTX2080Ti GPU、3.6 GHz 的 i7-9700K CPU、32 G RAM 的 PC 机。

### 4.1 数据集选择

本文选用 CASIA-WebFace<sup>[25]</sup> 人脸数据集中的 20 个不同身份的人,共计约 12 000 张图片作为数据集。由

于本文修复方法主要针对物体遮挡,因此在选择图像时主要挑选光线充足和正脸的图像以避免光线遮挡和自遮挡对实验的影响。实验将原图片的 $250\times 250$ 像素变换为 $256\times 256$ 像素。由于没有标准和通用的遮挡人脸数据集,并且现实中的遮挡类型具有多样性,因此在训练数据时采用图片点乘的方式给原图加上了不同类型和大小的遮挡作为训练集,如图4所示。



图4 不同遮挡类型的图片示例  
Fig.4 Examples of different occlusion types

4.2 模型训练

本文方法包括修复和识别2部分,两者独立训练,具体训练方式如下。

4.2.1 修复网络训练

在训练集中选取1 000张涵盖20个不同身份的图片并随机加上图4所示的遮挡(称为数据集1)作为生成器1的输入,再从上述训练集中随机挑选另外1 000张不同身份的图片(称为数据集2)做为第1个判别器的输入。因此,判别器1的输入包括生成器1的输出和数据集2,然后生成器1的输出再做为生成器2的输入;将生成器2的输出和数据集1再输入判别器2,这样循环训练直到生成器1的输出收敛为无遮挡的人脸图片。普通GAN的训练方式是判别器先于生成器开始训练,并且训练次数多于生成器。本文训练方式也不例外,略有差别的是判别器2要在生成器1生成几轮后才开始训练,以便得到较好的图片投入逆向训练中,一般200个epoch后网络开始收敛,生成器1得到较好的修复结果。

由于修复网络训练过程中并没有用到遮挡区域信息,因此达到了盲修复的目的,尽管此类修复后的图像整体效果较原图的清晰化程度低,但总体特征相较于遮挡图像而言具有较大程度的改善,因此识别准确率也会有质的提升,后续实验结果也验证了这一点。修复网络超参数设置如表1所示。

表1 修复网络训练的超参数设置  
Table 1 Hyper-parameter setting for the training of inpainting network

参数名	参数设置
优化算法	Adam
G学习率	0.000 2
D学习率	0.000 2
批次数量	250
迭代数量	300

4.2.2 识别网络训练

将所有12 000余张图片进行划分,其中4/5用做训练集、1/5用做测试集。为更好地对比识别网络性能,首先训练识别模型使其在正常人脸数据集上达到较高的识别准确率,并将此时训练好的的网络作为基准来识别修复后的人脸,这个基准可以随着人脸数据集的大小进行调整,最终修复后的识别准确率也会随着基准变化,但本文最终的目的是探究修复效果和提升识别准确率,故基准的大小不影响对比实验。网络超参数设置如表2所示。

表2 识别网络训练的超参数设置  
Table 2 Hyper-parameter setting for the training of identification network

参数名	参数设置
优化算法	Adam
学习率	0.000 5
权重衰减	$1\times 10^{-8}$
批次数量	128
迭代数量	100
平衡参数 $\lambda$	1

4.3 实验结果及分析

本文对不同遮挡类型、不同遮挡面积和不同遮挡区域的修复效果和识别准确率进行实验,并对比了DCGAN+CNN算法下的修复效果及识别准确率,分析引入RegularFace损失函数对识别效果的影响。

4.3.1 修复结果分析

修复效果分析具体如下:

1) 不同遮挡类型的修复结果

针对不同遮挡类型的部分人脸图像修复效果如图5所示。由图5可知不同类型的遮挡包括线性遮挡和矩形遮挡。对于线性遮挡,遮挡部分被还原得比较接近客观事实,且未遮挡的部分无明显变化,整个人脸看上去较为真实连贯。对于矩形遮挡,小块遮挡能很好地还原,但当矩形遮挡面积过大时,修复效果不甚理想。

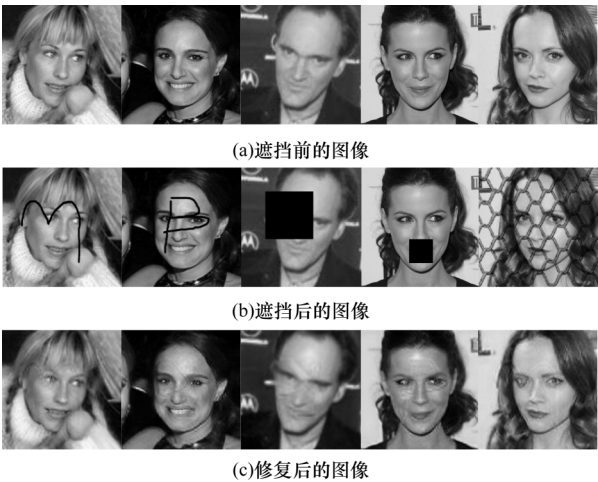


图5 不同遮挡类型的修复效果对比

Fig.5 Comparison of repair effects of different occlusion types



## 2) 不同遮挡面积的修复效果

对不同遮挡面积的分析建立在同一人的相同遮挡形状上,本文选用了不同面积的线性遮挡和不同面积的矩形遮挡进行实验,结果如图6所示。从修复效果可以看出,修复效果随着遮挡面积的增大而逐渐变差。

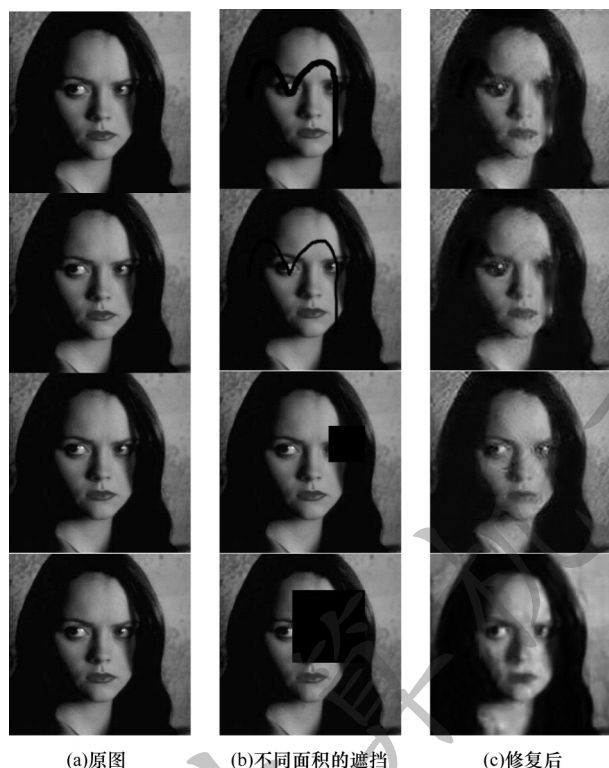


图6 不同遮挡面积的修复效果对比

Fig.6 Comparison of repair effect of different occlusion area

## 3) 不同遮挡部位的修复效果

五官特征对于人脸识别的准确率影响很大<sup>[26]</sup>,因此本文对不同遮挡区域的人脸修复效果进行了对比实验,实验结果如图7所示。

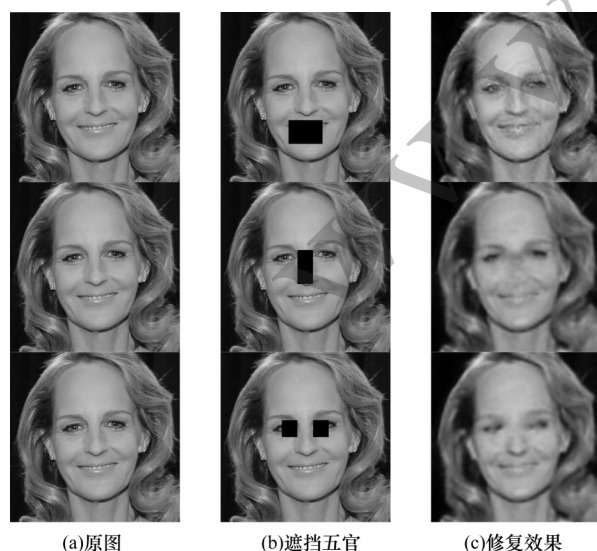


图7 不同遮挡部位的修复效果对比

Fig.7 Comparison of repair effect of different occlusion site

由图7可知,当嘴巴和鼻子受到遮挡时,修复效果较好,但当眼睛收到遮挡时,其修复后的细节丢失比较严重。

## 4) 与其他算法修复效果对比

本文对比了DCGAN+CNN算法的修复效果,实验结果如图8所示。

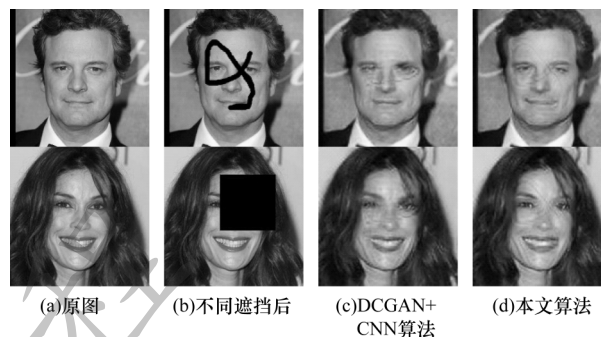


图8 本文算法与DCGAN+CNN算法的修复效果对比

Fig.8 Comparison of repair effect between the algorithm in this paper and the DCGAN+CNN algorithm

如图8所示,DCGAN+CNN算法脸部修复细节不如本文算法连贯,脸部甚至较原图较小,而本文算法保留了大量未遮挡区域的特征,看起来更加自然,这也从侧面证明了本文算法的有效性。

## 4.3.2 修复前后识别准确率的比较

本文的最终目的是提高带遮挡人脸的识别准确率,而不是过度追求复原成清晰无缺的图片。人脸识别准确率与网络提取到的人脸特征密切相关,本文修复方法保留和补充了大量的原图特征,且修复得到的图片也符合语义,因此相较于直接识别带遮挡人脸图片,准确率会有很大提升。

本文所用的识别模型在正常人脸数据集上达到了一个较高的识别准确率,以此为参照分别比较了不同遮挡类型的修复准确率,结果如图9所示。

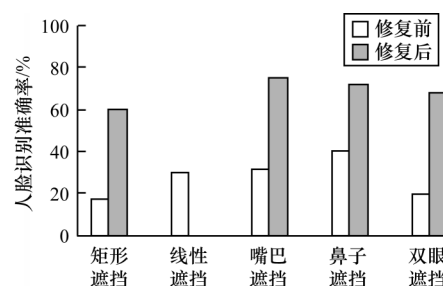


图9 相同面积下的不同遮挡类型对修复前后识别准确率的影响

Fig.9 Influence of different occlusion types in the same area on recognition accuracy before and after repair

由图9可知,在遮挡面积相同时,矩形遮挡修复前后的识别准确率都很低,但修复后的识别准确率较修复前仍提升了一倍多,表明修复后的人脸获得了更多有效特征。线性遮挡在修复后识别准确率提升得最多,修复后的效果也最好。而五官遮挡的人脸尽管修复后的效果不佳,但由于其特征复原得比较好,所以识别准确率也有很大提升,仅次于线性遮挡修复后的识别效果。由图10可知,当线性遮挡面

积小于30%时,由于遮挡面积较小,修复前的识别准确率已经达到较高状态,故修复后的准确率提升不大,但仍有约10%的提升。当遮挡面积在30%附近时,修复后的识别准确率有很大的提升,当遮挡面积进一步增大到40%后,修复后的识别准确率提升了约一倍。

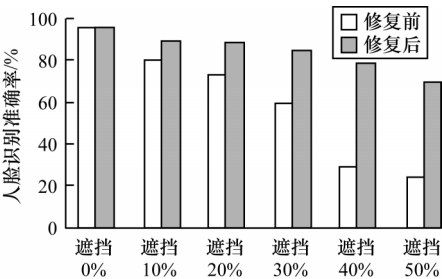


图10 不同线性遮挡面积对修复前后识别率的影响  
Fig.10 Influence of different linear occlusion areas on recognition rate before and after repair

识别正确和错误的示例分别如图11、图12所示。由图12(a)可以看出,当五官特征恢复较差时,容易导致识别失败。对比图11(b)、图12(b)可知,识别网络本身对识别效果也有重要影响。

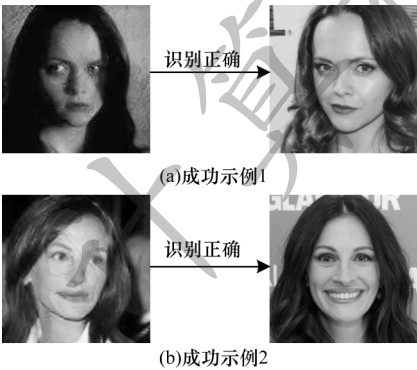


图11 识别成功的示例  
Fig.11 Identify successful examples

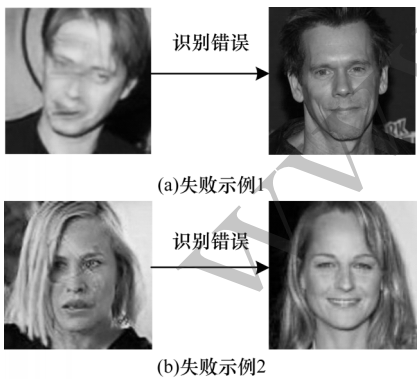


图12 识别失败的示例  
Fig.12 Identify examples of failures

4.3.3 识别网络性能分析

为验证 RegularFace 损失函数的性能,本文对使用 RegularFace+Softmax(简记为 R+S)损失函数和仅单使用 Softmax(简记为 S)损失函数情况下修复前后

的识别准确率进行了比较,实验结果如图13和图14所示。

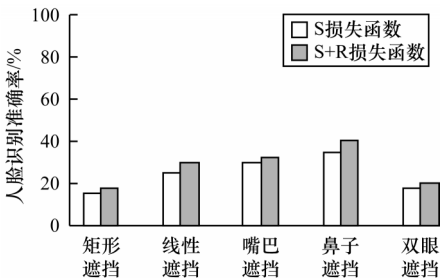


图13 不同损失函数修复前识别准确率对比  
Fig.13 Comparison of recognition accuracy before repair of different loss functions

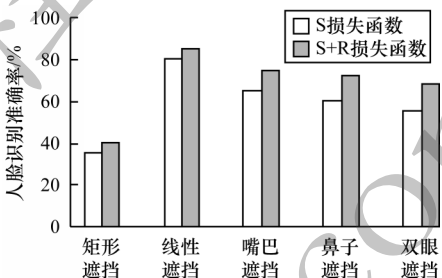


图14 不同损失函数修复后识别准确率对比  
Fig.14 Comparison of recognition accuracy after repair of different loss functions

由图13可知,修复前的人脸丢失了过多特征,无论用哪种损失函数的识别准确率均很低,但用 RegularFace+Softmax 损失函数较于 Softmax 仍有部分提升。由图14可知,修复后的识别准确率均有明显提升,使用 RegularFace+Softmax 损失函数的提升效果也更明显,尤其是五官遮挡,这是由于 RegularFace 有对类间距离进行量化的特点,对于五官这些难以分类的特征效果提升更加明显。

4.3.4 与其他算法的对比

本文选取了 DCGAN+CNN 算法进行对比实验,其中 DCGAN 作为修复算法来修复人脸图像,CNN 则选用与本文相同的 ResNet50,损失函数统一使用 softmax,识别结果如表3、表4所示。

由表3、表4可知,无论是在不同的遮挡类型还是不同遮挡面积条件下,本文算法的人脸识别准确率都更高,DCGAN+CNN 算法相较于本文算法留存的原图特征较少,当遮挡面积较小时,本文算法优势明显,当遮挡面积达到50%左右时,DCGAN+CNN 算法的识别准确率较低,而本文算法的识别准确率还能维持在一个较高的水平。

表3 不同算法对不同遮挡类型的识别准确率对比

Table 3 Comparison of recognition accuracy of different occlusion types by different algorithms %			
算法	无遮挡	矩形遮挡	线性遮挡
DCGAN+CNN算法	96.0	38.5	74.4
本文算法	96.0	60.0	85.0

表4 不同算法对不同遮挡面积的识别准确率对比

Table 4 Comparison of recognition accuracy of different algorithms for different occlusion areas %					
算法	线性遮挡 10%	线性遮挡 20%	线性遮挡 30%	线性遮挡 40%	线性遮挡 50%
DCGAN+CNN算法	86.3	85.0	74.4	70.8	68.2
本文算法	90.2	89.5	85.5	85.2	83.1

5 结束语

针对当前含遮挡人脸识别方法准确率较低的问题,本文提出一种基于循环生成对抗网络的含遮挡人脸识别算法。采用先修复再识别的方式,在修复阶段不需要遮挡信息,仅通过2个生成器和2个判别器循环训练,使生成的人脸与原遮挡人脸的特征重合度高且修复后语义连贯性较好,实现人脸遮挡图像的盲修复。同时,在识别阶段引入RegularFace损失函数,有效增加不同类的类间距离,减小同类间的类间距离。实验结果表明,与DCGAN+CNN算法相比,所提算法对不同遮挡类型和遮挡区域的人脸图像识别准确率均有所提高。但本文算法对大面积矩形遮挡图像,以及对随机遮挡和日常遮挡的修复效果也有待提升。下一步将引入遮挡区域估计方式以充分利用原始未遮挡图像信息,提升遮挡人脸的识别准确率。

参考文献

[ 1 ] MASII WU Y, HASSNER T, et al. Deep face recognition; a survey [C]//Proceedings of the 31st Conference on Graphics, Patterns and Images. Washington D. C. , USA; IEEE Press, 2018:471-478.

[ 2 ] CHENG Y, JIANG B, JIA K B. A deep structure for facial expression recognition under partial occlusion [C]//Proceedings of the 10th International Conference on Intelligent Information Hiding and Multimedia Signal Processing. Washington D. C. , USA; IEEE Press, 2014:211-214.

[ 3 ] HE K M, ZHANG X Y, REN S Q, et al. Deep residual learning for image recognition [C]//Proceedings of IEEE Conference on Computer Vision and Pattern Washington D. C. , USA; IEEE Press, 2016:770-778.

[ 4 ] ZHAO K, XU J Y, CHENG M M. RegularFace; deep face recognition via exclusive regularization [C]//Proceedings of IEEE/CVF Conference on Computer Vision and Pattern Recognition. Washington D. C. , USA; IEEE Press, 2019:1136-1144.

[ 5 ] CHEN Y, LIU S. Deep partial occlusion facial expression recognition via improved CNN [C]//Proceedings of the Advances in Visual Computing. Berlin, Germany: Springer, 2020:451-462.

[ 6 ] ZHANG T P, TANG Y Y, FANG B, et al. Face recognition under varying illumination using gradientfaces [J]. IEEE Transactions on Image Processing, 2009, 18(11):2599-2606.

[ 7 ] ZHANG W C, SHAN S G, WEN G, et al. Local gabor binary pattern histogram sequence; a novel non-statistical model for face representation and recognition [C]//Proceedings of the 10th IEEE International Conference on

Computer Vision. Washington D. C. , USA; IEEE Press, 2005:786-791.

[ 8 ] 李小薪,梁荣华. 有遮挡人脸识别综述:从子空间回归到深度学习[J]. 计算机学报,2018,41(1):177-207.

LI X X, LIANG R H. A review for face recognition with occlusion; from subspace regression to deep learning [J]. Chinese Journal of Computers, 2018, 41(1):177-207. (in Chinese)

[ 9 ] CEN F, WANG G H. Dictionary representation of deep features for occlusion-robust face recognition [J]. IEEE Access, 2019, 7:26595-26605.

[ 10 ] LI Y, LIU S, YANG J, et al. Generative face completion [EB/OL]. [2021-02-10]. [https://www.researchgate.net/publication/316272513\\_Generative\\_Face\\_Completion](https://www.researchgate.net/publication/316272513_Generative_Face_Completion).

[ 11 ] 张雪菲,程乐超,白升利,等. 基于变分自编码器的人脸图像修复[J]. 计算机辅助设计与图形学学报,2020,32(3):401-409.

ZHANG X F, CHENG L C, BAI S L, et al. Face image inpainting via variational autoencoder [J]. Journal of Computer-Aided Design & Computer Graphics, 2020, 32(3):401-409. (in Chinese)

[ 12 ] 曹真,杨云,齐勇,等. 基于多损失约束与注意力块的图像修复方法[J]. 陕西科技大学学报,2020,38(3):158-165.

CAO Z, YANG Y, QI Y, et al. Image inpainting method based on multi-loss constraint and attention block [J]. Journal of Shaanxi University of Science & Technology, 2020, 38(3):158-165. (in Chinese)

[ 13 ] ZHANG S, HE R, SUN Z N, et al. DeMeshNet; blind face inpainting for deep MeshFace verification [J]. IEEE Transactions on Information Forensics and Security, 2018, 13(3):637-647.

[ 14 ] 陈俊周,王娟,龚勋. 基于级联生成对抗网络的人脸图像修复[J]. 电子科技大学学报,2019,48(6):910-917.

CHEN J Z, WANG J, GONG X. Face image inpainting using cascaded generative adversarial networks [J]. Journal of University of Electronic Science and Technology of China, 2019, 48(6):910-917. (in Chinese)

[ 15 ] 曾阳艳,叶柏龙. 基于支持向量机的人脸识别系统的研究[J]. 计算机工程与应用,2008,44(15):182-184.

ZENG Y Y, YE B L. Face recognition based on support vector machines [J]. Computer Engineering and Applications, 2008, 44(15):182-184. (in Chinese)

[ 16 ] 冯小荣,惠康华,柳振东. 基于卷积特征和贝叶斯分类器的人脸识别[J]. 智能系统学报,2018,13(5):769-775.

FENG X R, HUI K H, LIU Z D. Face recognition based on convolution feature and Bayes classifier [J]. CAAI Transactions on Intelligent Systems, 2018, 13(5):769-775. (in Chinese)

[ 17 ] KRIZHEVSKY A, SUTSKEVER I, HINTON G E. ImageNet classification with deep convolutional neural networks [J]. Communications of the ACM, 2017, 60(6):84-90.



(上接第296页)

- [18] GOODFELLOW I J, POUGET-ABADIE J, MIRZA M, et al. Generative adversarial nets [C]//Proceedings of the 27th International Conference on Neural Information Processing Systems. Washington D. C. , USA ; IEEE Press, 2014: 2672-2680.
- [19] LATA K, DAVE M, NISHANTH K N. Image-to-image translation using generative adversarial network [C]//Proceedings of the 3rd International Conference on Electronics, Communication and Aerospace Technology. Washington D. C. , USA ; IEEE Press, 2019: 186-189.
- [20] WEN Y D, ZHANG K P, LI Z F, et al. A discriminative feature learning approach for deep face recognition [C]//Proceedings of Conference on Computer Vision. Berlin, Germany: Springer, 2016: 499-515.
- [21] MIKOLOV T, KOMBRINK S, BURGET L, et al. Extensions of recurrent neural network language model [C]//Proceedings of IEEE International Conference on Acoustics, Speech and Signal Processing. Washington D. C. , USA ; IEEE Press, 2011: 5528-5531.
- [22] LIU W Y, WEN Y D, YU Z D, et al. SphereFace: deep hypersphere embedding for face recognition [C]//Proceedings of IEEE Conference on Computer Vision and Pattern Recognition. Washington D. C. , USA ; IEEE Press, 2017: 6738-6746.
- [23] WANG H, WANG Y T, ZHOU Z, et al. CosFace: large margin cosine loss for deep face recognition [C]//Proceedings of IEEE/CVF Conference on Computer Vision and Pattern Recognition. Washington D. C. , USA ; IEEE Press, 2018: 5265-5274.
- [24] DENG J K, GUO J, XUE N N, et al. ArcFace: additive angular margin loss for deep face recognition [C]//Proceedings of IEEE/CVF Conference on Computer Vision and Pattern Recognition. Washington D. C. , USA ; IEEE Press, 2019: 4685-4694.
- [25] YI D, LEI Z, LIAO S C, et al. Learning face representation from scratch [EB/OL]. [2021-02-10]. <https://arxiv.org/abs/1411.7923v1>.
- [26] ELMAHMUDI A, UGAIL H. Deep face recognition using imperfect facial data [J]. Future Generation Computer Systems, 2019, 99: 213-225.

编辑 赖玉玲