

Research on the Robustness of Face Recognition Algorithm to Different Lighting Technologies^{*}

SUN Jinlong, WU Zhenning, XIAO Zhongzhe, HUANG Min^{*}

(School of Optoelectronic Science and Engineering, Soochow University, Suzhou Jiangsu 215006, China)

Abstract: The existing face recognition algorithms, such as MTCNN algorithm and RetinaFace algorithm, have been able to achieve quite high recognition rate for the face images collected under normal lighting technology. However, in some special applications, there is no exact experimental result that can give us a conclusion on whether the existing face recognition algorithms have good robustness and maintain a high recognition rate for the face images collected under special lighting technology. We collected more than 6 000 face images obtained under different lighting technologies as our hybrid face database, and choose four effective face recognition networks designed by using LBPH algorithm, convolutional neural network (CNN), MTCNN algorithm and RetinaFace algorithm, which are tested on the open source WIDER FACE database and our hybrid face database respectively. Finally, it is found that RetinaFace algorithm is robust to face images obtained through different lighting technologies. We further use the deep learning labeling tool to label the face images misjudged by RetinaFace algorithm, and send the labeled images to RetinaFace face recognition network for retraining. The optimized RetinaFace face detection model achieves 98.6% recognition accuracy, which further improves the robustness of RetinaFace algorithm to different lighting conditions.

Key words: LBPH algorithm; convolutional neural network; MTCNN algorithm; RetinaFace algorithm; facial identification; robustness

EEACC: 6135E

doi: 10.3969/j.issn.1005-9490.2022.05.018

关于人脸识别算法对不同照明技术的鲁棒性研究^{*}

孙金龙, 吴振宁, 肖仲喆, 黄 敏^{*}

(苏州大学光电科学与工程学院, 江苏 苏州 215006)

摘 要: 对于在正常照明技术下采集得到的人脸图像, 现有的人脸识别算法如 MTCNN 算法、RetinaFace 算法, 已经能够取得相当高的人脸识别率。然而在某些特殊应用中, 对于在特殊照明技术下采集得到的人脸图像, 现有的人脸识别算法是否具有很好的鲁棒性, 保持较高的人脸识别率, 并没有确切的实验结果能够给出我们结论。本文收集了 6 000 多张在不同照明技术下得到的人脸图像作为我们的混合人脸数据库, 并利用 LBPH 算法、卷积神经网络 (CNN)、MTCNN 算法、RetinaFace 算法设计出了四种有效的人脸识别网络, 分别对开源的 WIDER FACE 人脸数据库和我们的混合人脸数据库进行了测试。最终发现 RetinaFace 算法对于不同照明技术得到的人脸图像具有较好的鲁棒性。我们进一步利用深度学习标注工具对 RetinaFace 算法误判的人脸图像进行了标注, 并将标注后的图像送入到 RetinaFace 人脸识别网络中重新训练, 优化后的 RetinaFace 人脸检测模型得到了 98.6% 的人脸识别准确率, 使得 RetinaFace 算法对不同光照条件的鲁棒性取得了进一步的提升。

关键词: LBPH 算法; 卷积神经网络; MTCNN 算法; RetinaFace 算法; 人脸识别; 鲁棒性

中图分类号: TN957.52

文献标识码: A

文章编号: 1005-9490(2022)05-1123-06

目前广泛应用的几种人脸识别算法对于 WIDER FACE 人脸数据库都能够获得较高的人脸识别率。然而 WIDER FACE 人脸数据库中的人脸

图像大部分都是在日光照明条件下获得的。在某些特殊应用如皮肤检测中, 为了对面部深层次的肌肤问题有一个更好的判别, 我们往往需要采用特殊的

项目来源: 国家自然科学基金项目 (61906128); 江苏省自然科学基金项目 (BK20180834)

收稿日期: 2021-06-21 修改日期: 2021-07-27

光照技术对人脸进行拍摄并识别。现有的人脸识别算法对于混合光照技术得到的人脸图像是否具有好的鲁棒性是本文重点研究的问题。

我们收集了 6 000 多张采用不同光照技术得到的人脸图像作为我们的混合人脸数据库。图 1 分别为采用日光、交叉偏振光、平行偏振光、伍氏光、紫外光照明技术得到的同一人的五张人脸图像。

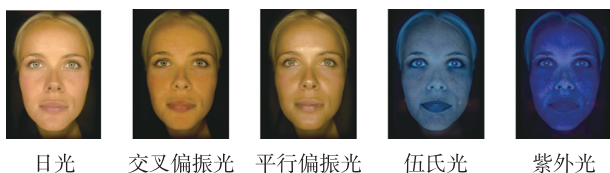


图 1 采用不同照明技术得到的人脸图像

本文介绍了几种常用的人脸识别算法的基本原理,并在我们自己的工作站上利用这些算法搭建了人脸识别网络,然后对 WIDER FACE 人脸数据库和我们的混合人脸数据库进行了人脸识别准确度测试。利用卷积神经网络(CNN)、MTCNN 算法、RetinaFace 算法设计的人脸识别网络,在前一数据库上取得了 84.3%,98.7%,99.4%的人脸识别准确率,而在混合数据库上仅取得了 72.7%,64.5%,95.8%的人脸识别准确率。显然 RetinaFace 算法具有较好的鲁棒性。

最后我们对搭建的 RetinaFace 人脸识别网络进行了优化,优化后的 RetinaFace 人脸识别网络在混合数据库上达到了 98.6%的人脸识别准确率,进一步提高了 RetinaFace 算法的鲁棒性。

1 人脸识别网络算法与原理

1.1 LBPH 人脸识别网络基本原理

LBPH 算法^[1]首先对彩色图像进行灰度化处理,使得图像中的每个像素点都用灰度值来表示。然后以每个像素点为中心,使用双线性差值法求得两个相邻像素点的灰度值大小关系,并对其进行二进制编码,从而得到原始图像的 LBP 编码图像。

将得到的 LBP 编码图像分割成多个方格区域,通过方格区域中每个像素点的 LBP 和 ULBP 的值获取该区域的 LBPH 图。把每个区域的 LBPH 图连接起来,继而得到整幅图像的 LBP 编码直方图,如图 2 所示。从图中可以看出 LBPH 算法具有很好的鲁棒性。

LBPH 人脸识别网络通过将被检测对象的 LBPH 图与构建好的 LBP 编码直方图进行比较,通过二者的契合度来判断被检测图像是否为人脸,达到人脸识别的目的。

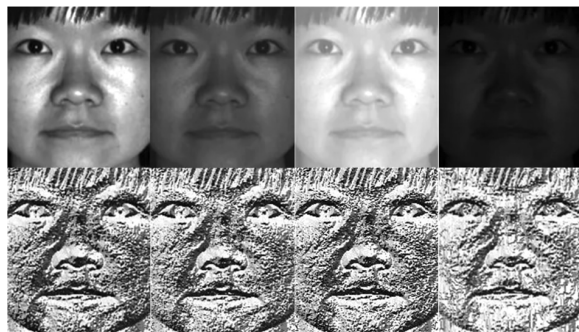


图 2 原始图像以及对应的 LBP 图像

1.2 CNN 人脸识别网络基本原理

1.2.1 卷积操作

卷积神经网络(Convolutional Neural Network, CNN)^[2]是采用卷积操作进行特征提取的神经网络。我们为 CNN 卷积神经网络设定一个初始滑动窗口,又称内核,过滤器,或者特征检测器。CNN 将过滤器的元素与数据库中人脸图像的像素矩阵相乘,然后相加的过程被称为卷积。

如图 3 所示,左侧 5×5 矩阵代表一张人脸的灰度图像,每个条目对应一个像素,0 代表黑色,1 代表白色(灰度图像通常在 0~255 之间),中间 3×3 矩阵代表初始滑动窗口。为了得到人脸图像完整的卷积,CNN 卷积神经网络在整个矩阵上滑动过滤器来对每个元素进行卷积,最终得到右边人脸图像的卷积特征。

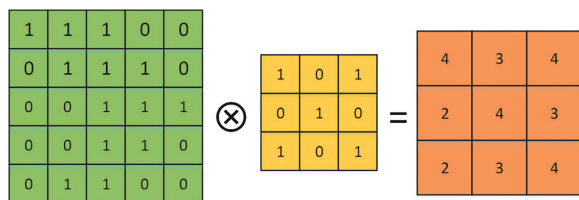


图 3 利用卷积对图像进行特征提取

1.2.2 利用 CNN 进行人脸识别的原理

利用卷积神经网络^[3]识别人脸图像,主要包括四个步骤:卷积层初步提取特征、池化层提取主要特征、全连接层汇总各部分特征、生成人脸识别分类器,从而对输入分类器的人脸图像进行预测识别。

卷积层:对于输入的 40×40 的人脸图像,每个像素点都存储着人脸图像的信息。我们首先随机初始化一个卷积核,再从人脸图像中提取一定特征(数字矩阵)。数字矩阵与卷积核进行卷积操作,得到卷积层输出结果。然后 CNN 卷积神经网络通过不断地改变卷积核,初步确定表征人脸特征的有用的卷积核是哪些,再得到人脸图像数字矩阵与相应卷积核进行卷积操作的卷积层输出结果。

池化层:在尽可能保留人脸图像空间信息的前提下,减少训练参数数量,降低卷积层输出的特征向

量维度,增大卷积核感受野,提取高层特征,同时减少过拟合^[4]现象,只保留有用图片信息,减少噪声传递。分为最大池化和平均池化,在我们的 CNN 人脸识别网络中选用了最大池化。

全连接层:将池化层输出的张量重新切割成一些向量,乘以权重矩阵,加上偏置值,然后对其使用 ReLU 激活函数^[5],用梯度下降法优化参数,最终生成我们所需要的分类器。

如图 4 所示,参考 LeNet5 模型^[6],我们自己设计的 CNN 网络利用两层“卷积层+池化层”来提取人脸特征,通过一个全连接层生成一个人脸识别的二分类器(输出层),最终实现了人脸识别的功能。

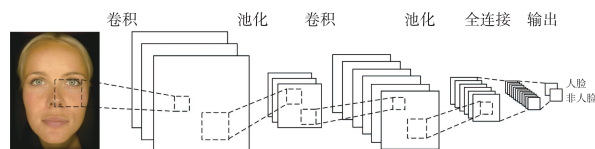


图 4 人脸识别卷积神经网络结构

1.3 MTCNN 人脸识别网络基本原理

MTCNN 算法^[7-8]是一种基于深度学习的人脸检测和人脸对齐算法,能够同时完成人脸检测和人脸对齐的任务,相较于传统的算法,MTCNN 算法性能更好,检测速度更快。

如图 5 所示,MTCNN 算法包括三级多任务深度卷积网络:

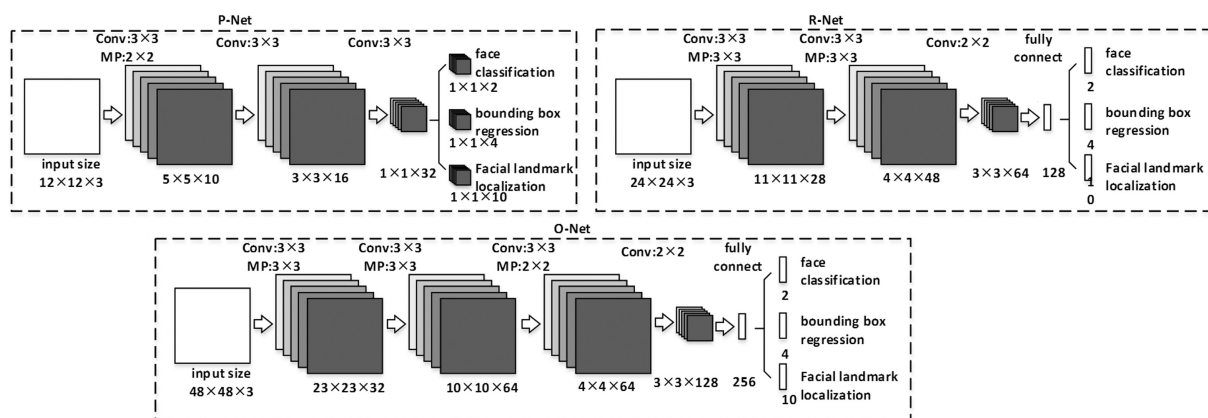


图 5 三级多任务深度卷积网络体系结构

MTCNN 算法首先对图片进行 Resize 操作,将原始图像缩放成不同的尺度,生成图像金字塔。然后将不同尺度的图像送入到这三个子网络中进行训练,目的是为了可以检测到不同大小的人脸,从而实现多尺度目标检测。

P-Net 网络:P-Net 网络是一个人脸区域的候选网络,P-Net 网络的输入是一个 $12 \times 12 \times 3$ 的图像,通过 3 层的卷积之后,判断这个 12×12 的图像中是否存在人脸,并且给出人脸框的回归和人脸关键点。

P-Net 网络的第一部分输出是用来判断输入图像中是否存在人脸,输出向量大小为 $1 \times 1 \times 2$,也就是两个值。

P-Net 网络的第二部分给出框的精确位置,一般称为框回归。输入 P-Net 网络的 12×12 的图像可能并不是完美的人脸框的位置,有的时候人脸并不正好为方形,有可能 12×12 的图像偏左或偏右,因此需要输出当前框位置相对完美的人脸框位置的偏移。这个偏移大小为 $1 \times 1 \times 4$,即表示框左上角的横坐标的相对偏移,框左上角的纵坐标的相对偏移、框的宽度误差、框的高度误差。

P-Net 网络的第三部分给出人脸的 5 个关键点

的位置。5 个关键点分别对应着左眼位置、右眼位置、鼻子位置、嘴巴边缘位置、嘴巴右边缘位置。每个关键点需要用两维向量来表示,因此输出向量的大小为 $1 \times 1 \times 10$ 。

最后,P-Net 网络采用 NMS^[9]来合并重叠率高的候选框,以提高 MTCNN 网络整体的运行速度。

R-Net 网络:R-Net 的网络结构和 P-Net 的网络结构相比多了一个全连接层,可以更好地取得抑制 false-positive 的作用。图像在输入 R-Net 网络之前,都需要缩放到 $24 \times 24 \times 3$ 的大小,R-Net 网络的输出与 P-Net 网络是相同的,R-Net 网络的目的是去除大量的非人脸框。

O-Net 网络:O-Net 网络比 R-Net 网络又多了一层卷积层,可以使得 O-Net 网络处理的结果更加精细。输入 O-Net 网络的图像大小为 $48 \times 48 \times 3$,O-Net 网络的输出包括 N 个边界框的坐标信息,score 以及关键点位置。

从 P-Net 网络到 R-Net 网络,再到最后的 O-Net 网络,网络输入的图像越来越大,卷积层的通道数越来越多,网络的深度(层数)也越来越深,因此识别人脸的准确率应该也是越来越高。

1.4 RetinaFace 人脸识别网络基本原理

RetinaFace 算法^[10-11]是一种鲁棒的单级人脸检测器。RetinaFace 算法采用轻量级的 backbone 网络,使得其能在单个 CPU 上实时运行 VGA 分辨率的图像。与此同时 RetinaFace 算法进一步添加自监督网络解码器分支,与已有的监督分支并行预测像

素级的 3D 形状人脸信息,使其对各种尺度条件下的人脸都可以做到像素级别的定位。

如图 6 所示,RetinaFace 算法主要包括四个模块,分别是骨干网络,特征金字塔,上下文模块以及多任务损失。

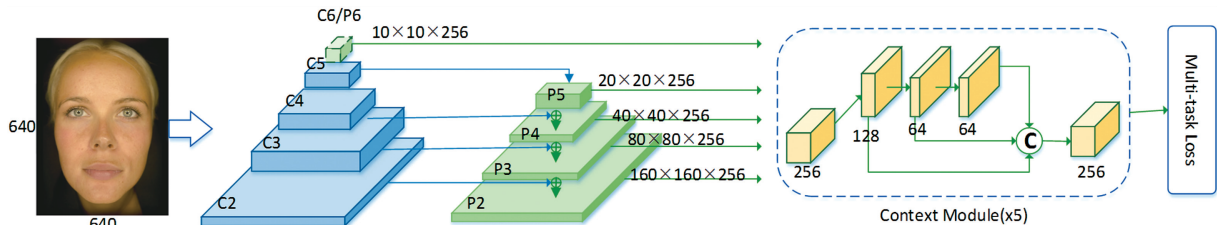


图 6 RetinaFace 人脸识别网络体系结构

骨干网络:RetinaFace 算法主要采用了两种骨干网络,分别是 Resnet50 和 mobilenetV1-0.25^[12]。通过采用轻量级骨干网络,使得 RetinaFace 算法可以在单个 CPU 内核上实时运行,以实现 VGA 分辨率的显示。且 mobilenetV1-0.25 利用了深度可分离卷积,具有很强的实时性。

特征金字塔:RetinaFace 算法采用从 P2 到 P6 的特征金字塔层,其中 P2 到 P5 通过使用自顶向下和横向连接计算相应的 ResNet 残差阶段(C2 到 C5)的输出。P6 是在 C5 处通过一个步长 2 的 3×3 卷积计算得到的。C1-C5 是在 ImageNet-11k 数据集上预先训练好的 ResNet-152^[13] 分类网络,P6 是用“Xavier”方法^[14] 随机初始化的。RetinaFace 算法通过特征金字塔层对人脸特征进行融合。

上下文模块^[15]:RetinaFace 算法将独立上下文模块应用于 5 个特征金字塔上,用于提高感受野和加强严格上下文建模能力。其结构是将输入的 feature map 分为三个部分,分别进行不同深度的特征提取,每一个部分的最后一次卷积操作不带激活函数,最后将三个部分联结,再通过激活函数。上下文模块使得人脸特征的融合进一步加强。

多任务损失:对于任何训练的 anchor i ,RetinaFace 的目标是最小化下面的多任务的 Loss:

$$L = L_{cls}(p_i, p_i^*) + \lambda_1 p_i^* L_{box}(t_i, t_i^*) + \lambda_2 p_i^* L_{pts}(l_i, l_i^*) + \lambda_3 p_i^* L_{pixel}$$

式中: p_i 表示 anchor i 预测为人脸的概率, p_i^* 表示真值,正样本 anchor 为 1,负样本 anchor 为 0, t_i, t_i^* 分别表示与正样本 anchor 对应的预测框和真实标注框的位置坐标。 l_i, l_i^* 分别表示正样本 anchor 五个人脸关键点的预测坐标和真实坐标。 L_{pixel} 表示渲染得到的人脸与原始人脸的像素差异。

主要包括四个部分:

①人脸分类 Loss:获取所有种类的预测结果的交叉熵 Loss。

②人脸框回归 Loss:获取所有正标签框预测结果的回归 Loss。

③人脸的 landmark 回归 Loss:获取所有正标签人脸关键点预测结果的回归 Loss。

④自监督 3D 人脸回归 Loss:Loss 调节参数 $\lambda_1 \sim \lambda_3$ 设置为 0.25, 0.1 和 0.01,这意味着在监督信号中,RetinaFace 算法增加了边界框和关键点定位的重要性。

2 实验基本条件和实验测试结果

2.1 实验基本条件

2.2.1 实验硬件平台

我们在自己的工作站上分别搭建了 TensorFlow 和 pytorch 的深度学习框架,作为利用不同算法设计的人脸识别网络的 Python 工作环境。并在我们的工作站上完成了四种人脸识别网络对不同的人脸图像数据库的训练和测试。我们的工作站配置如下:

- Intel(R) Xeon(R) Gold 6128 CPU.
- 512GB DDR4 memory.
- NVIDIA Geforce RTX 2080Ti GPU.

2.2.2 人脸数据集合理性说明

在我们的实验中,我们利用 WIDER FACE 人脸数据集和我们自己的混合人脸数据集对不同的人脸识别网络分别进行了训练和测试。

WIDER FACE 人脸数据集作为人脸识别领域经典的开源数据集,其所包含的人脸图像在尺度、姿态、表情、遮挡方面都有很大的变化范围,使用它对网络进行训练和测试能够很好地表征各种人脸识别

网络在正常光照条件下的基本性能差异。

我们自己的混合人脸数据集包含了各种年龄、各种角度的男女比例均衡的人脸图像,同时确保对于每一张在日光(正常光)照明条件下得到的人脸图像,都有与之对应的采用交叉偏振光、平行偏振光、伍氏光、紫外光照明技术得到的四张人脸图像,保证了同一组人脸图像只存在照明条件的不同,排除其他外在因素的干扰。在人脸识别算法对不同照明条件的鲁棒性研究中,确保了实验结果的准确性和可靠性。

2.2 实验测试结果对比

2.2.1 正常光照条件下的基本性能测试

我们利用工作站分别对搭建好的四种人脸识别网络进行了测试。其中虽然 LBPH 人脸识别网络对人脸图像达到了极高的识别准确率,且使用 LBPH 算法进行人脸识别不会受到光照、缩放、旋转和平移的影响,但是通过 LBPH 算法设计的人脸识别网络需要先对人脸图像进行训练,然后通过对比判断被检测图像是否与已经训练的数据相匹配,从而判断被检测图像是否为人脸。因此采用 LBPH 算法设计的人脸识别网络具有很大的局限性。

对于使用 CNN、MTCNN、RetinaFace 算法设计的人脸识别网络,我们利用 WIDER FACE 人脸数据库中的 WIDER FACE train 数据集对网络进行了训练,并利用训练好的网络模型对 WIDER FACE test 数据集进行了测试,测试得到的 TP (True Positives, 真正:预测为正,实际为正)、TN (True Negatives, 真负:预测为负,实际为负)、FP (False Positives, 假正:预测为正,实际为负)、FN (False Negatives, 假负:预测为负,实际为正)如表 1 所示。

表 1 不同人脸识别网络基本性能测试结果

人脸识别网络	TP	TN	FP	FN
CNN 网络	84.5%	95.33%	4.67%	15.5%
MTCNN 算法	98.4%	99.83%	0.17%	1.6%
RetinaFace 算法	99.3%	99.89%	0.11%	0.7%

通过表 1 我们可以看出,对于正常光照条件下的人脸图像,RetinaFace 算法的网络性能是最好的,MTCNN 算法的网络性能略次之,我们自己设计的 CNN 卷积神经网络的性能较差。

2.2.2 不同光照条件下的性能对比测试

为了更好地反映不同算法设计的人脸识别网络对不同光照技术的鲁棒性,我们使用训练好的网络模型参数分别对 WIDER FACE test 数据集(部分)

和我们的混合人脸数据集的识别准确率进行了对比测试,测试结果如表 2 所示。

表 2 不同光照条件下的性能对比测试结果

人脸识别网络	识别准确度/%	
	WIDER FACE test 数据集	混合人脸数据集
CNN 卷积神经网络	84.3	72.7
MTCNN 算法	98.7	64.5
RetinaFace 算法	99.4	95.8

通过表 2 可以看出,RetinaFace 人脸识别网络在对不同光照条件下的人脸图像识别时仍能保持较高的准确率;而 MTCNN 人脸识别网络和我们自己设计的 CNN 人脸识别网络的识别准确率受到了较大的影响,出现了一定程度的下降,其中 MTCNN 人脸识别网络受照明条件影响较大。

2.2.3 不同人脸识别网络增强训练结果对比

为了进一步反映三种人脸识别网络在对不同照明技术下获取的人脸图像进行识别时的抗干扰能力,我们对网络检测错误的人脸图像手动标注了五个面部标志,并将其与原来的监督信号一起重新送入网络中进行网络增强训练。我们利用增强训练后的模型参数对混合人脸数据库进行了进一步的测试,得到的三种人脸识别网络的识别准确率如表 3 所示。

表 3 不同人脸识别网络增强训练结果对比

人脸识别网络	识别准确度/%	
	增强训练前	增强训练后
CNN 卷积神经网络	72.7	78.4
MTCNN 算法	64.5	63.7
RetinaFace 算法	95.8	98.6

通过表 3 我们可以看出,在进行网络增强训练后,RetinaFace 人脸识别网络的准确率得到了一定提升,其性能得到了进一步优化;我们自己设计的 CNN 人脸识别网络也得到了一定程度的提升;而 MTCNN 人脸识别网络由于对光照不敏感,需要数据量大,人脸识别准确率反而有所降低。

3 测试结果分析

对测试结果进行分析发现,相较于其他人脸识别算法,RetinaFace 算法不仅对正常光照下的人脸图像具有优良的人脸识别性能,同时对采用不同照明技术得到的人脸图像进行识别时同样具有优良的鲁棒性。在进行网络增强训练后,RetinaFace 算法网络模型的性能仍然可以得到进一步优化。Reti-

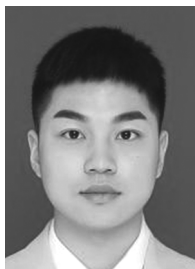
naFace 算法作为一个强大的单阶段人脸检测器,利用联合监督和自我监督的多任务学习方式,实现了在各种人脸尺度上执行像素方面的人脸定位。同时 RetinaFace 算法采用了上下文模块等结构,使得训练得到的网络模型更能适应光照条件的变化。因此 RetinaFace 算法对不同照明条件下的人脸图像具备优良的鲁棒性。

而 MTCNN 算法虽然在正常光照条件下具有良好的识别性能,然而不同的照明条件对 MTCNN 算法的识别性能造成了极大的影响,同时在进行网络增强训练之后,MTCNN 算法对不同照明条件的人脸图像的识别性能也并未得到有效提升。通过进一步对检测出错的人脸图像进行分析,我们发现大部分检测错误的人脸图像都是由伍氏光和紫外光照明技术拍摄得到的。因此 MTCNN 算法对于不同照明技术的人脸图像并不具有鲁棒性。

在我们设计的 CNN 人脸识别网络中,由于我们仅仅采用了两层“卷积层+池化层”对人脸特征进行提取,简单的网络结构无法有效地提取出人脸图像的大部分特征,人脸识别性能并不显著,同时也会受到光照条件的影响。在进行网络增强训练后,对不同光照条件人脸图像的识别性能也取得了一定程度的提高。

4 结束语

为了分析不同照明技术对常用人脸识别算法的影响,从而定性描述常用的人脸识别网络的鲁棒性。本文对几种常用的人脸识别网络对在不同照明技术下获取的人脸图像的识别准确率进行了对比测试,最终发现相较于其他人脸识别算法,RetinaFace 算法本身具有较高的人脸识别性能和对不同照明技术的抗干扰能力。同时我们对 RetinaFace 算法进行了网络增强训练,使其性能得到了进一步的优化,在混合人脸数据库上的识别准确率最终也达到了 98.6%,进一步提高了 RetinaFace 算法对不同照明条件的鲁棒性,使得其能够进行更多的实际应用。



孙金龙(2000—),男,硕士研究生,研究方向为图像及声音信号处理,20224239033@stu.suda.edu.cn,1823403022@stu.suda.edu.cn;

参考文献:

- [1] 朱峰. 一种基于多尺度 LBPH 特征的快速人脸识别算法[J]. 计算机应用与软件, 2015, 8(32): 315-319.
- [2] 冯友兵, 陆轶秋, 仲伟波. 基于 CNN 和 SVM 的人脸识别系统的设计与实现[J]. 计算机与数字工程, 2021, 49(2): 378-382, 420.
- [3] 柯岩, 林小竹, 廖蕊, 等. 卷积神经网络层级分解研究[J]. 计算机工程, 2019, 45(11): 191-197.
- [4] 李俭川, 秦国军, 温熙森, 等. 神经网络学习算法的过拟合问题及解决方法[J]. 振动、测试与诊断, 2002, 22(4): 260-264.
- [5] 王双印, 滕国文. 卷积神经网络中 ReLU 激活函数优化设计[J]. 信息通信, 2018(1): 42-43.
- [6] Yann L C, Bottou L, Bengio Y, et al. Gradient-Based Learning Applied to Document Recognition[J]. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 1998, 86(11): 2278-2324.
- [7] Zhang K P, Zhang Z P, Li Z F, et al. Joint Face Detection and Alignment using Multi-Task Cascaded Convolutional Networks[J]. IEEE Signal Processing Letters, 2016, 23(10): 1499-1503.
- [8] 蓝雯飞, 张盛兰, 朱容波, 等. 基于改进 MTCNN 的人脸检测算法[J]. 中南民族大学学报(自然科学版), 2020, 39(6): 637-641.
- [9] Salscheider N O. FeatureNMS: Non-Maximum Suppression by Learning Feature Embeddings[C]//2020 25th International Conference on Pattern Recognition (ICPR), Milan, Italy, 2021: 7848-7854.
- [10] Deng J K, Guo J, Zhou Y X, et al. RetinaFace: Single-Stage Dense Face Localisation in the Wild[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, Seattle, WA, USA, 2020: 5203-5212.
- [11] 随玉腾, 阎志远, 戴琳琳, 等. 基于 RetinaFace 的人脸多属性检测算法研究[J]. 铁路计算机应用, 2021, 30(3): 1-4.
- [12] Cai K W, Miao X Y, Wang W, et al. A Modified YOLOv3 Model for Fish Detection Based on MobileNetV1 as Backbone[J]. Aquacultural Engineering, 2020, 91, 102117.
- [13] He K M, Zhang X Y, Ren S Q, et al. Deep Residual Learning for Image Recognition[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Las Vegas, NV, USA, 2016: 770-778.
- [14] Glorot X, Bengio Y. Understanding the Difficulty of Training Deep Feedforward Neural Networks[C]//Proceedings of the Thirteenth International Conference on Artificial Intelligence and Statistics, PMLR, Sardinia, Italy, 2010, 9: 249-256.
- [15] 郑晨斌, 张勇, 胡杭, 等. 目标检测强化上下文模型[J]. 浙江大学学报(工学版), 2020, 54(3): 529-539.



黄敏(1986—),男,博士,副教授,研究方向为计算机体系结构,智能系统,自动化测试,医学图像处理及识别,声音信号处理,声学目标探测及识别, hmin@suda.edu.cn。