



基于面部特征图对称的人脸正面化生成对抗网络算法

李虹霞^{1,2}, 秦品乐^{1,2*}, 闫寒梅³, 曾建潮^{1,2}, 鲍骞月^{1,2}, 柴锐^{1,2}

(1. 中北大学 大数据学院, 太原 030051; 2. 山西省医学影像与数据分析工程研究中心(中北大学), 太原 030051;

3. 山西警察学院 刑事科学技术系, 太原 030401)

(* 通信作者电子邮箱 qpl@nuc.edu.cn)

摘要: 目前人脸正面化研究主要解决人脸偏转问题, 而对监控视频等现实场景中同时受偏转和俯仰变化影响的侧脸的正面化生成关注较少, 针对这个问题和多角度侧脸生成的正面人脸图存在身份信息保留不全的问题, 提出了一种基于特征图对称模块和眼周特征保留损失的生成对抗网络(GAN)。首先, 根据人脸对称性先验, 提出特征图对称模块, 先使用人脸关键点检测器检测出侧脸鼻尖点位置, 再将编码器提取到的特征图依照鼻尖位置进行镜像对称, 从而在特征层面上缓解面部信息缺失的问题。其次, 借鉴眼周识别思想, 在现有的生成图身份保留方法中加入了眼周特征保留损失以训练生成器生成逼真的且保留身份信息的人脸正面图像。实验结果表明, 所提算法得到的生成图面部细节保留较好, 且在 CAS-PEAL-R1 数据集的所有俯角下人脸的平均 Rank-1 识别率为 99.03%, 可见该算法能够有效解决多角度侧脸的正面化问题。

关键词: 人脸识别; 人脸正面化; 生成对抗网络; 深度学习; 眼周识别

中图分类号: TP389.1 **文献标志码:** A

Face frontalization generative adversarial network algorithm based on face feature map symmetry

LI Hongxia^{1,2}, QIN Pinle^{1,2*}, YAN Hanmei³, ZENG Jianchao^{1,2}, BAO Qianyue^{1,2}, CHAI Rui^{1,2}

(1. School of Data Science and Technology, North University of China, Taiyuan Shanxi 030051, China;

2. Shanxi Medical Imaging and Data Analysis Engineering Research Center (North University of China), Taiyuan Shanxi 030051, China;

3. Science and Technology Department, Shanxi Police College, Taiyuan Shanxi 030401, China)

Abstract: At present, the research of face frontalization mainly solves the face yaw problem, and pays less attention to the face frontalization of the side face affected by yaw and pitch at the same time in real scenes such as surveillance video. Aiming at this problem and the problem of incomplete identity information retained in front face image generated by multi-angle side faces, a Generative Adversarial Network (GAN) based on feature map symmetry and periocular feature preserving loss was proposed. Firstly, according to the prior of face symmetry, a symmetry module of the feature map was proposed. The face key point detector was used to detect the position of nasal tip point, and mirror symmetry was performed to the feature map extracted by the encoder according to the nasal tip, so as to alleviate the lack of facial information at the feature level. Finally, benefiting from the idea of periocular recognition, the periocular feature preserving loss was added in the existing identity preserving method of generated image to train the generator to generate realistic and identity-preserving front face image. Experimental results show that the facial details of the images generated by the proposed algorithm were well preserved, and the average Rank-1 recognition rate of faces with all angles under the pitch of CAS-PEAL-R1 dataset is 99.03%, which can effectively solve the frontalization problem of multi-angle side faces.

Key words: face recognition; face frontalization; Generative Adversarial Network (GAN); deep learning; periocular recognition

0 引言

随着深度学习的发展,人脸正面识别的性能有了很大提升,而监控视频中获得的人脸图像往往同时受到偏转角和俯仰角的影响,对于这种包含多角度变化的侧脸图像,现有人脸识别系统的识别准确率较低。目前针对多角度侧脸识别主要

有两种方法:一种是学习不受姿态影响的人脸身份特征^[1-2];另一种是利用单一侧脸图像恢复出保留身份的逼真正脸图像,也称为人脸正面化^[3-8]。就第一种方法而言,由于现有人脸数据集呈现长尾分布,在姿态变化上分布不均匀,因此很难学习到较为鲁棒的姿态不变性人脸特征^[3]。第二种方法人脸正面化可以作为一种预处理操作部署在人脸识别系统中,使

收稿日期:2020-06-09;修回日期:2020-09-16;录用日期:2020-10-09。 基金项目:山西省重点研发项目(201803D31212-1)。

作者简介:李虹霞(1996—),女,山西大同人,硕士研究生,主要研究方向:深度学习、计算机视觉; 秦品乐(1978—),男,山西长治人,教授,博士,CCF会员,主要研究方向:机器视觉、大数据、医学影像分析; 闫寒梅(1968—),女,山西天镇人,副教授,主要研究方向:刑事科学技术、刑事影像学; 曾建潮(1963—),男,陕西大荔人,教授,博士生导师,博士,CCF会员,主要研究方向:复杂系统的维护决策、健康管理; 鲍骞月(1998—),男,山西朔州人,主要研究方向:深度学习、机器视觉; 柴锐(1985—),男,山西运城人,讲师,博士,主要研究方向:医学影像处理。



得现有性能强的识别系统不需要重新训练也可以在一定程度上提升识别准确率,同时在某些场景中,正面化的结果可以作为关键的参考材料,例如在刑侦领域,可以为办案人员提供参考。此外,人脸正面化在其他领域也具有实用的意义。

目前针对监控摄像头得到的侧脸进行正面化主要有两方面的困难:被遮挡信息的补全和保留身份信息问题。

对于解决多角度侧脸被遮挡信息补全的问题,现有方法可以分为两种:基于二维方法^[4-8]和基于三维方法^[9-11]。

基于三维的方法需要二维图像到三维模型的映射,这一过程会引入较大误差,且三维方法得到的正面化图像通常不够真实,会出现人工伪影和严重的面部纹理信息丢失情况,并且在大角度姿态下,三维方法的性能会严重下降^[4]。

基于二维的方法在大角度姿态下相比三维方法性能较好,是目前人脸正面化算法中使用最多的。Kan等^[5]利用自编码器提出堆叠渐进式自编码器(Stacked Progressive Auto-Encoder, SPAE),以渐进方式把大角度侧脸逐渐转换为小角度最终实现正面化。Yim等^[6]提出包括主深度神经网络(Main Deep Neural Network, Main DNN)和辅深度神经网络(Auxiliary Deep Neural Network, Auxiliary DNN)的多任务学习模型,Main DNN负责生成目标姿态下的人脸图像,Auxiliary DNN利用生成图作为输入来重建原始输入图像,以这种多任务方式保留更多的身份信息。但这些方法在训练时由于引入的约束较少,得到的生成图像较为模糊。近几年,生成对抗网络(Generative Adversarial Network, GAN)^[12]提升了二维图像生成任务的视觉效果,很多基于GAN的人脸正面化方法被提出。两路生成对抗网络(Two-Pathway Generative Adversarial Network, TP-GAN)^[7]提出了双支路生成对抗网络,通过全局和局部两条路径分别处理侧脸图像,再将两部分结果进行融合得到最终的正面人脸,同时根据人脸对称性先验提出对称损失来补全被遮挡部分,但这是从图像层面上补全信息。人脸正面化模型(Face Normalization Model, FNM)^[4]提出了一系列人脸注意力机制判别器来判别生成图的真实性,从而增加对生成器的惩罚,得到逼真的正面人脸图像。用于面部姿态分析的多偏转多俯仰高质量数据库(Multi-yaw Multi-pitch high-quality database for Facial Pose Analysis, M2FPA)^[8]采用软注意力的方式利用人脸解析网络分割出发、五官和皮肤,分别给每一部分增加判别器判别其真假,以补全丢失的面部信息从而实现正面化。上述这些基于GAN的方法大多通过改进判别器结构,从而增大对生成器的惩罚,促使生成器生成被遮挡部分,但即使判别器能力强,多角度姿态的自遮挡问题使得输入图像本身有效的面部特征较少,造成生成器可利用的信息较少,使得生成过程困难。因此,为了增加可利用信息,缓解自遮挡问题,本文根据人脸对称性这一先验知识,设计了特征图对称模块,即根据鼻尖位置判断输入侧脸的朝向,然后将编码器卷积层提取的特征图依据人脸朝向进行镜像对称,从特征层面上补全被遮挡的关键人脸信息,帮助生成器生成缺失的面部信息。

对于保留身份信息问题,现有正面化方法^[3-4,7-8]利用性能强的人脸识别网络提取生成图和真实图的深层身份特征,通过降低特征之间的感知距离使得生成图的身份信息保留,这种方法通过全局方式保留生成图的面部整体特征,但在局部特征保留上关注较少,文献[13-14]指出眼周区域即包含眼睛和眉毛的矩形区域是面部中具有高判别性的区域,因此本文在文献[3-4,7-8]提出的身份保留损失基础上,加入了眼周特征保留损失,分别从全局和局部的角度来提高生成图的身份信息保留。

本文的主要工作如下:

1)本文提出了一个端到端的特征图对称人脸正面化生成对抗网络(Symmetric Generative Adversarial Network, Sym-GAN),可以根据单张大角度下的侧脸图像生成保留其身份的正面化图像。生成器中加入了特征图对称模块,可以对多角度变化的侧脸图像进行一定程度的信息补全。

2)为了使得生成图保留更多关键的身份信息,同时利用人脸全局特征和眼周局部特征,通过缩小生成图像和真实图像对在面部区域和眼周区域的特征向量距离,使得生成正脸图保留更多的身份信息。

3)即使Sym-GAN未包含任何人脸识别网络,但它可以轻松地部署在人脸识别系统中,作为一种预处理操作来专门处理多角度变化的侧脸图像,以提高现有人脸识别系统的性能。

在公开数据集CAS-PEAL-R1^[15]上进行了定性和定量实验,实验结果证明本文提出的Sym-GAN可以有效提高人脸识别模型在多角度侧脸下的识别性能。

1 人脸正面化

目前基于生成对抗网络的人脸正面化方法主要包括TP-GAN^[7]、M2FPA^[8]、双注意力生成对抗网络(Dual-Attention Generative Adversarial Network, DA-GAN)^[16]。如图1所示,TP-GAN提出包括全局和局部路径的双支路生成器,全局路径处理侧脸图像得到正面脸型轮廓,局部路径处理各五官图像块输出对应正面图像,最后将全局和局部路径的结果进行融合得到正面图像。TP-GAN的生成器采用编-解码结构,同时加入跳层拼接利用编码器提取的侧脸特征进行特征融合,但侧脸特征包含的信息较少,使得解码器上采样过程中可用的有效特征较少。对于TP-GAN的双路径生成器,网络结构庞大,生成图像受最慢路径影响,在训练和预测时效率低,针对这一问题,M2FPA和DA-GAN只采用全局路径的单支路网络也可以达到相同效果。对于生成图身份信息保留,TP-GAN、M2FPA和DA-GAN使用的身份保留损失是以面部全局角度提高生成图身份信息保留,对局部具有高辨别能力的五官信息关注较少。

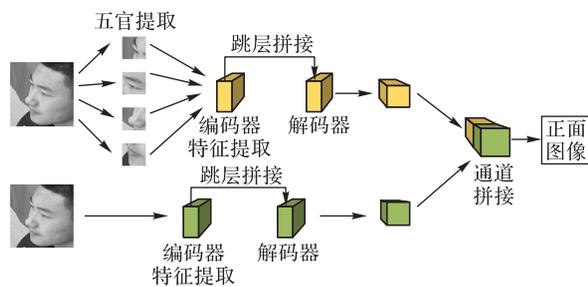


图1 TP-GAN网络框架

Fig. 1 Network framework of TP-GAN

2 Sym-GAN

本文提出的Sym-GAN用于对受多角度变化的侧脸实现正面化生成,网络框架如图2所示。Sym-GAN由生成器-判别器组成。生成器和TP-GAN全局路径生成器一样采用编解码结构。编码器对输入侧脸图像提取特征,同时借助特征图对称模块,将侧脸特征依据鼻尖位置镜像对称;解码器将编码器提取的深层特征进行多次反卷积,在上采样过程中加入跳层拼接,将原始侧脸特征图、对称特征图、输入图像在通道维度上进行拼接,最终输出生成正脸图像。判别器采用深层卷积网络判定生成图和基准图的真实性,促使生成器生成逼真的



正面图像。Sym-GAN 主要包含两个模块:特征图对称模块和改进的解码器跳层拼接模块。根据人脸对称性这一先验提出了特征图对称模块,应用于解码器跳层拼接,以补全丢失的人脸关键信息,从而缓解自遮挡造成的生成器对丢失部分生成困难的问题。改进的解码器跳层拼接是指对 TP-GAN 解码器跳层拼接进行改进,在反卷积过程中拼接了编码器提取的侧脸特征图、对称特征图以及输入图,可以起到补全信息和增强

面部细节学习的作用。此外,本文在身份特征保留损失的基础上,加入了眼周特征保留损失来训练生成器。身份特征从全局角度表征一些面部纹理特征和五官间的拓扑结构,眼周特征从局部角度表征眉毛、眼型和眼间距等局部身份信息,从而提高生成图全局和局部区域身份信息保留能力。下面将详细介绍特征对称模块和改进的解码器跳层拼接模块,眼周特征保留损失在 3.2 节眼周特征保留损失介绍。

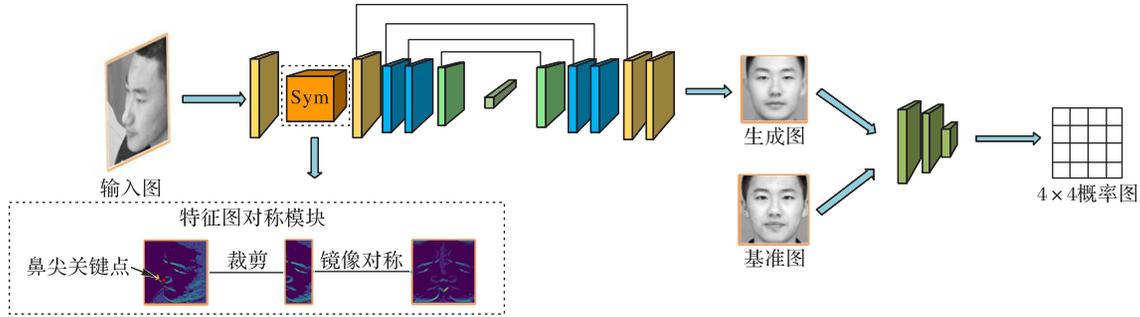


图 2 Sym-GAN 网络框架

Fig. 2 Network framework of Sym-GAN

2.1 特征图对称编码器

编码器的作用是将输入图像通过多层卷积提取特征,得到高维人脸特征表达向量,解码器对特征向量进行上采样恢复出正面图。对于多角度变化下的侧脸正面化任务,为了保证生成正面人脸逼真且保留足够多的身份信息,编码器一方面需要对输入图像进行信息过滤,滤去背景噪声和与人脸身份无关的角度信息,为解码器提供更为紧凑的面部特征向量;另一方面需要提供给解码器较为丰富有效的面部信息,以缓解自遮挡导致的人脸信息丢失问题。为了实现这一目的,本文在编码器对侧脸图像进行下采样过程中加入了特征图对称模块,首先使用预训练的人脸特征点检测器视网膜人脸(Retinaface)^[17]对输入侧脸图像进行特征点检测,通过鼻尖点位置坐标来判断当前输入侧脸的朝向,将侧脸图像中信息包含较多的一侧以鼻尖横坐标为中心做镜像对称,编码器提取的特征图和对称后的特征图可视化效果如图 3 所示,对称特征图应用于解码器跳层拼接,以补全丢失的人脸信息。编码器继续对侧脸特征图进行下采样得到 256 维的深层人脸特征表达向量。

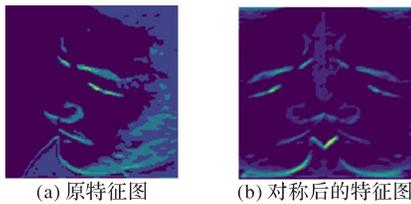


图 3 特征图对称模块

Fig. 3 Feature map symmetry module

2.2 改进的解码器跳层拼接

解码器作用是将编码器得到的高层特征依次经过多个反卷积层进行上采样,直到生成图的分辨率和输入图一致。本文对 TP-GAN 解码器跳层拼接做出改进,TP-GAN 在跳层拼接时只将侧脸特征图、输入原图和反卷积特征图进行拼接,如图 4 所示。本文的改进方法为:在解码过程中,每层反卷积输入的特征图由编码器中对应尺度侧脸特征图、特征图对称模块得到的对称特征图、同一尺寸的输入侧脸图以及前一层反卷

积的输出特征图拼接得到。相较于 TP-GAN 的拼接方式,本文得到的对称特征图也参与了拼接,这种拼接方式不仅可以实现特征融合,也可以在特征层面上对丢失侧脸特征进行模拟和补全,以缓解生成器生成过程中能利用的有效面部信息较少的困难,从而生成出具有高细节保真的正面图像。

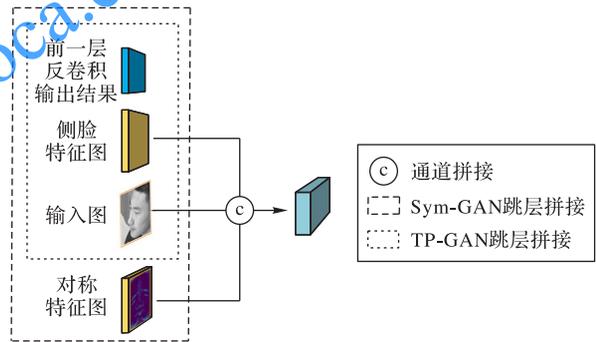


图 4 改进的解码器跳层拼接

Fig. 4 Skip layer splicing of improved decoder

2.3 判别器结构

判别器的作用是判别生成图和真实图的真假性,从而对生成器造成一定的惩罚,使得生成器生成的结果更接近真实图像。本文采用的判别器结构如图 5 所示,由 6 层卷积层组成,其中在第 4 层和第 5 层卷积层之后加入残差块,使得网络层数加深,从而提取到更为抽象的面部细节特征,提高判别器的判别能力。从第 5 层到第 6 层采用 1x1 卷积核用于降低特征通道,最后输出 4x4 概率图。本文采用 1x1 卷积核代替全连接层,是因为全连接层会破坏图像的空间结构,而 1x1 卷积核可以保证在降低特征图维度的同时不改变图像的空间结构,即不破坏面部拓扑结构。此外,相较于传统生成对抗网络中判别器只输出一个标量值来表示生成图或真实图整幅图像的真实性,本文的判别器输出 4x4 概率图,其中每一个概率值对应图像中一个局部感受野,可以恰好将人脸五官位置进行分离判别,采用这种方法可以提高判别器对局部细节的判别能力,从而促使生成器生成高细节保留、高分辨率的正面人脸图像。

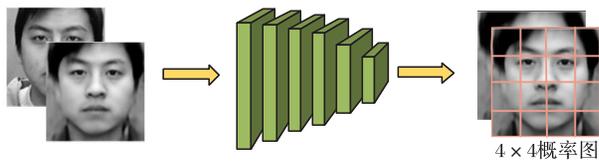


图5 判别器网络结构

Fig. 5 Network structure of discriminator

3 损失函数设计

3.1 整体目标函数

用于训练Sym-GAN的目标函数 L_{syn} 由一系列损失函数加权组成,公式如下所示:

$$L_{syn} = \lambda_1 L_{ip} + \lambda_2 L_{eyefp} + \lambda_3 L_{pixel} + \lambda_4 L_{adv} + \lambda_5 L_{sym} + \lambda_6 L_{tv} \quad (1)$$

其中: L_{ip} 表示身份保留损失,以全局角度保留生成图身份信息;在 L_{ip} 的基础上,为了增加对生成图局部身份信息的关注,本文提出眼周特征保留损失 L_{eyefp} ; L_{pixel} 表示多尺度像素级损失,确保生成图和真实图的内容一致性; L_{adv} 表示对抗损失,使生成图在视觉效果上较好; L_{sym} 表示对称损失,用于缓解自遮挡问题; L_{tv} 表示总差分正则化,降低生成图的人工伪影,提高视觉效果; $\lambda_1, \lambda_2, \lambda_3, \lambda_4, \lambda_5, \lambda_6$ 为对应的权重参数。下面将详细介绍每一个损失函数。

3.2 眼周特征保留损失

为了保留生成图身份信息,文献[7-8]利用轻卷积神经网络(Light Convolutional Neural Network, Light CNN)^[18]提取生成图和真实图的高维特征向量构成感知损失来确保生成图身份信息保留,这种方式是以全局角度关注生成图的身份特征,而对面部局部区域的身份信息关注较少;文献[13-14]指出眼周区域(即包含眉毛和眼睛的矩形区域)是面部最具有判别能力的区域之一,表明眼周区域相较于面部其他五官,可以表征出更加独特的人脸身份特征。因此,为了提高生成图全局和局部区域的身份特征保留,在文献[7-8]身份保留损失的基础上,提出眼周特征保留损失,文献[19]分析了传统特征提取方法,如局部二值模式(Local Binary Pattern, LBP)^[20]、方向梯度直方图(Histogram of Oriented Gradient, HOG)^[21]、尺度不变特征变换(Scale-Invariant Feature Transform, SIFT)^[22]和现有卷积特征在眼周图像识别效果,提出了现有的卷积特征在眼周图像上也有较好的识别性能,且实验表明使用残差网络(Residual Network, ResNet)^[23]ResNet50在眼周图像上的第73层特征识别效果相对较好。因此,本文按照如图6所示设计了眼周特征保留损失,利用ResNet50分别提取生成图和真实图的左右眼区域的第73层特征,然后分别计算对应特征之间的余弦距离,将两部分距离相加得到眼周特征保留损失,作为训练生成器损失函数的一部分,约束局部区域身份保留。

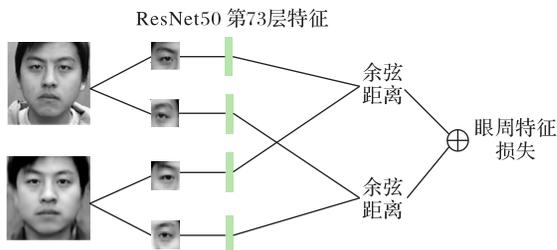


图6 眼周特征保留损失计算

Fig. 6 Calculation of preserving loss of periorcular feature
眼周特征保留损失函数公式如下所示:

$$L_{eyefp} = dist(\phi(I_{le}^{gen}), \phi(I_{le}^{st})) + dist(\phi(I_{re}^{gen}), \phi(I_{re}^{st})) \quad (2)$$

其中: $\phi(\cdot)$ 定义为ResNet50提取的第73层网络的特征; I_{le}^{gen} 和 I_{re}^{st} 分别表示生成图像和基准图像; I_{le}^{*} 和 I_{re}^{*} 分别表示 I^* 的左右眼区域图像; $dist$ 表示余弦距离。

3.3 身份保留损失

为了保留生成图像全局身份信息,按照文献[7-8],使用预训练人脸识别网络提取的高维特征向量构成感知损失来保留生成图的身份信息。

$$L_{ip} = \|\varphi(I^{gen}) - \varphi(I^{st})\|_2^2 \quad (3)$$

其中: $\varphi(\cdot)$ 定义为预训练的人脸识别网络Light CNN提取的最后一层全连接层的特征,由于Light CNN是在大规模人脸数据集上进行训练,它可以提取到更为普遍、更为显著的人脸特征; $\|\cdot\|_2^2$ 表示向量2范数。

3.4 多尺度像素级损失

按照文献[7],加入多尺度像素级损失来确保生成图像 I^{gen} 和基准图像 I^{st} 在内容上的一致性,此外也可提高生成图像 I^{gen} 的逼真程度。

$$L_{pixel} = \frac{1}{3} \sum_{i=1}^3 \frac{1}{W_i H_i C} \sum_{w,h,c=1}^{W_i, H_i, C} |I_{i,w,h,c}^{gen} - I_{i,w,h,c}^{st}| \quad (4)$$

其中: i 表示第 i 个图像尺寸,分别在图像尺寸为 128×128 、 64×64 、 32×32 大小的生成图 I^{gen} 和真实图 I^{st} 之间计算像素值的L1距离; C 表示图像的通道数, W_i 和 H_i 分别表示对应分辨率图像的宽和高。

3.5 对抗损失

Sym-GAN由两部分组成:生成器G和判别器D。判别器D的目的是从真实图和生成图中区分出真实图。生成器G的目的是生成接近真实的图像以欺骗判别器。生成器和判别器之间的训练^[12]可表示为:

$$L_{adv} = \min_{\theta_G} \max_{\theta_D} E[\log D_{\theta_D}(I^{st})] + E[\log(1 - D_{\theta_D}(I^{gen}))] \quad (5)$$

对抗损失 L_{adv} 可以帮助生成图在视觉感官上产生较好的效果。

3.6 对称损失

文献[7]提出的对称损失,可以有效地缓解自遮挡问题以提高模型在大角度姿态下的正面化生成性能。

$$L_{sym} = \frac{1}{W/2 \times H} \sum_{w=1}^{W/2} \sum_{h=1}^H |I_{w,h}^{gen} - I_{W-(w-1),h}^{gen}| \quad (6)$$

其中: $W - (w - 1)$ 表示生成图 I^{gen} 中 w 位置的对称横坐标。

3.7 总变分正则化

通常,GAN模型生成的图像会产生许多的人工伪影,干扰视觉质量和识别性能,按照文献[3],在生成图像中加入了总变分正则化项以缓解这一问题。

$$L_{tv} = \sum_{c=1}^C \sum_{w,h=1}^{W,H} |I_{w+1,h,c}^{gen} - I_{w,h,c}^{gen}| + |I_{w,h+1,c}^{gen} - I_{w,h,c}^{gen}| \quad (7)$$

其中: W 和 H 表示最终生成图像的宽和高。

4 实验与结果分析

Sym-GAN的核心目标是根据包含多角度变化的侧脸合成其逼真的正面人脸,同时能够保留足够的身份信息。在4.2节中展示了提出的Sym-GAN在公开数据集上的定性正面化合成效果,在4.3节中定量评估了Sym-GAN对人脸识别网络的增强效果。在4.4节中,针对Sym-GAN不同结构变体和



损失函数变体进行了消融实验,以分析其各自的作用。

4.1 实验设置

CAS-PEAL-R1 数据集是一个公开的大规模中国人脸数据集,它涵盖了姿态、表情、装饰物和光照等变化。该数据集总共包含 1 040 个人的 30 863 张灰度图像,其中男性有 595 人、女性有 445 人。本文中只使用含有姿态变化的图像,包含 7 个偏转角度(0°, ± 15°, ± 30°, ± 45°)和 3 个俯仰角度(0°, ± 30°),共 21 种角度变化。使用前 600 人构成训练集,剩下的 440 人构成测试集。

使用人脸检测模型 Retinaface 对图像进行预处理,将 CAS-PEAL-R1 数据集中用于实验评估的所有图像裁剪到分辨率为 128×128,并且为了保证正面化的效果,对所有基准正面图像进行对齐处理,即保证左右眼特征点处于同一水平线上,且确保两眼连线的中点落在裁剪后图像的中点上。用于

计算身份损失的预训练人脸识别网络 Light CNN-29 在训练期间参数固定,眼周区域大小选取数据集眼周矩形框的平均尺寸。本文的模型采用 Pytorch 深度学习框架实现。选择 β_1 为 0.9, β_2 为 0.99 的 Adam 优化器,学习率为 2^{-4} ,图像分辨率为 128×128 时设置 batch size 为 16。在一张 NVIDIA Tesla P100 GPU(16 GB)上训练网络。在所有实验中,依据经验分别设置权重参数 $\lambda_1, \lambda_2, \lambda_3, \lambda_4, \lambda_5, \lambda_6$ 为 0.01, 0.05, 150, 0.1, 0.01, 1.0。

4.2 人脸正面化效果

本文以侧脸图像作为输入图像,以偏转角和俯仰角都为 0° 的正面图像作为基准图像训练网络。图 7 所示为 Sym-GAN 在各个俯仰角的生成结果。可以观察到,生成图像和基准图像不仅在全局的面部结构上保持一致,同时在眉毛和眼型等局部细节上也与基准图像接近,证明了本文提出的眼周特征保留损失对局部区域特征保留具有一定的作用。

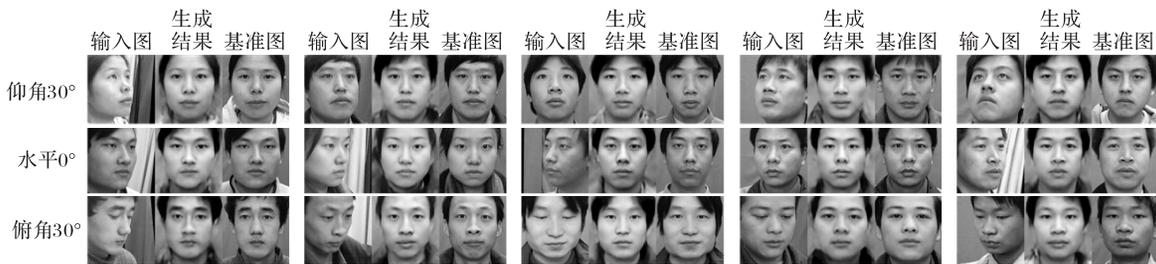


图 7 Sym-GAN 在 CAS-PEAL-R1 数据集上不同俯仰角的生成正面人脸图像

Fig. 7 Frontal images generated by Sym-GAN on images with different pitch angles in CAS-PEAL-R1 dataset

本文提供了一系列人脸正面化方法的生成结果与 Sym-GAN 进行比较,包括 TP-GAN^[7]、完整表示生成对抗网络(Complete Representations Generative Adversarial Network, CR-GAN)^[24]、M2FPA^[8]、DA-GAN^[16],生成效果对比结果如图 8 所示。由图 8 可以看出,在眼周部分 Sym-GAN 模型的生成效果相较其他方法与基准图像更为接近。此外,相较其他方法在判别器加入一些区域约束,会造成其他区域出现模糊的现象,本文方法使用相对简单的判别器结构,保证了面部整体细节的一致性。

4.3 姿态无关的人脸识别

人脸正面化可以作为人脸识别模型的一种预处理操作,所以使用人脸识别准确率作为评价指标来评估不同正面化方法的身份保留能力。识别准确率越高,人脸合成过程中保留下的人脸特征越多,正面化合成的效果越好。因此,本文在 CAS-PEAL-R1 数据集上定量验证 Sym-GAN 模型“通过生成再进行识别”的有效性。把 Sym-GAN 模型作为预处理操作,然后使用预训练人脸识别模型 Light CNN 作为人脸特征提取器,并使用距离度量计算生成图像和真实图像对之间的相似度。表 1 所示为不同方法在 CAS-PEAL-R1 数据集上的 Rank-1 识别准确率。从表 1 中可以观察到,在一些极端角度上,本文的方法可以得到更高的识别准确率,证明了本文方法在侧脸正面化任务上的有效性。

4.4 消融实验

为了验证 Sym-GAN 模型的优势以及其各个组成部分的贡献,进行了消融对比实验。图 9 为本文方法与其 3 个不完全变体在 CAS-PEAL-R1 数据集上的可视化比较结果。从图 9 可以看出,若消去总变分正则化损失,会产生更多的人工伪影;若消去身份信息保留损失,在脸型轮廓上和基准图像差距较大;此外,加入眼周特征保留损失,可以使眉毛和眼型上与基准图像更加接近,并且网络训练中有无眼周特征保留损失对身份信息保留损失收敛的对比如图 10(a)和(b)所示,加入眼周特征保留损失有助于身份信息保留损失的收敛,从而对面部全局身份信息保留起到一定积极作用。表 2 为本文方法与其 3 个不同变体在 CAS-PEAL-R1 数据集上的 Rank-1 识别率,也可以展示出身份信息保留损失和眼周特征保留损失在身份信息保留过程中的有效作用。

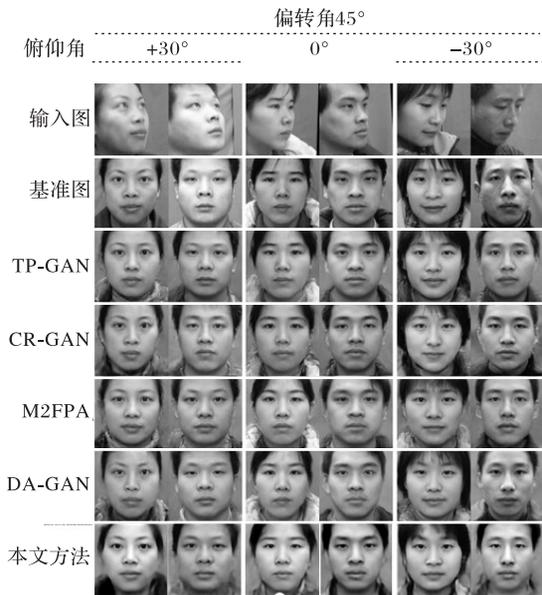


图 8 CAS-PEAL-R1 数据集上不同方法的比较

Fig. 8 Comparison of different methods on CAS-PEAL-R1 dataset



表 1 CAS-PEAL-R1 数据集上不同方法的 Rank-1 识别率
Tab. 1 Rank-1 recognition rates of different methods on CAS-PEAL-R1 dataset

方法	俯仰-30°偏转					俯仰0°偏转				俯仰+30°偏转				
	0°	±15°	±30°	±45°	平均	±15°	±30°	±45°	平均	0°	±15°	±30°	±45°	平均
TP-GAN	98.86	98.94	98.89	97.62	98.58	100.00	99.94	98.71	99.55	97.68	97.73	97.45	95.83	97.17
CR-GAN	83.98	83.91	83.17	80.38	82.86	97.61	95.80	89.73	94.38	89.74	89.44	87.95	83.90	87.76
M2FPA	99.38	99.42	99.30	98.53	99.16	100.00	99.94	99.36	99.77	98.60	98.69	98.58	97.84	98.43
DA-GAN	99.71	99.72	99.65	98.99	99.52	100.00	100.00	99.70	99.90	98.96	98.98	98.86	98.13	98.73
本文方法	99.72	100.00	99.72	95.25	98.67	100.00	100.00	99.72	99.91	100.00	99.72	99.44	96.94	99.03

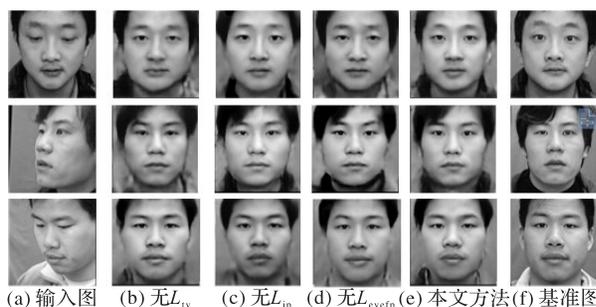


图 9 Sym-GAN 模型及其变体在 CAS-PEAL-R1 数据集上的生成效果

Fig. 9 Generation results of Sym-GAN model and its variants on CAS-PEAL-R1 dataset

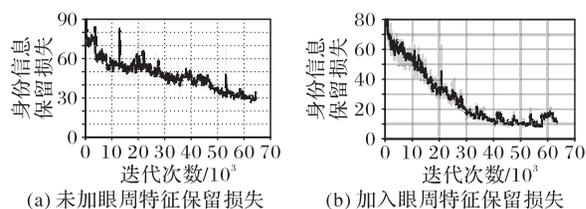


图 10 身份信息保留损失曲线对比

Fig. 10 Comparison of identity information preserving loss curves

表 2 CAS-PEAL-R1 数据集上各消融结果的 Rank-1 识别率 单位:%
Tab. 2 Rank-1 recognition rates of ablation results on CAS-PEAL-R1 dataset

方法	俯仰-30°偏转		俯仰0°偏转		俯仰+30°偏转	
	±0°	±15°	±15°	±45°	±15°	±30°
无 L_{tv}	98.89	99.72	100.00	98.06	99.44	98.61
无 L_{ip}	97.21	96.67	99.44	93.61	98.06	95.28
无 L_{eyefp}	98.41	97.73	100.00	97.73	98.85	98.18
本文方法	99.72	100.00	100.00	99.72	99.72	99.44

5 结语

本文针对多角度侧脸正面化问题提出了一种基于面部特征图对称和眼周特征保留损失的生成对抗网络模型。首先,根据人脸对称性这一先验知识,本文提出了特征图对称模块,用于缓解大角度姿态下人脸存在的自遮挡现象造成正面化生成较为困难的问题。此外,受益于眼周区域相较于面部其他五官,具有较高的判别能力,加入了眼周特征保留损失使得在生成过程中对眼周区域增大关注,从而提高生成图的身份信息保留程度。定性的正面化结果和定量的人脸识别性能的实验验证了本文方法的有效性。同时,将对面部各五官特征对人脸识别的贡献度进行调研,在眼周保留损失的基础上尝试加入面部其他高辨别区域的身份特征保留从而进一步提高生成人脸的效果。

参考文献 (References)

[1] TRAN L, YIN X, LIU X. Disentangled representation learning

GAN for pose-invariant face recognition [C]// Proceedings of the 2017 IEEE Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE, 2017: 1283-1292.

[2] CAO K, RONG Y, LI C, et al. Pose-robust face recognition via deep residual equivariant mapping [C]// Proceedings of the 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE, 2018: 5187-5196.

[3] HU Y, WU X, YU B, et al. Pose-guided photorealistic face rotation [C]// Proceedings of the 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE, 2018: 8398-8406.

[4] 钱一琛. 基于生成对抗的人脸正面化生成[D]. 北京:北京邮电大学, 2019: 73. (QIAN Y C. Generative adversarial networks based face frontalization [D]. Beijing: Beijing University of Posts and Telecommunications, 2019: 73.)

[5] KAN W, SHAN S, CHANG H, et al. Stacked Progressive Auto-Encoders (SPA-E) for face recognition across poses [C]// Proceedings of the 2014 IEEE Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE, 2014: 1883-1890.

[6] YIM J, JUNG H, YOO B, et al. Rotating your face using multi-task deep neural network [C]// Proceedings of the 2015 IEEE Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE, 2015: 676-684.

[7] HUANG R, ZHANG S, LI T, et al. Beyond face rotation: global and local perception GAN for photorealistic and identity preserving frontal view synthesis [C]// Proceedings of the 2017 IEEE International Conference on Computer Vision. Piscataway: IEEE, 2017: 2458-2467.

[8] LI P, WU X, HU Y, et al. M2FPA: a multi-yaw multi-pitch high-quality database and benchmark for facial pose analysis [C]// Proceedings of the 2019 IEEE/CVF International Conference on Computer Vision. Piscataway: IEEE, 2019: 10042-10050.

[9] HASSNER T, HAREL S, PAZ E, et al. Effective face frontalization in unconstrained images [C]// Proceedings of the 2015 IEEE Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE, 2015: 4295-4304.

[10] ZHU X, LEI Z, YAN J, et al. High-fidelity pose and expression normalization for face recognition in the wild [C]// Proceedings of the 2015 IEEE Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE, 2015: 787-796.

[11] 丁立军,冯浩,华亮. Clifford代数3D人脸姿态矫正方法[J]. 小型微型计算机系统, 2013, 34(4): 906-909. (DING L J, FENG H, HUA L. Clifford algebra approach for 3D face pose correction [J]. Journal of Chinese Computer Systems, 2013, 34 (4) : 906-909.)

[12] GOODFELLOW I, POUGET-ABADIE J, MIRZA M, et al.



- Generative adversarial nets [C]// Proceedings of the 27th International Conference on Neural Information Processing Systems. Cambridge: MIT Press, 2014: 2672-2680.
- [13] ALONSO-FERNANDEZ F, BIGUN J. A survey on periocular biometrics research[J]. Pattern Recognition Letters, 2016, 82(Pt 2): 92-105.
- [14] ZHAO Z, KUMAR A. Improving periocular recognition by explicit attention to critical regions in deep neural network [J]. IEEE Transactions on Information Forensics and Security, 2018, 13(12): 2937-2952.
- [15] GAO W, CAO B, SHAN S, et al. The CAS-PEAL large-scale Chinese face database and baseline evaluations [J]. IEEE Transactions on Systems, Man, and Cybernetics - Part A: Systems and Humans, 2008, 38(1): 149-161.
- [16] YIN Y, JIANG S, ROBINSON J P, et al. Dual-attention GAN for large-pose face frontalization [EB/OL]. [2020-04-25]. <https://arxiv.org/pdf/2002.07227.pdf>.
- [17] DENG J, GUO J, ZHOU Y, et al. RetinaFace: single-stage dense face localisation in the wild [EB/OL]. [2020-04-25]. <https://arxiv.org/pdf/1905.00641.pdf>.
- [18] WU X, HE R, SUN Z, et al. A light CNN for deep face representation with noisy labels [J]. IEEE Transactions on Information Forensics and Security, 2018, 13(11): 2884-2896.
- [19] HERNANDEZ-DIAZ K, ALONSO-FERNANDEZ F, BIGUN J. Periocular recognition using CNN features off-the-shelf [C]// Proceedings of the 2018 International Conference of the Biometrics Special Interest Group. Piscataway: IEEE, 2018: 1-5
- [20] OJALA T, PIETIKAINEN M, MAENPAA T. Multiresolution gray-scale and rotation invariant texture classification with local binary patterns[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2002, 24(7): 971-987.
- [21] DALAL N, TRIGGS B. Histograms of oriented gradients for human detection [C]// Proceedings of the 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE, 2005: 886-893.
- [22] LOWE D G. Distinctive image features from scale-invariant keypoints[J]. International Journal of Computer Vision, 2004, 60(2): 91-110.
- [23] HE K, ZHANG X, REN S, et al. Deep residual learning for image recognition [C]// Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE, 2016: 770-778.
- [24] ZHAO J, CHENG Y, XU Y, et al. Towards pose invariant face recognition in the wild [C]// Proceedings of the 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE, 2018: 2207-2216.

This work is partially supported by the Shanxi Provincial Key Research and Development Program (201803D31212-1).

LI Hongxia, born in 1996, M. S. candidate. Her research interests include deep learning, computer vision.

QIN Pinle, born in 1978, Ph. D., professor. His research interests include machine vision, big data, medical image analysis.

YAN Hanmei, born in 1968, associate professor. Her research interests include criminal science and technology, criminal imageology.

ZENG Jianchao, born in 1963, Ph. D., professor. His research interests include maintenance decision of complex system, health management.

BAO Qianyue, born in 1998. His research interests include deep learning, computer vision.

CHAI Rui, born in 1985, Ph. D., lecturer. His research interests include medical image processing.