

基于生成对抗网络的多姿态人脸识别算法

蒋文豪¹

JIANG Wenhao

摘要

头部姿态角转换会造成人脸成像多姿态变化,人脸离散数据的高斯分布混乱,无法准确地反映人脸多姿态的任意性和连续性,存在识别效果差的问题。引入生成对抗网络理论,设计多姿态人脸识别算法。对获取到的不同角度人脸图像,实施多姿态人脸校正与旋转残差注意力计算,解决当前头部姿态估计方法对不同人脸感兴趣区域不稳健的问题。设计生成对抗网络进行双路循环优化,在生成的对抗网络中,参考CASIA-Net网络结构,使用深层次网络结构,每一层都有一个3*3的卷积核。所提出的设计可以降低网络参数,增强网络的非线性度,实现高效的面部特征提取,构建人脸多姿态识别模型,并完成人脸识别。通过实验结果表明,所提算法针对多姿态人脸识别效果好,在人脸不同姿态变化过程中,识别率始终在97%以上,更适用于多姿态人脸识别。

关键词

生成对抗网络;残差注意力;人脸识别;多姿态;循环优化;识别模型

doi: 10.3969/j.issn.1672-9528.2024.02.042

0 引言

在现实生活中,人们经常使用图片来实现面孔的识别。通过人脸对身份信息^[1]的识别优势在于以下几个方面。第一,可以实现对象与设备之间的零接触。与指纹识别技术的区别在于,在面部识别的过程中,识别对象只需要站在识别设备前一段距离就可以完成识别任务,而不需要与识别设备产生直接的接触。第二,具有快速性和简便性。在被识别的物体通过识别装置时,识别装置内的图像采集系统会将所获取的人脸图像迅速地传送给识别系统,以便对其进行身份分析和验证。第三,具备丰富的数据库。在进行人脸识别的研究时,为了对多种不同因素下的人脸识别技术进行充分研究,学者们构建了多种类型丰富的人脸信息库,这大大推动了人脸识别技术的发展。但是在实际应用中,人脸的姿态经常会发生变化,例如在人脸识别、人脸跟踪、人脸动画等领域中。如果人脸识别算法只能处理正面人脸或者固定姿态的人脸,那么在实际应用中就会受到很大的限制。因此,研究多姿态人脸识别算法可以提高识别准确性,使其能够更好地适应实际应用中的需求。

古振猛等人^[2]为了解决卷烟零售业在收货拍照过程中身

份验证不一致的难点问题,提出了基于深度学习的识别算法,结合机器学习和卷积神经网络完成人脸特征的识别。李云红等人^[3]提出了基于超分辨重建的识别算法,基于生成对抗网络结构,提取网络特征,采用公共特征子空间法获取目标函数,结合损失函数实现最终的人脸识别。这些人脸识别方法已经大大提升了识别能力,但在实际应用中人脸与采集摄像机之间的角度变换会使得采集的人脸图像发生位置上的变化,进而引起自遮挡效应,最终造成采集到的人脸图像与真实采集对象的人脸发生较大程度的变化,而变化越大,人脸识别的正确率越低。

针对这一问题,本文下面将结合生成对抗网络,开展对多姿态人脸识别算法的研究。

1 多姿态人脸校正与旋转残差注意力计算

考虑到在多姿态下获取到的人脸可能是不同头部姿态角度下的人脸图像,为了确保后续对其识别时的准确性,需要对多姿态人脸进行校正和旋转,并提取旋转相关的特征。在获取侧面人脸图像后,从正面人脸生成器上生成具有相同身份信息的正常人脸图像^[4]。

将人脸校正阶段中使用的判别器看作是一种多任务判别器,所需要完成的任务包括以下两方面。一是逐步地改进输入人脸图像源的判别能力^[5]。二是对输入的人脸影像进行残差注意力计算,以确保校正后的人脸影像与原人脸影像属于同一训练目标,解决当前头部姿态估计方法对不同人脸感兴趣区域不稳健的问题。多姿态人脸校正后的结果可表示为:

1. 重庆航天职业技术学院 重庆 400021

[基金项目] 项目来源:重庆市教育委员会科学技术研究项目,项目名称:基于深度学习的人脸识别关键算法研究,项目编号: KJZD-K202003001

$$D_N = [D_N^{(r)}, D_N^{(d)}] \quad (1)$$

式中： D_N 表示判别器输出的多姿态人脸校正结果； $D_N^{(r)}$ 表示获取到的图像可表示为真实图片的概率， $D_N^{(r)}$ 的取值越大，说明生成的图像越接近真实图像数据； $D_N^{(d)}$ 表示被识别图像的身份分类结果，参与到训练中的样本数量，用来保证随后的模型输出和输入是相同的训练目标。

在多姿态的人脸校正阶段，由于人脸图像只包含多个姿态，所以需要对其得到的姿态进行分类^[6]。对判别器当中获取到的图像可表示为真实图片的概率 $D_N^{(r)}$ 部分的对抗损失函数进行计算，以此实现在使产生图像的输出概率最小的情况下，使图像的输出概率最大，以增强图像的识别能力。损失函数的表达式为：

$$L_{D_N}^{(adv)} = E[-\ln[D_N(x)]] \quad (2)$$

式中： $L_{D_N}^{(adv)}$ 表示校正过程中产生的损失量； x 表示真实图像分布数据； E 表示身份信息标签值。在图像制作等工作中，减少两幅图像像素之间的差异，可以改善人脸图像品质^[7]。

针对这一问题，本文提出在双循环生成对抗网络中引入像元残差注意力函数，通过该函数来表征低频信息，衡量生成器输出与真实目标图像之间的像元差异，从而提高网络的特征抽取与场景构建能力^[8]。类似地，面部旋转部分的判别器也属于多任务判别器，其主要完成三个功能：第一，判定输入人脸的源头；第二，识别输入人脸的位置；第三，识别输入人脸的姿态。人脸旋转最终输出的部分可表示为：

$$D_E = L_{D_N}^{(adv)}[D_E^{(r)}, D_E^{(d)}, D_E^{(p)}] \quad (3)$$

式中： D_E 表示人脸旋转最终输出结果； $D_E^{(r)}$ 表示通过判定输入侧面部影像的来源，从实际画面资料中提取的概率； $D_E^{(d)}$ 表示分类结果，以确保输入和输出是同一训练对象； $D_E^{(p)}$ 表示通过对输入人脸图像的姿态分类，通过该参数的加入，确保输出结果与给定姿态相符合^[9]。

在此基础上，进行第二个步骤，对判别器在生成多姿态人脸图像分类时的残差注意力函数进行计算，其公式为：

$$L_{D_E}^{(adv)} = -\ln[D_E(x, c)] \quad (4)$$

式中： $L_{D_E}^{(adv)}$ 表示判别器在生成多姿态人脸图像分类时的残差注意力函数； c 表示对应姿态。综合上述人脸旋转部分涉及残差注意力函数，在校正和旋转时进行补偿，确保校正和旋转后的人脸图像更接近实际，解决当前头部姿态估计方法对不同人脸兴趣区域不稳健的问题。

2 人脸识别生成对抗网络中双路循环方法设计

人脸在不同头部变化姿态下的形状和纹理都会发生变化，通过对多姿态人脸进行校正和旋转，可以使得不同姿态下的人脸图像具有相同的姿态和形状，从而减少了姿态变化对人脸识别的影响，获取更加精准的双路循环优化函数，提高识别精度。为了能够进一步提高人脸识别算法的识别性能，

引入生成对抗网络，将上述人脸校正和旋转内容相结合，通过增加的对抗网络进行网络约束^[10]。生成对抗网络中包含两条循环路径，利用Cycle-1实现人脸侧面—正面—侧面的转换。假设输入的侧面人脸图像为 $x_i \sim P_{data}$ ，在这一过程中双路循环优化函数为：

$$L_{cyc-1} = \frac{1}{n} \sum_i \|x_i - G_E[G_N(x_i), L_{D_E}^{(adv)}]\| \quad (5)$$

式中： L_{cyc-1} 表示在人脸侧面—正面—侧面的转换双路循环中的优化函数； n 表示循环次数； x_i 表示某一真实图像 i 的分布数值； G_E 表示侧面人脸生成器； G_N 表示正面人脸图像。还需要利用Cycle-2实现人脸正面—侧面—正面的转换，利用侧面人脸生成器对正面人脸进行处理，再将输入的侧面人脸图像还原为与输入图像相似的正面人脸图像。在这一过程中双路循环优化函数为：

$$L_{cyc-2} = \frac{1}{n} \sum_i \|x_i - G_N[G_E(x_i), L_{cyc-1}]\| \quad (6)$$

式中： L_{cyc-2} 表示在人脸正面—侧面—正面的转换双路循环中的优化函数。将上述两种训练的优化函数作为目标，在生成的对抗网络当中，实现对面脸识别的双路循环优化。

3 基于对抗网络的人脸识别算法设计

通过第2小节对生成对抗网络进行双路循环优化，可以提高识别模型在处理多姿态人脸时的稳定性。这是因为双路循环优化可以突出多姿态人脸特征，从而保证在处理不同姿态下的人脸时，能够保持较好的性能。对已有的多姿态人脸识别网络进行探究，该网络对特定姿态的人脸具有很好的识别能力。但是，还有两个地方可以改进。第一，在面部姿态改变超过60°的情况下，相应的人脸识别精度仍然会有很大的降低，需要进一步地改进^[11]。第二，已有的多姿态面孔识别网络将人脸校正和人脸旋转两个独立的过程融合到一个网络中，增加了网络训练的难度。在生成的对抗网络中，参考CASIA-Net网络结构，使用深层次网络结构，每一层都有一个3*3的卷积核。该设计可以降低网络参数，增强网络的非线性度，实现高效的面部特征提取，使用卷积层取代原来的MaxPool层，并使用ReLU函数来增强TC-GAN网络的稳定性。表1为TC-GAN网络的基本结构。

表1 TC-GAN网络的基本结构

序号	层数	过滤器	步幅	输出尺寸
(1)	Conv11	3×3	1	24×24×128
(2)	Conv12	3×3	1	24×24×96
(3)	Conv21	3×3	2	24×24×48
(4)	Conv22	3×3	1	48×48×64
(5)	Conv23	3×3	1	48×48×128
(6)	Conv31	3×3	2	96×96×32
(7)	Conv32	3×3	1	96×96×64
(8)	Conv33	3×3	1	96×96×128

在该网络结构当中, 所得到的编码器由一系列卷积层所组成, 这些卷积层卷积在 96×96 的输入大小的人脸图像上, 并且在每个卷积层卷积之后, 进行批量归一化处理。其中, 采用 ELU 作为启动函数, 保证了 TC-GAN 网络的稳定。通过一系列的卷积运算, 得到一个 $(1, 1, 320)$ 的三维张量。在此基础上, 提取人脸图像的特征点, 基于特征点与人脸局部部位之间存在的关系, 识别每张人脸图像 i , 获取相应重要部位的图像集, 公式如下所示:

$$P_i = \left\{ L_{\text{cyc}-2} \left(P_i^j, t_i^j \right) \right\}_{j=1}^n \quad (7)$$

式中: P_i^j 表示人脸图像 i 的第 j 个重要部位图像; t_i^j 表示重要部位图像对应的面部部位编码标签。最后, 对获取到的重要部位图像集的特征编码进行解码处理, 输出的结果即为识别得到的多姿态人脸识别结果。

4 对比实验与结果分析

通过上述论述, 完成对基于生成对抗网络的人脸识别算法的理论设计。为了验证新的算法在实际应用中的可行性, 以及是否能够有效解决现有识别方法在实际应用中存在的问题, 将该识别算法设置为实验组, 将基于深度学习的识别算法设置为对照 A 组, 将基于超分辨重建的识别算法设置为对照 B 组, 开展下述对比实验研究。

设置实验环境配置条件为 64 bit Windows10 操作系统; 内存为 32 GB, GPU 为 NVIDIA GeForce RTX 1080Ti。将 Multi-PIE 作为实验过程中所使用的人脸数据集, 其中包含了 565 名身份不同的对象, 并且每一个对象的人脸图像包含了多种姿态变化、多种光照变化和表情变化。从人脸数据集中选择 500 张图像, 将其作为三组识别算法的训练数据集。在训练之前, 首先要对图像进行预处理, 即从图像中抽取人脸, 减少非人脸区域的干扰。然后为了减少在训练中出现的过拟合问题, 采用了人脸检测器来实现人脸的识别, 在对齐所有图片尺寸后, 采用随机裁切器将其剪切成标准的 $96 \text{ pixel} \times 96 \text{ pixel}$ 尺寸, 从而减少了在训练时出现的过拟合问题。最后为实现对三组识别算法应用性能的验证, 以不同姿态下识别算法的正确识别率为评价指标, 分别设置五组不同姿态变化范围: 0° 、 $-30^\circ \sim 30^\circ$ 、 $-45^\circ \sim 45^\circ$ 、 $-60^\circ \sim 60^\circ$ 、 $-90^\circ \sim 90^\circ$, 如图 1 所示。

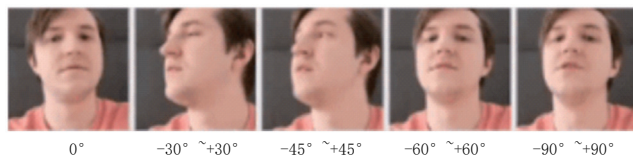


图 1 五组不同姿态人脸变化示意图

针对多姿态人脸校正与旋转残差注意力计算和双循环约束, 人脸三维特征的变化过程如图 2 所示。

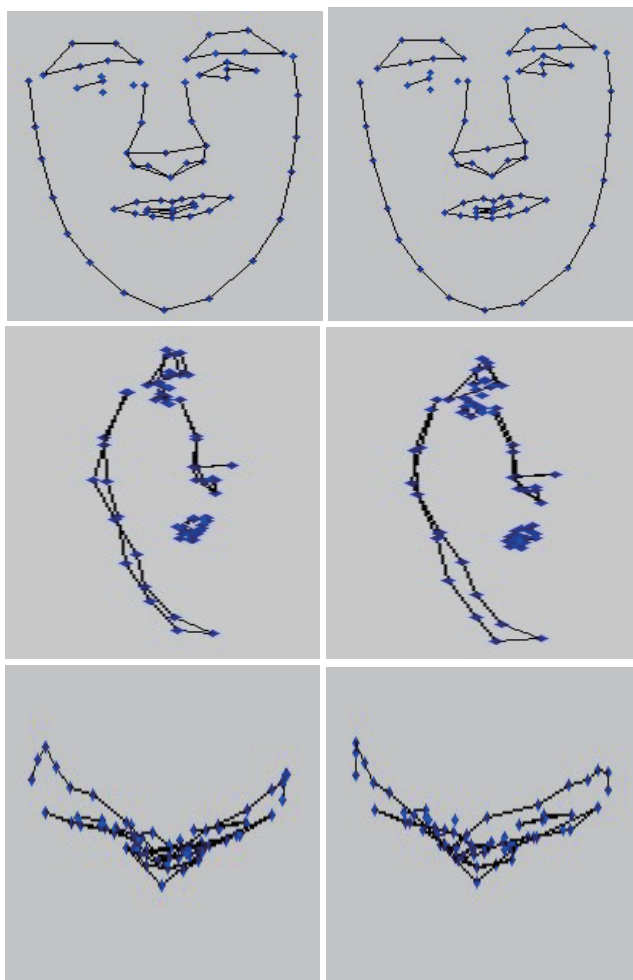


图 2 加入双循环和残差注意力约束的特征稳定对比

人脸旋转部分涉及残差注意力函数和双循环约束后, 在校正和旋转时进行补偿, 确保校正和旋转后的人脸图像更接近实际, 解决当前头部姿态估计方法对不同人脸兴趣区域不稳健的问题。下面进行识别算法测试。

利用本文提出的识别算法, 结合上述五组不同姿态人脸, 对图 1 所示的人脸图像进行识别, 并得到识别结果。图 2 中给出了左脸的正视、侧视以及俯视三个视角下的特征识别结果比较图, 图中红色圆圈表示未加入残差注意力约束下的特征识别结果, 蓝色点表示加入残差注意力约束后的识别结果。侧视图和俯视图能够清晰地反映出红圈表示的眼睛过于向里凹陷, 与眉毛的相对位置不符, 而蓝点表示的眼睛与眉毛的相对位置较准确。这说明加入约束项后识别过程更加连续, 识别位置更加准确。从图 3 得到的识别结果可以看出, 应用本文提出的识别算法可以针对目标人脸图像刻画基本轮廓, 进而准确识别到相对应的人脸图像, 具备极高识别精度。

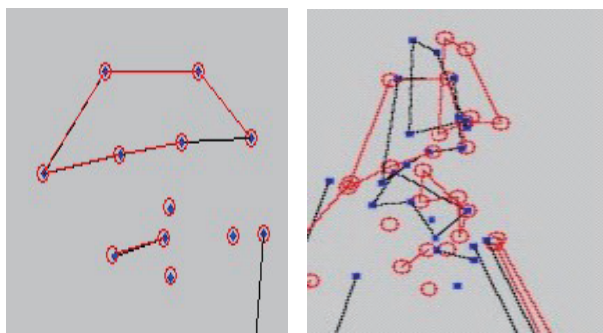


图3 基于生成对抗网络的识别算法识别结果图

在此基础上,为了进一步验证本文基于生成对抗网络的识别算法在实际应用中与现有识别算法相比具备的优势,针对上述五组不同的姿态变化条件,依据下述公式,计算得出三组识别算法的识别率:

$$T = \frac{P - F}{P} \times 100\%$$
 (8)

式中: T 表示识别算法的识别率; P 表示识别总数; F 表示错误识别的数量。根据上述公式,计算得出不同实验条件下的识别算法识别正确率,为了方便比较,将结果记录在表2。

表2 三组识别算法识别率记录表

序号	姿态变化	实验组识别率 /%	对照 A 组识别率 /%	对照 B 组识别率 /%
(1)	0°	98.62	82.54	83.25
(2)	-30° ~ 30°	98.26	74.26	74.68
(3)	-45° ~ 45°	98.03	70.24	70.35
(4)	-60° ~ 60°	97.65	68.25	68.56
(5)	-90° ~ 90°	97.54	65.41	65.85

从表2中记录的实验结果可以看出,随着人脸姿态变化范围的不断增加,三组识别算法的识别率均呈现出逐渐递减的变化趋势,但实验组识别算法的识别率始终高于另外两组,并且实验组识别算法识别率的递减幅度也最低,在姿态变化范围在-90°~90°时,其识别正确率仍然达到了97.00%以上,而对照A组和对照B组的识别率最高仅控制在82.00%~83.50%范围内。由此可以看出,实验组识别算法的识别正确率更高,识别能力更强,更适用于应用到实际情况中实现人脸在多姿态环境下的识别。

5 结语

本文上述研究从多姿态人脸识别问题角度出发,结合生成对抗网络,提出了一种全新的人脸识别算法。通过对比实验证明,新的识别算法在实际应用中可以实现对不同姿态变化下的高准确率识别。由于研究时间有限,新的识别算法当前存在网络训练时间长、训练所需样本数量多等问题。因此,

在后续的研究中,为了能够进一步提高该识别算法的应用适用性,还将在确保识别准确率的前提条件下,针对如何进一步缩短网络训练时间、减少训练样本数量,开展更深入的探索研究。

参考文献:

[1] 陈侃松,郑园,许立君,等.基于 ThetaMEX 全局池化的人脸识别神经网络: ShuffaceNet[J]. 计算机应用, 2023, 43(8): 2572-2580.

[2] 古振猛,孙士日.基于深度学习的人脸识别技术在烟草商业内部监管的应用[J]. 信息与电脑(理论版), 2023, 35(6): 181-183.

[3] 李云红,刘杏瑞,谢蓉蓉,等.基于超分辨重建和公共特征子空间的低分辨率人脸识别[J]. 西北大学学报(自然科学版), 2023, 53(2):241-247.

[4] 王大彬,栗兰林,李茂圣,等.采用人脸识别支付的地铁自动售检票系统避免重复计费的方案研究[J]. 城市轨道交通研究, 2023,26(7):87-91+98.

[5] 李畅.基于特征融合的人脸识别技术在城市轨道交通安检系统中的应用[J]. 交通世界, 2023(19):10-12.

[6] 罗傲,贾萍萍,张增一.国外公众对人脸识别技术的认知与态度:基于 Twitter 相关话题讨论的质性分析[J]. 北京理工大学学报(社会科学版), 2023,25(4):192-200.

[7] 耿雅玲,袁申,徐向阳.人脸识别技术在高校智慧图书馆建设中的应用研究[J]. 电脑知识与技术, 2023,19(17):30-32+35.

[8] 江柏,李含雁,李勇滔,等.基于人脸识别的工程机械作业人员疲劳检测系统研究[J]. 装备制造技术, 2023(5):81-83+108.

[9] 桑高丽,肖述笛,赵启军.联合软阈值去噪和视频数据融合的低质量3维人脸识别[J]. 中国图象图形学报, 2023, 28(5):1434-1444.

[10] 徐浙君.基于改进萤火虫算法的卷积神经网络在人脸识别中的研究[J]. 科技通报, 2023,39(4):32-37+50.

[11] 杨彦青,宋星.不同光照强度变化下融合深度学习的人脸识别技术研究[J]. 自动化与仪器仪表, 2023(4):49-52+57.

【作者简介】

蒋文豪(1982—),女,四川大竹人,硕士,副教授,研究方向:大数据和大数据安全。

(收稿日期:2023-11-10)