

DOI:10.19651/j.cnki.emt.2210776

# 基于轻量型网络的口罩遮挡人脸识别方法\*

移 浩<sup>1</sup> 侯 劲<sup>1,2</sup> 石浩德<sup>1</sup>

(1. 四川轻化工大学自动化与信息工程学院 宜宾 644000; 2. 四川轻化工大学人工智能四川省重点实验室 宜宾 644000)

**摘要:** 由于口罩的遮挡会大幅降低人脸可供识别的特征,使得之前提出的人脸识别算法在现有外部环境下的识别性能大幅下降。因此,针对现有人脸识别技术在当前应用场景中的不足,本研究采用 MobileNet v2 轻量级卷积神经网络替换 InceptionResNet-v1 网络作为骨干网络对 FaceNet 人脸识别方法进行了改进,在简化模型参数的同时提高了模型的运算速度,并且在 MobileNet V2 网络中引入一种轻量型的混合注意力模块,同时将 Softmax Loss 与 Triplet Loss 加权融合作为网络模型的联合损失函数,通过调整权值达到最优后作为损失函数进行训练,提高网络的识别准确率。实验结果表明:本研究所提出的人脸识别网络在进行口罩遮挡人脸识别时,识别准确率达到 92.1%,较原有人脸识别网络有大幅提升,同时识别速度也明显优于原有网络。

**关键词:** 人脸识别; MobileNet v2; 注意力机制; 损失函数

**中图分类号:** TP391 **文献标识码:** A **国家标准学科分类代码:** 510.4050

## Mask occlusion face recognition method based on lightweight network

Yi Jie<sup>1</sup> Hou Jin<sup>1,2</sup> Shi Haode<sup>1</sup>

(1. School of Automation and Information Engineering, Sichuan University of Science and Engineering, Yibin 644000, China;

2. Artificial Intelligence Key Laboratory of Sichuan Province, Sichuan University of Science and Engineering, Yibin 644000, China)

**Abstract:** As the mask will greatly reduce the features available for face recognition, the recognition performance of the previously proposed face recognition algorithm will be greatly reduced in the existing external environment. Therefore, given the shortcomings of the existing face recognition technology in the current application scenarios, this study uses MobileNet v2 lightweight convolutional neural network to replace the Inceptionresnet-V1 network as the backbone network to improve the FaceNet face recognition method, which simplifies the model parameters and improves the operation speed of the model. In addition, a lightweight mixed attention module is introduced into the Mobilenet v2 network, and the weighted fusion of Softmax Loss and Triplet Loss is used as the joint Loss function of the network model, which is trained as the Loss function after the adjustment of weight reaches the optimal value to improve the recognition accuracy of the network. The experimental results show that the face recognition network proposed in this study achieves 92.1% recognition accuracy in face mask masking, which is significantly improved compared with the original face recognition network, and the recognition speed is also significantly better than the original network.

**Keywords:** face recognition; MobileNet v2; attention mechanism; loss function

## 0 引 言

目前,现有的各类人脸识别技术已逐渐趋于成熟,可以在多数应用场景中取得较高的识别精度。然而,使用现有面部识别算法进行某些特定条件下的面部识别时,由于使用者面部遮挡、表情变化以及光线干扰等因素的影响,需要与设备充分配合来完成面部识别操作,限制了现有识别算法在这类识别环境下的识别效率,也影响了算法的识别准

确度。同时,在新型冠状病毒肺炎疫情期期间,口罩的普遍佩戴也使现有人脸识别系统趋于失效,摘下口罩进行识别或改由人工进行识别会产生疫情传播风险,增加抗疫成本。因此,针对面部口罩遮挡等外部不利因素对现有面部识别算法进行改进研究既可以降低疫情防控工作的人力投入,改善防疫机制,对于研究其他类型遮挡下的人脸识别也有一定的参考价值<sup>[1]</sup>。

在现阶段应用背景下,使用传统人脸识别技术提取人

收稿日期:2022-07-22

\* 基金项目:四川省科技厅项目(2021YFG0055)、四川省人工智能重点实验室项目(2021RYY04)资助

脸图像的深层特征,已经不能取得很好的提取效果。而随着计算机硬件设备条件以及相关技术的发展,深度学习技术<sup>[2]</sup>在目标检测、模式识别和图像生成等领域应用越来越广泛,快速推动了基于深度学习的人脸识别技术的发展。使用基于深度学习进行人脸识别,可以获得更深层次的人脸图像特征,而且在深度模型上性能优越,已经成为当前的主流方法。

Google 于 2015 推出的 FaceNet<sup>[3]</sup>是一个被广泛应用于人脸识别、检测和聚类的系统,该系统利用神经网络对映射到欧式空间的人脸图像所获得的欧氏距离与阈值之间的大小进行对比从而判断是否为同一张人脸<sup>[4]</sup>。同时,FaceNet 创新的提出了使用三元损失函数替代 Softmax Loss,将 LFW 数据集中的面部识别率提高到了 99% 以上<sup>[5]</sup>,但由于模型的参数量大,同时三元组损失函数存在一些运算复杂度较高的问题,使模型的训练较为困难且对计算机的硬件资源要求较高。2016 年 Liu 等<sup>[6]</sup>在 Softmax Loss 中引入角度距离,提出了一种 L-Softmax 人脸损失函数,通过将特征之间的直线距离问题转化为角度问题,从而进一步减少类内距离,增大类间距离。2022 年,谷凤伟等<sup>[7]</sup>基于 FaceNet 网络提出了皮尔森判别网络的人脸识别方法,通过采用皮尔森相关系数判别模块替换 FaceNet 网络中的欧氏距离判别模块来实现人脸的深度判别特征,对非限制场景下的人脸目标具有良好的识别效果。

深度学习的一般趋势是使用更加复杂和深层的网络来获得更高的准确率<sup>[8]</sup>。但随着网络深度的加深,其分布会逐渐发生偏移,导致梯度消失进而使模型的收敛速度减缓,训练速度降低,模型准确度下降,识别时间增加<sup>[9]</sup>。为了解决上述缺陷,参数较少的模型正在受到越来越多的关注<sup>[10]</sup>,Howard 等<sup>[11]</sup>在 2017 年提出了轻量级神经网络 MobileNet,提出了深度可分离卷积层,在保证精度高的情况下,大大减少计算量提高了运算效率。Sandler 等<sup>[12]</sup>2019 提出的 MobileNet V2,在继续使用深度可分离卷积的基础上,提出了一种使用倒置残差和线性瓶颈的模块。2018 年 Zhang 等<sup>[13]</sup>针对计算能力有限的移动设备,使用分组卷积和通道洗牌操作,在保证模型精度、减少计算量的情况下构建了 ShuffleNet V1 网络。

基于对当前深度学习技术的研究,针对现有人脸识别方法存在的不足,本文提出了一种基于改进 FaceNet 人脸识别网络模型的口罩遮挡下的人脸识别方法,本研究首先使用 MobileNet v2 轻量级卷积神经网络替换 InceptionResNet-v1 作为 FaceNet 模型的骨干网络,加快模型的运算速度,同时在 MobileNet v2 网络中引入轻量级混合卷积注意力机制,优化网络的特征表达能力,提高模型的识别准确率;其次在原有的人脸识别模型 FaceNet 的损失函数 Triplet Loss 的基础上,采用 Softmax Loss 和 Triplet Loss 加权的方法作为其损失函数,从而进一步改善人脸识别的性能。

## 1 相关研究

### 1.1 模型整体框架

口罩遮挡条件下人脸识别的研究可以划分为两个阶段:人脸检测与人脸识别。具体流程如图 1 所示。第 1 步,人脸图像的获取,对口罩下的人脸图像进行采集,然后进行图像分类与标签处理,保存并建立数据库。第 2 步,将图像照片或者视频中的人脸进行检测,并对检测出的人脸进行框选标记,对其预处理,然后进行配准和归一化,以便于之后的人脸识别;第 3,人脸识别对比,比较在前面步骤中获得的人脸图像与对比库中的图像,利用网络模型将二者进行特征提取,再使用损失函数求差,从而判定是否为对照库中所存在的待检测人脸。

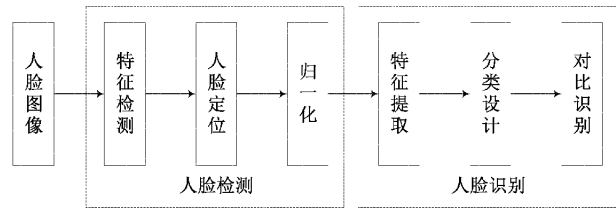


图 1 人脸识别流程系统框图

#### 1) 人脸检测

人脸检测首先要做的是获取面部图像并对其进行检测。本文采用 MTCNN 网络<sup>[14]</sup>来对人脸图像进行检测。人脸图像采集是利用摄像头对包含有面部的照片进行采集,并对其进行特征点检测,通过回归眼睛、鼻子和嘴巴五点关键点定位人脸,从而提取到人脸图像。MTCNN 通过 P-Net, R-Net 以及 O-Net3 个 CNN 级联,对人脸由粗到细依次进行处理<sup>[15]</sup>。其中,P-Net 网络的主要功能是获取面部区域的边框和边框的回归向量;R-Net 网络与 P-Net 相比多了一个全连接层,能够有效抑制 false-positive,主要通过边框和非极大抑制(non maximum suppression, NMS)去掉 false-positive 区域;而 O-Net 网络比 R-Net 多出了一个卷积层,功能与 R-Net 相似,但是精细度会更高一些<sup>[16]</sup>。

#### 2) 人脸识别

人脸识别是指将 MTCNN 所检测到的人脸图像输入到 FaceNet 网络进行识别,即把采集到的人脸图像映射到多维空间,将输入图像转化成为欧式空间中的特征向量,并对两个特征向量之间的欧氏距离进行比较,从而衡量两者之间的相似度,图 2 为 FaceNet 的网络结构。Deep Architecture 是一种稳定的基于深度学习的卷积神经网络架构,是深度学习的骨干网络,FaceNet 的骨干网络选取了深层卷积神经网络 InceptionResNet-v1,把输入的人脸图像转化为向量,获得了不错的识别准确率。然而,由于其 Triplet Loss 的具有较高的运算复杂度以及拥有很大的参数数量,给模型的训练和计算机的硬件条件增添了很大的挑战性。

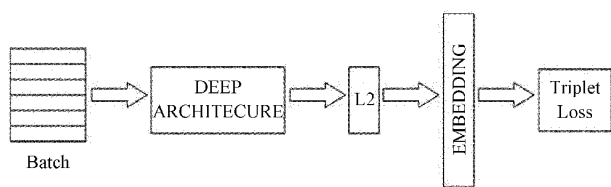


图 2 FaceNet 网络结构

1.2 MobileNet v2 网络

在 2017 年,Google 发布了一种 MobileNet 轻量级卷积神经网络,该网络采用深度可分离卷积<sup>[17]</sup>,将传统的卷积分解为深度卷积和逐点卷积。深度可分离卷积能够显著降低网络的参数量和计算量,同时还可以保证网络的精确度不会被损失掉。但 MobileNet v1 在减少计算量的同时也存在结构简单,性价比较低以及在训练时部分 kernel 为空等其他问题。

MobileNet v2 在 MobileNet v1 的基础上引入了线性瓶颈层和反向残差块,加深了网络层次,但使得体积变小、运行速度更快,同时检测精度也有了有较高的提升。MobileNet v2 网络结构如表 1 所示。其中,第一个卷积层输入大小为  $224 \times 224$ ,通道数为 3,经过卷积后的输出通道数为 32,此时的步长为 2,即卷积核每次移动 2 个距离,以此类推; $f$  为输入输出通道数。MobileNet v2 在深度卷积之前进行升维,加入一个点卷积,可使网络在更高维空间提取特征。此外,为解决 MobileNet v1 激活函数 ReLU 损耗方面的问题,MobileNet v2 在点卷积之后把激活函数 ReLU 换成线性激活函数。并且 MobileNet v2 借鉴了 ResNet 思想使输入和输出相加。这种结构可以使得信息在各层之间流动更加容易,在前向传播时提供特征重用作用,当反向传播时也可以起到缓解梯度消失的作用。

表 1 MobileNet V2 网络结构

Input	Operator	t	c	n	s
$224 \times 224 \times 3$	Conv2d	—	32	1	2
$112 \times 112 \times 32$	bottleneck	1	16	1	1
$112 \times 112 \times 16$	bottleneck	6	24	2	2
$56 \times 56 \times 24$	bottleneck	6	32	3	2
$28 \times 28 \times 32$	bottleneck	6	64	4	2
$28 \times 28 \times 64$	bottleneck	6	96	3	1
$14 \times 14 \times 96$	bottleneck	6	160	3	2
$7 \times 7 \times 160$	bottleneck	6	320	1	1
$7 \times 7 \times 320$	Conv2d1 × 1	—	1 280	1	1
$7 \times 7 \times 1280$	Avgpool7 × 7	—	—	1	—
$1 \times 1 \times f$	Conv2d1 × 1	—	f	—	—

MobileNet v2 网络模型中所提出的模型压缩算法,在保持网络模型的准确率不受影响的同时,运算时间和参数量也大大减少,MobileNet v2 为在移动资源的使用环境受

到限制条件下计算机视觉系统的有效运行提供了一种合理的解决方案,包含了图像分类、目标检测等一些常见的领域。本文之所以选择 MobileNet v2 作为图像分类任务,是因为该网络自身所具备的优点,该方法不但可以降低网络训练的计算量,提高训练速度,还能使得准确率得到有效提高。

2 基于改进的 FaceNet 人脸识别算法

本文基于 FaceNet 网络模型进行人脸识别,并将其骨干网络使用 MobileNet v2 轻量级卷积神经网络替换掉原本的 InceptionResNet-v1,该方法针对原特征提取网络中计算复杂度高、参数量大的问题进行提出,并且为了同时提高识别准确率,在 MobileNet v2 网络中引入轻量型混合卷积注意力机制,优化网络的特征表达能力;其次,由于三元组损失函数收敛困难、网络训练不易,将采用一种基于 SoftmaxLoss 和 TripletLoss 加权融合的联合损失函数,通过调整权值达到最优,替代三元组损失函数来训练网络模型。

2.1 结合轻量型混合注意力模块的 MobileNet-v2 网络

在对口罩遮挡下的人脸进行识别时,人脸的部分特征会由于遮挡而丢失,未遮挡区域会包含更多识别所需要的特征信息。因此,本文在 MobileNet-v2 网络中引入轻量型混合注意力机制<sup>[18]</sup>,将注意力模块插入到 MobileNet v2 网络的瓶颈层之间,让模型将注意力更多地放在未遮挡部分的人脸特征上,根据评判待提取特征的重要性,进一步分配特征图中空间和通道各部分的权重,并获取最具辨识度的人脸特征<sup>[19]</sup>。混合注意力机制是一种结合了空间和通道的注意力机制。如图 3 所示,包含两个独立的模块,空间注意力机制模块和通道注意力机制模块,分别在这两个模块进行空间与通道上的注意力自主学习,可以在节省网络参数的同时节约训练所需要的计算量。

通道注意力模块旨在突出不同通道特征图之间的相关性,并利用网络的自主学习途径,分析每个特征通道的重要性并赋予其不同的权重系数,从而重点强调重要特征,并对非重要的特征进行抑制。使用  $F \in R^{C \times H \times W}$  来表示输入的特征图,其中  $C$  是通道数, $H$  和  $W$  代表特征图的空间维度,通道注意力模块输出特征图的计算方式如下,其中  $PWConv$  代表点卷积,  $\oplus$  代表按位相加,  $\odot$  表矩阵乘法:

$$A_{out} = \text{Softmax}(\text{Transpose}(\text{AvgPool}^{2 \times 2}(PWConv^{1 \times 1}(F_{in})))) \tag{1}$$

$$\hat{F}_{out} = \text{Transpose}(\text{AvgPool}^{2 \times 2}(F_{in})) \tag{2}$$

$$F_{out} = F_{in} \oplus \text{Expand}(PWConv^{1 \times 1}(\hat{F}_{out} \odot A_{out})) \tag{3}$$

在获取通道注意力的过程中,在特征图的空间维度上利用平均池化来压缩特征图的尺寸这一操作进行降维。所用池化核的大小为  $2 \times 2$ ,步长为 2,缩小后的特征图变为原来的  $1/4$  大小,从而减少了后续矩阵乘法的运算时间。通

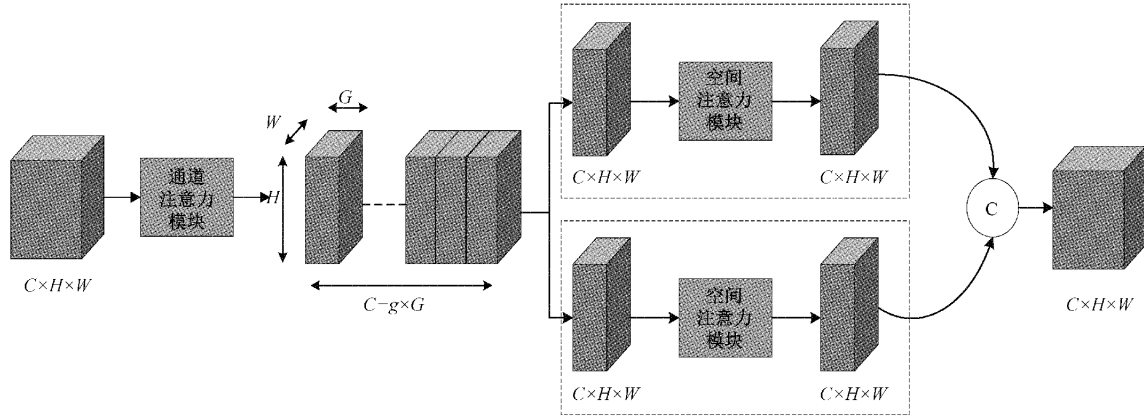


图 3 轻量型混合注意力模块整体结构图

道注意力模块通过为所有查询位置共享全局注意力图来减少计算量,明显降低了参数数量和计算复杂度,使其能够应用于轻量级骨干网络的多个层中,以便更好地捕获通道间依赖关系,不过计算成本会略有增加。

空间注意力模块作为通道注意力模块的补充,其主要是为了突出最具信息量的那部分。该模块通过沿着通道采用平均池化操作和最大池化操作,计算空间注意力,然后将它们连接成一个有效的特征图,最后通过卷积操作得到一个大小为  $R \times H \times W$  的空间注意力权重特征图。对于任意一组子空间中输入的特征图,其输出特征图的计算方式如下所示,其中  $\otimes$  代表按位相乘:

$$\hat{F}_{out} = \text{AvgPool}^{3 \times 3}(\text{DWConv}^{1 \times 1}(F_{in})) \oplus \text{MaxPool}^{3 \times 3}(\text{DWConv}^{1 \times 1}(F_{in})) \quad (4)$$

$$A_{out} = \text{Expand}(\text{Softmax}(\text{PWConv}^{1 \times 1}(\hat{F}_{out}))) \quad (5)$$

$$F_{out} = (F_{in} \otimes A_{out}) \oplus F_{in} \quad (6)$$

通过将输入特征图划分到多个不同卷积核的子空间,使得网络可以生成多频、多尺度的特征表示。原始图像中包含了低频和高频分量,低频分量是指局部变化缓慢的特征,高频分量则表示图像中的细微细节,特别是当被遮挡人脸图像样本中存在比较大的类内差异时,多频特征学习方法可以让模型有能力关注更精细的图像高频特征,即网络可以将注意力重点放在人脸非遮挡部分的细节特征上。

注意力机制的本质就是利用网络的自主学习能力来获得所需的权重系数,并使用动态加权的方式来赋予其权重,以此来强调网络在学习过程中抑制不重要特征、突出重要特征的机制。在对人脸识别的研究中,网络通过引入注意力机制可以更好的将注意力集中在面部特征关键点上。将人脸图像输入网络之后,通过卷积层对图像进行逐层处理,而注意力机制则会作用在每个特征图中,通过其自主学习能力,可以使网络在特征提取时将注意力放在更具区分性的特征向量上,不同人脸特征向量之间的距离更大。

## 2.2 损失函数

在进行算法模型的训练过程中,损失函数是可以反映

出模型实际数据与模型输出值二者间误差距离的关键性因素。本研究在 FaceNet 网络模型的训练过程中,针对网络模型的优化问题,采用了 softmaxLoss<sup>[20]</sup> 和 TripletLoss<sup>[21]</sup> 加权融合的方式。其中,softmax Loss 更加强调的是类与类之间的可分性,而对于分类内部的分布效果并不是很好,Triplet 损失函数通常被用于训练差异性较小的示例<sup>[21]</sup>。本文通过采用联合训练两个损失函数的方法,可以得到更好的模型收敛效果。

Triplet 损失函数可以由式(7)表示:

$$L_t = \sum_i^N [d(a, p)^2 - d(a, n)^2 + \text{margin}]_+ \quad (7)$$

式中:  $L_t$  表示损失;  $a$  用来表示锚示例;  $p$  用来表示正示例;  $n$  用来表示负示例;  $N$  代表一组 Batch 中的全部示例个数。

在处理图像多分类任务时,通常使用 softmax 损失函数,其具体表示如式(8)所示。

$$L_s = -\frac{1}{N} \sum_i^N \log \frac{\exp(\mathbf{W}_{y_i}^T \mathbf{X}_i + b_{y_i})}{\sum_j^C \exp(\mathbf{W}_j^T \mathbf{X}_i + b_j)} \quad (8)$$

式中:  $L_s$  被用来表示损失;  $\mathbf{X}_i$  是第  $i$  个步态图像的特征向量;  $y_i$  用来代表  $\mathbf{X}_i$  的真实标签;  $b$  为偏置;  $\mathbf{W}_{y_i}^T$  与  $\mathbf{W}_j^T$  分别表示将  $X_i$  判别为  $y_i$  类和  $j$  类的权向量;  $C$  用来表示总类别数量;  $\mathbf{W}_{y_i}^T \mathbf{X}_i + b_{y_i}$  表示步态图像在类别  $y_i$  上的得分;  $N$  代表整个训练样本的数量。

最终本文所采用的联合损失函数为:

$$\text{Loss} = \gamma L_t + \lambda L_s \quad (9)$$

式(9)中,  $\gamma$  和  $\lambda$  分别被用来表示 TripletLoss 和 SoftmaxLoss 两个损失函数的权重;当  $\gamma = 0$  且  $\lambda \neq 0$  时为 Softmax 损失函数;当  $\gamma \neq 0$  且  $\lambda = 0$  时为 Triplet 损失函数;当  $\gamma \neq 0$  且  $\lambda \neq 0$  时则为联合损失函数。

## 3 实验结果与分析

### 3.1 实验环境和数据集

本文实验平台是在 Windows 10 64 位专业版服务器上

进行,处理器为 Intel(R) Xeon(R) CPU E5-2695 v4@2.10 GHz 2.10 GHz(2 处理器)、内存(RAM)256GB、GPU(NVIDIA TITAN Xp)共 4 块、显卡内存 12 GB;软件环境为 Pycharm2020、Pytorch1.7、Python3.6、CUDA10.2 等。

本实验使用来自武汉大学开源口罩遮挡数据集来验证研究的有效性,为从网络爬取样本,经过整理、清洗和标注后的真实口罩人脸数据集,含 525 位不同名人的人脸图像,其中包含 5 千张戴口罩的人脸和 9 万张正常人脸,用于训练和测试口罩遮挡下人脸识别相关的机器学习模型。部分佩戴口罩和未佩戴口罩的人脸图像分别如图 4、5 所示。



图 4 数据集中没有戴口罩的人脸图像



图 5 数据集中戴口罩的人脸图像

### 3.2 实验结果及数据分析

通过现有人脸识别算法与本文所提出的算法进行实验对比,以验证本实验所提出的结合轻量型混合注意力模块的 MobileNet-v2 网络和联合损失函数对口罩遮挡下人脸识别速度和准确率的有效性。

#### 1) 识别速率对比

表 2 为 MobileNet v2 和 Inception-ResNet-v1 网络模型

参数量与计算量的对比,从其中可以看出 MobileNet v2 与 Inception-ResNet-v1 网络相比,参数量小并且计算量减少了 2.5 倍以上。参数量和计算量是用来衡量深度学习网络的重要指标,参数量对应网络的空间复杂度即所消耗的计算机内存资源,而计算量所对应的是网络的时间复杂度即网络所耗费的时间。因此,可以得出 MobileNet v2 网络相比原网络实现了识别速率的有效提升。

表 2 网络性能对比

网络模型	计算量	参数量
Inception-v1	1 600	7.5
MobileNet v2	569	4.2

#### 2) 口罩遮挡对不同损失函数的影响分析

本文采用 FaceNet, SphereFace, CosFace 三种人脸识别算法在本实验选取的数据集上进行集中测试,这 3 种算法主要是在其损失函数的不同上。FaceNet 使用 Triplet Loss 来衡量输入人脸图像在欧式空间上的距离,其目的是为了减少相同人脸样本间的差距,并增大不同人脸样本之间的差距。SphereFace 采用角度 Softmax Loss 作为其损失函数,其作用于 Softmax Loss 类似,可以增加网络生成的特征中同一类的相似性和不同类的区分度。而 CosFace 则是引入了余弦距离。在本类实验中,采用准确率作为人脸识别评判指标,准确率指标是指网络识别到的正确样本数量占总样本数量的百分比,实验结果如表 3 所示。

表 3 人脸识别网络在数据集上的实验结果

对比方法	无遮挡	有遮挡
FaceNet	99.63%	87.68%
SphereFace	99.42%	88.3%
CosFace	99.6%	88.7%

可以看出,现有人脸识别准确率在无遮挡条件下非常高,但是在口罩遮挡的条件下,各类人脸识别算法都受到了不同程度的影响,使得识别准确率下降,其主要原因是由于口罩的遮挡,人脸大部分可供识别的特征无法被提取。

因此,在本文算法引入了轻量级混合卷积注意力机制并且采用联合损失函数来优化 FaceNet 网络模型,让网络更加关注人脸未遮挡部分的特征,抑制对无效特征的关注。表 4 为不同算法在本实验数据集上的识别准确率对比,在 40%~50% 的人脸口罩遮挡面积占比下,通过对比实验,可以明显看到与原算法相比,识别准确率得到了一定程度的提升,可以进一步优化网络的特征表达能力。

模型通过给定的优化处理方案达到了准确度、大小和能耗等性能参数上的均衡。通过对以上实验结果的分析,可以验证本算法的有效性,其对口罩遮挡下的人脸识别速度和准确率都有了一定程度的提升。

表 4 改进网络与原网络的识别准确率对比

使用算法	识别准确率
Inception-v1	87.68%
mobileNet	88%
本文算法	92.1%

## 4 结 论

针对现有人脸识别系统进行大面积遮挡的人脸图像识别时存在的精度低、速度慢的问题,本文提出了一种基于改进 FaceNet 人脸识别网络模型的口罩遮挡下的人脸识别方法。本研究中对口罩遮挡下的人脸识别问题基于 MTCNN 进行人脸检测并且使用 FaceNet 完成人脸识别以及人脸特征向量提取并对其从识别速率和识别准确率两方面进行改进。使用轻量级网络 MobileNet v2 替代了原本 FaceNet 人脸识别网络中的 Inceptionv1 作为其骨干网络,通过向 MobileNet v2 网络中引入轻量级混合卷积注意力机制,优化了网络的特征表达能力,并且对 FaceNet 损失函数进行了进一步改进,在提高网络识别速率的同时,大幅提升了网络的识别准确率。同时,本研究通过对比实验,证明了在同等遮挡条件下,所提出的改进算法在识别准确率、识别速率、模型参数量以及模型计算量方面均优于原有的人脸识别网络。

本文所提出的改进算法实现了在口罩大面积遮挡条件下面部识别准确率的有效提升。然而,该模型的移植性和泛化能力依旧需要继续完善,这同时也是目前网络学习模型中普遍存在并需要解决的问题。在下一阶段的研究中,需要继续对网络模型进行改进,在确保人脸识别网络性能的基础上,进一步提高识别的精度与速度。

## 参考文献

- [1] 王慧星,黄勃,高永彬,等.部分遮挡人脸识别的方法综述[J].武汉大学学报(理学版),2020,66(5):451-461, DOI:10.14188/j.1671-8836.2020.0148.
- [2] LECUN Y, BENGIO Y, HINTON G. Deep learning[J]. Nature, 2015, 521(7553): 436-444.
- [3] SCHROFF F, KALENICHENKO D, PHILBIN J. Facenet: A unified embedding for face recognition and clustering[C]. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2015: 815-823.
- [4] HUANG D, WANG C D, WU J S, et al. Ultra-scalable spectral clustering and ensemble clustering [J]. IEEE Transactions on Knowledge and Data Engineering, 2019, 32(6): 1212-1226.
- [5] WILLIAM I, RACHMAWANTO E H, SANTOSO H A, et al. Face recognition using facenet (survey, performance test, and comparison)[C]. 2019 Fourth International Conference on Informatics and Computing(ICIC), IEEE, 2019: 1-6.
- [6] LIU W, WEN Y, YU Z, et al. Large-margin softmax loss for convolutional neural networks [J]. ArXiv Preprint, 2016, ArXiv:1612.02295.
- [7] 谷凤伟,陆军,夏桂华. facenet 皮尔森判别网络的人脸识别方法[J]. 智能系统学报, 2022, 17(1): 107-115.
- [8] SIMONYAN K, ZISSERMAN A. Very deep convolutional networks for large-scale image recognition[J]. ArXiv Preprint, 2014, ArXiv:1409.1556.
- [9] 石翠萍,谭聪,左江,赵可新.基于改进 AlexNet 卷积神经网络的人脸表情识别[J]. 电讯技术, 2020, 60(9): 1005-1012.
- [10] IANDOLA F N, HAN S, MOSKEWICZ M W, et al. SqueezeNet: AlexNet-level accuracy with 50x fewer parameters and < 0.5 MB model size [J]. ArXiv Preprint, 2016, ArXiv:1602.07360.
- [11] HOWARD A G, ZHU M, CHEN B, et al. Mobilenets: Efficient convolutional neural networks for mobile vision applications [J]. ArXiv Preprint, 2017, ArXiv:1704.04861.
- [12] SANDLER M, HOWARD A, ZHU M, et al. Mobilenetv2: Inverted residuals and linear bottlenecks[C]. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2018: 4510-4520.
- [13] ZHANG X, ZHOU X, LIN M, et al. Shufflenet: An extremely efficient convolutional neural network for mobile devices [C]. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2018: 6848-6856.
- [14] ZHANG K, ZHANG Z, LI Z, et al. Joint face detection and alignment using multitask cascaded convolutional networks [J]. IEEE Signal Processing Letters, 2016, 23(10): 1499-1503.
- [15] 赖欣,王储,陈航.低照度下人脸检测 MSRCR 光频分段滤波增强算法[J]. 电子测量与仪器学报, 2022, 36(2): 96-106, DOI:10.13382/j.jemi.B2104626.
- [16] 陈斐凡,关浩杰,田海丽.基于 FaceNet 算法的公交车人脸识别系统设计与实现[J]. 信息与电脑, 2021, 33(5): 3.
- [17] 王丹峰,陈超波,马天力,等.基于深度可分离卷积的 YOLO3 行人检测算法[J]. 计算机应用与软件, 2020, 37(6): 218223.
- [18] 张泽哈.基于局部遮挡人脸识别的移动端人员管控系统设计与实现 [D]. 北京邮电大学, 2021, DOI: 10.26969/d.cnki.gbydu.202.000896.

- [19] 冯宇平,管玉宇,杨旭睿,等.融合注意力机制的实时行人检测算法[J].电子测量技术,2021,44(17):123-130. DOI:10.19651/j.cnki.emt.2107174.
- [20] LI X, WANG W. Learning discriminative features via weights-biased softmax loss[J]. Pattern Recognition, 2020, 107: 107405.
- [21] 赵佳鑫,张雅丽,施新凯,等.基于联合损失函数的多视角步态识别方法研究[J].智能计算机与应用,2022,12(2):13-17.

### 作者简介

**移洁**,硕士研究生,主要研究方向为计算机视觉和图像处理。

E-mail:yijie0411@163.com

**侯劲**,教授,主要研究方向为人工智能、智能检测等。

E-mail:houjin828@sina.com

**石浩德**,硕士研究生,主要研究方向为计算机视觉、图像修复和图像分割。

E-mail:1843945104@qq.com