

张拓. 基于多路径融合的人脸识别算法研究[J]. 智能计算机与应用, 2024, 14(7): 64-70. DOI: 10.20169/j.issn.2095-2163.240709

## 基于多路径融合的人脸识别算法研究

张拓

(中国矿业大学(北京)人工智能学院, 北京 100083)

**摘要:** 本文针对 FaceNet 算法在低光照和有遮挡情况下的人脸识别准确率较低的问题, 提出了一种多路径融合的人脸识别算法, 该算法在 FaceNet 算法基础上引入多路径融合模块, 包括 3 条并行的路径: 采用 Ghost 瓶颈模块加强特征提取, CBAM 注意力机制模块提取图像的重要特征以及 ASPP 模块对图像的特征进行多尺度提取, 并通过 Sigmoid 门控单元对每个路径自身进一步提取来加强有用信息; 对经由三条路径处理后的特征图进行加权融合。在此基础上引入 FPN 模块, 加强对不同感受野下的特征提取。除此之外, 还在原有 Tripletloss 损失函数的基础上加入交叉熵分类损失, 以确保模型能更快地收敛。实验结果表明, 本文提出的多路径人脸识别算法取得了较好的识别效果, 在 LWF 数据集的准确率达到 99.7%, 在自建的低光照有遮挡数据集的准确率达到 98.2%, 可准确地识别低光照、遮挡人脸目标。

**关键词:** 人脸识别; 多路径融合; FPN

中图分类号: TP391.41

文献标志码: A

文章编号: 2095-2163(2024)07-0064-07

### Research on face recognition algorithm based on multi-path fusion of human features

ZHANG Tuo

(School of Artificial Intelligence, China University of Mining and Technology(Beijing), Beijing 100083, China)

**Abstract:** This paper addresses the issue of lower accuracy in face recognition of the FaceNet algorithm under low-light conditions and occlusion by proposing a multi-path fusion face recognition algorithm. This algorithm integrates a multi-path fusion module into the FaceNet framework, comprising three parallel paths: employing Ghost bottleneck structures to enhance feature extraction, utilizing the CBAM attention mechanism module to extract crucial image features, and using the ASPP module for multi-scale feature extraction, further enhanced through Sigmoid gate control units for each path to strengthen useful information. Subsequently, the feature maps of the three paths are fused with weights. Furthermore, an FPN module is incorporated to enhance feature extraction for different receptive fields. Additionally, a cross-entropy classification loss is added to the original Tripletloss loss function to ensure faster convergence of the model. Experimental results demonstrate that the proposed face recognition algorithm achieves excellent recognition performance, with an accuracy of 99.7% on the LWF dataset and 98.2% on a self-built dataset with low-light and occlusion conditions, accurately identifying faces under low-light and occluded scenarios.

**Key words:** face recognition; multi-path fusion; FPN

## 0 引言

人脸识别是基于人面部特征信息进行身份识别的一种技术<sup>[1]</sup>, 在重要场景的安全防护、交通运输等领域有着十分广泛的应用<sup>[2]</sup>。人脸识别技术已经广泛应用于各种场景, 其能够更精准与快速地对采集到的目标人脸进行识别, 是近年来在该领域较为热门的研究方向<sup>[3]</sup>。

人脸识别算法从早期的基于几何特征、全局信

息和局部信息的传统算法, 逐渐发展为基于深度学习的方法。文献[4]首次将深度学习方法用于人脸识别, 提出了利用卷积神经网络生成人脸图片特征向量编码的方法, 在人脸对比时, 首先构建待对比人脸库并生成对应的编码向量, 通过计算拍照生成的人脸编码向量与人脸库中的编码向量之间的欧几里得距离最接近的距离来确认人脸身份; 文献[5]在文献[4]的基础上融合了注意力机制模块, 能较精准地识别无遮挡人脸, 但是对有遮挡人脸识别的效

果并不明显;文献[6]在 FaceNet 的基础上优化了特征提取网络与损失函数,将主干特征提取网络换成轻量级网络,并在损失函数中融合分类损失,缩短了网络训练的时间。除了使用主流图像进行人脸识别,还出现了基于点云算法的人脸识别方法,文献[7]提供了一种单模态点云算法,通过将训练集替换为点云数据训练与预测。但点云算法也存在一些弊端,公开数据集较为有限,数据标注困难。文献[8]提出加性角边损失函数,通过在损失函数中引入边值, ArcFace 增强了类别间的可分性,使得算法更有效地识别和区分不同人脸特征;文献[9]提出了一种无需额外人工注释即可平衡每个群组中的样本数量的方法,解决了传统面部识别模型在处理大量面部图像时的信息编码不足问题;文献[10]提出一种改进的 FaceNet 算法,使用 Mish 激活函数训练来提高网络学习能力,除此之外,还引入了注意力机制模块与接受场模块以适应复杂环境中的不同场景,在 FaceNet 的基础上有了一定的提升;文献[11]提出了一种加权平均池化算法,解决了 FaceNet 算法中未能充分考虑面部各个区域重要性差异的问题。

本文提出了一种基于多路径融合的人脸识别算法,在 FaceNet 算法的基础上串入多路径融合模块,共分 3 条路径:路径一采用 GhostNet 中的 Ghost 瓶颈模块加强特征提取;路径二采用 CBAM (Convolutional Block Attention Module) 注意力机制

模块提取图像的重要特征;路径三采用 ASPP (Atrous Spatial Pyramid Pooling) 模块对图像进行多尺度特征提取,并通过 Sigmoid 门控单元对每个路径的特征进一步提取来突出关键信息,再将三条路径的特征图进行融合。随后串入 FPN (Feature Pyramid Network) 模块,加强对不同感受野下的特征提取。在损失函数方面,为保证模型能尽快收敛,本文还在原有损失函数的基础上加入交叉熵损失。除此之外,本文还对多路径融合模块进行了对照试验,比较了门控单元与加权融合单元的先后顺序对实验结果的影响,结果表明在加权单元之前使用门控单元能够取得更好的识别效果。总之相较于 FaceNet 算法,本文算法有更好的识别效果。

## 1 相关工作

### 1.1 FaceNet 简述

FaceNet 算法的原理是将人脸 RGB 图像通过深度神经网络映射到高维空间,以两张人脸编码的直接欧几里得距离为评价标准,判断两张人脸是否来自同一人。该算法的处理过程如图 1 所示,算法的第一部分将人脸通过骨干特征提取网络得到人脸特征向量;第二部分为了方便地计算两张人脸图像之间的欧氏距离,采用 L2 标准化将人脸特征向量进行归一化;第三部分设置阈值,并通过欧几里得距离来识别人脸。

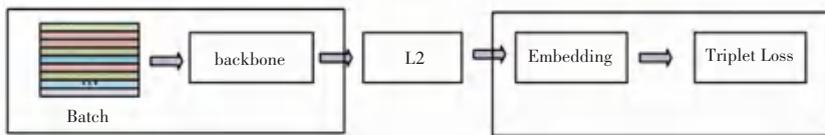


图 1 FaceNet 处理过程

Fig. 1 FaceNet processing process

FaceNet 使用 TripletLoss 损失作为损失函数,通过设置阈值的方式来控制正负样本之间的欧氏距离。TripletLoss 示意图如图 2 所示,每次选择 3 个样本进行损失计算,一个样本标记为  $A_i$ ,另一个是同一人的正样本标记为  $P_i$ ,最后一个是其他人的负样本标记为  $N_i$ 。在训练过程中,通过优化模型,确保  $A_i$  到  $P_i$  的距离逐渐小于  $A_i$  到  $N_i$  的距离。

TripletLoss 损失函数的计算公式(1):

$$L = \sum \max(d(a,p) - d(a,n) + \alpha, 0) \quad (1)$$

其中,  $a$  代表每个样本标记  $A_i$ ;  $p$  代表正样本  $P_i$ ;  $n$  代表负样本  $N_i$ ;  $d$  则代表样本之间的欧几里得距离;  $\alpha$  是一个超参数,规定了损失对比过程中的阈值。

TripletLoss 损失函数的作用就是通过网络的训练,使得标记样本到正样本之间的欧氏距离与阈值相加后的结果小于标记样本到负样本之间的欧氏距离,即:  $d(a,p) + \alpha \leq d(a,n)$ 。

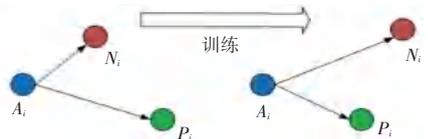


图 2 TripletLoss 示意图

Fig. 2 TripletLoss diagram

### 1.2 FaceNet 算法改进

#### 1.2.1 多路径融合模块

为了加强对人脸特征的提取,本文在 FaceNet

算法的基础上串入多路径融合模块,如图3所示。

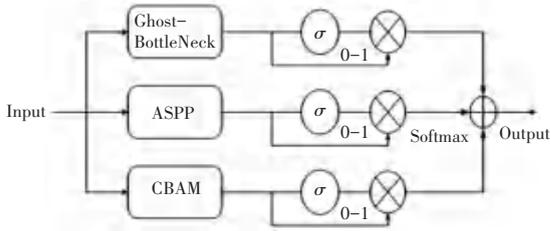


图3 多路径融合模块

Fig. 3 Multi-path fusion module

### 1) Ghost 瓶颈模块

Ghost 模块是 GhostNet 中的特征提取模块<sup>[12]</sup>。首先,通过卷积层变化原始特征图的通道数;其次,对经由卷积层处理后的特征图的每个通道维度进行廉价卷积处理,再将每个处理后的特征图与最初经由卷积层处理后的特征图在通道维度上拼接。将 Ghost 模块级联处理,先升维再降维,再通过 shortcut 连接,即为 Ghost 瓶颈模块,本文在第一个路径采用 Ghost 瓶颈模块加强对图像特征的提取,Ghost 模块和 Ghost 瓶颈结构如图4所示。

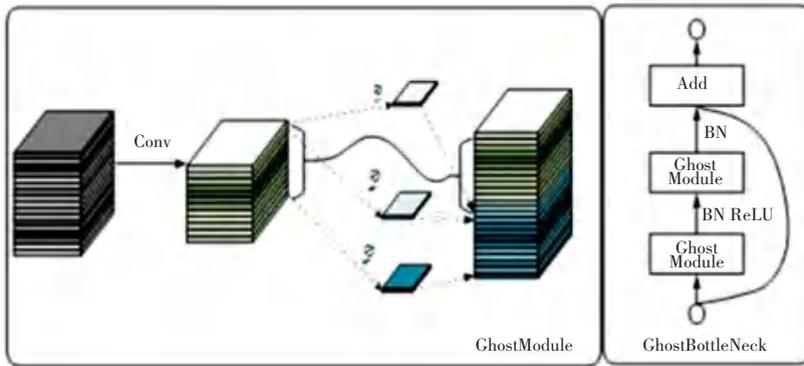


图4 Ghost 模块和 Ghost 瓶颈模块

Fig. 4 Ghost module and Ghost bottleneck module

### 2) ASPP 模块

ASPP 模块由多个卷积支路和一个池化支路构成,每个卷积支路使用不同的采样率,以捕获多尺度上下文信息<sup>[13]</sup>。这些支路产生的特征图具有不同的感受野,通过融合操作将其合并成一个更丰富的

特征图。融合后的特征表示包含了来自不同尺度和不同上下文信息的特征,有助于提高网络的性能,ASPP 模块如图5所示,Conv 代表卷积操作, $k$  代表卷积核尺度, $s$  代表步长, $p$  代表膨胀率。

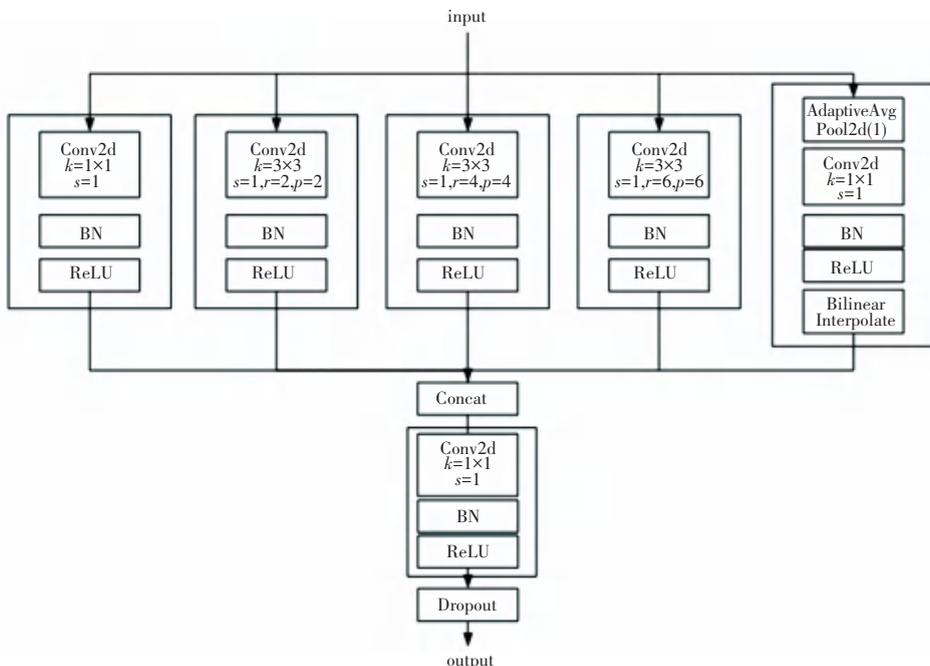


图5 ASPP 模块

Fig. 5 ASPP module

### 3) CBAM 注意力机制

CBAM 模块是一种可以对输入特征图动态调整的注意力机制模块,旨在增强网络对重要特征的感知力,提高模型对复杂场景和对象的表达能力<sup>[14]</sup>。CBAM 模块分为通道注意力模块 (CAM) 和空间注意力模块 (SAM),如图 6 所示。

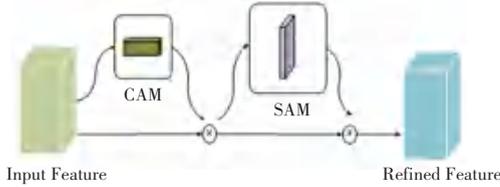


图 6 CBAM 注意力机制  
Fig. 6 CBAM attention mechanism

引用通道注意力机制模块可以加强网络对特征图中高维信息的关注,将特征图分别进行平均池化与最大池化处理,之后对平均池化与最大池化的结果利用多层感知机 (MLP) 进行处理并将结果相加,通过 Sigmoid 函数得到权重矩阵,再与原特征图对应位置的元素相乘,如图 7 所示。

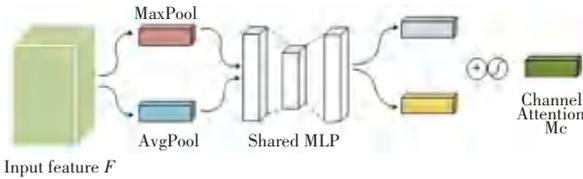


图 7 通道注意力机制模块  
Fig. 7 Channel attention mechanism module

CAM 计算公式(2):

$$M_c(F) = \sigma(\text{MLP}[\text{AvgPool}(F) + \text{MaxPool}(F)]) = \sigma\{W_1[W_0(F_{\text{avg}})] + W_1[\alpha(F_{\text{max}})]\} \quad (2)$$

其中,  $W_0$  与  $W_1$  都是可学习的参数; AvgPool 代表平均池化操作; MaxPool 代表最大池化操作; MLP 代表多层感知机;  $\sigma$  代表 Sigmoid 函数的运算操作。

引入空间注意力机制模块可以加强网络对重要位置所包含信息的关注。该模块对输入特征层的每个特征点的通道分别计算最大值和平均值,并将计算结果拼接。通过一次通道数为 1 的卷积运算对拼接后的结果进行调整。用 Sigmoid 函数获取权重矩阵,并与原始特征图相乘。空间注意力机制模块如图 8 所示。

SAM 的计算公式(3):

$$M_s(F) = \sigma(f^{7 \times 7}[\text{AvgPool}(F); \text{MaxPool}(F)]) =$$

$$\sigma\{f^{7 \times 7}[F_{\text{avg}}^s; F_{\text{max}}^s]\} \quad (3)$$

其中,  $f^{7 \times 7}$  代表  $7 \times 7$  的卷积运算。

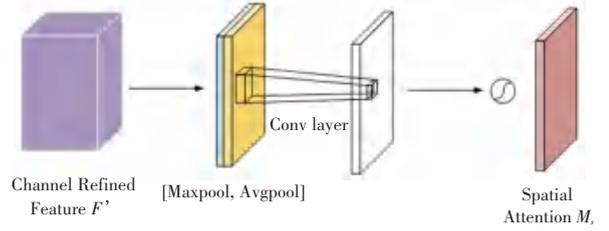


图 8 空间注意力机制模块  
Fig. 8 Spatial attention mechanism module

CBAM 的整体计算公式,公式(4)和公式(5):

$$F' = M_c(F) \otimes F \quad (4)$$

$$F'' = M_s(F') \otimes F' \quad (5)$$

其中,  $M_c$  和  $M_s$  分别代表通道注意力机制运算函数与空间注意力机制运算过程,经由通道注意力机制模块处理后得到的结果为  $F'$ ,经由空间注意力机制处理后得到的结果为  $F''$ 。

在每个路径处理完成后,本文通过对每个路径输出的特征图的每一个位置的参数用 Sigmoid 函数计算,生成相应的权重矩阵;将对应路径的特征图与对应路径的权重矩阵做点乘处理,生成新的特征图。为了更灵活地调整每个路径在每个位置的重要性,本文引入了 3 个可学习参数,分别为 3 个特征图的权重矩阵,由神经网络进行训练得出。为了确保在加权过程中权重之和为 1,本文对每个通道每个对应位置的权重参数应用 Softmax 函数进行处理,使得模型能够自适应地学习路径之间的权重分布,从而提高网络的适应性。

多路径融合模块中融合过程的计算方法:假设  $F_1, F_2, F_3$  为 3 个路径的特征图,  $F_{ij}$  为特征图中每个元素的值,  $\sigma\{\cdot\}$  为利用 Sigmoid 函数进行计算的定义,则经特征融合之后的特征图中的每个元素  $O_{ij}$  可以表示为式(6):

$$O_{ij} = \frac{e^{A_{ij}} \sigma\{(F_1)_{ij}\} + e^{B_{ij}} \sigma\{(F_2)_{ij}\} + e^{C_{ij}} \sigma\{(F_3)_{ij}\}}{e^{A_{ij}} + e^{B_{ij}} + e^{C_{ij}}} \quad (6)$$

其中:  $A_{ij}, B_{ij}, C_{ij}$  为可学习的权重参数矩阵。

#### 1.2.2 FPN 模块

FPN 是一种专门解决多尺度信息处理的模块,如图 9 所示,利用自顶向下的特征传播构建了一个特征金字塔,通过下采样操作将语义丰富的高层信息传递到所有层次,提高了特征提取的全面性,从而更好地适应不同尺寸物体的识别和检测任务。

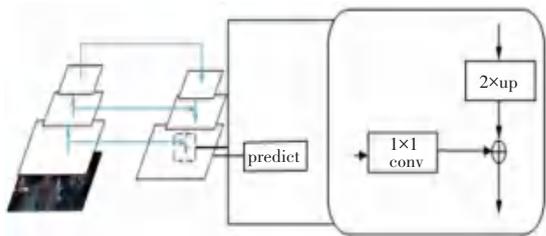


图9 FPN 模块

Fig. 9 FPN module

### 1.2.3 TripletLoss 损失函数改进

在 FaceNet 的训练过程中,存在网络模型规模庞大,训练难度较大且收敛缓慢等问题。为了解决这些问题,本文采用了融合交叉熵损失的策略:将交

叉熵损失融入 TripletLoss 损失函数,通过提供更加丰富的监督信号来加速网络的收敛,克服大规模网络训练的挑战,促使网络更有效地学习特征判别的相关内容,从而提高模型整体的性能。改进后损失函数的公式(7):

$$L = L_t - \frac{1}{N} \sum_{i=1}^N \sum_{k=1}^m y_k^{(i)} \log(p_k^{(i)}) \quad (7)$$

其中,  $L_t$  代表 TripletLoss 损失函数;  $N$  代表样本个数;  $m$  人脸代表类别个数;  $y_k^{(i)}$  代表二值指示函数;  $p_k^{(i)}$  代表类别概率。

### 1.3 整体网络结构图

本文在 FaceNet 的基础上,串入多路径融合模块与 FPN 特征金字塔模块来加强特征提取,以提高模型的性能,网络结构如图 10 所示。

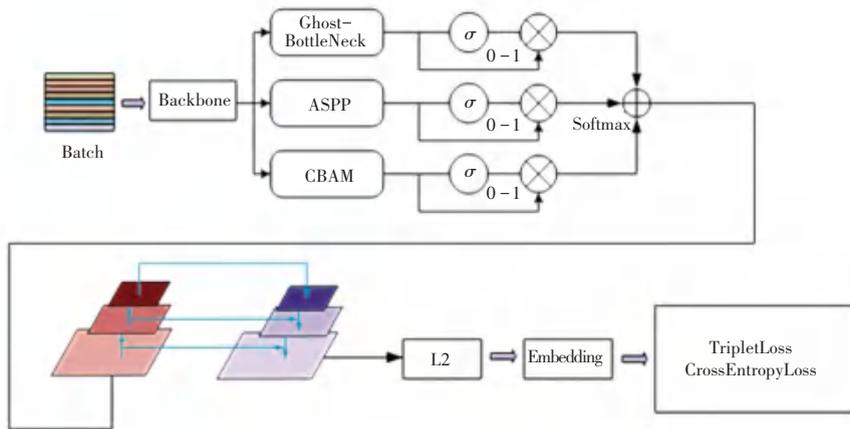


图10 网络结构图

Fig. 10 Network structure diagram

## 2 实验验证

本文实验采用 CASIA-WebFace 数据集进行训练,采用 LFW 数据集和自建低光照有遮挡 BLFW 数据集作为测试集来验证模型的效果。LFW 数据集是人脸识别领域经典数据集之一,其中包含数千人在自然场景中的数万张照片;BLFW 数据集是自建数据集,于煤矿井下采集,共采集了 100 个人 3 000 多张人脸图片。为了测试模型的鲁棒性,将部分图像做翻转、遮挡、贴图处理,BLFW 数据集的格式参照 LFW 数据集进行标注。

在实验中,为了验证本文所提出的多路径融合模块与 FPN 模块对网络性能提升的影响,针对 FPN 模块与本文所提出的多路径融合模块进行模块消融实验,消融实验结果见表 1。可见相较于 FaceNet 算法,并入多路径融合模块后,模型在 FLW 数据集上

的准确率提高了 0.6 个百分点,在 BFLW 数据集上的准确率提高了 0.6 个百分点,但是在模型运行时间方面,模型的运行时间多消耗了 15.1 ms;而串入 FPN 模块之后,模型在 FLW 数据集上的准确率达到了 99.7%,在 BFLW 数据集上的准确率达到 98.2%,但是在模型运行时间方面,模型运行时间多消耗了 19.3 s,相较于只串入多路径融合模块的网络相比,在准确率上有了进一步的提升。所以可以得出结论,本文所提出的模型相较于 FaceNet 有助于更精准地识别人脸,且在有遮挡、低光照的人脸识别方面有着较好的表现能力。

除此之外,本文还对单张人脸图片进行了测试,效果图如图 11。从单张人脸的测试结果同样不难看出,本文算法有更好的识别效果。

多路径融合模块首先通过多个支路对特征图进行处理,随后将处理结果进行融合,本文对加权融合

与 Sigmoid 门控单元处理的次序进行了调整,将每个支路产生的特征图先进行加权融合,再应用

Sigmoid 门控单元处理。两种融合方式的对比如图 12 所示。

表 1 子模块消融实验  
Table 1 Submodule ablation experiment

算法	多路径融合模块	FPN 模块	准确率/%		时间/ms
			FLW 数据集	BFLW 数据集	
FaceNet 算法	-	-	99.0	97.5	28.6
本文算法	✓	-	99.6	98.1	43.7
	✓	✓	<b>99.7</b>	<b>98.2</b>	47.9



(a) 本文算法



(b) FaceNet 算法

图 11 单张人脸测试效果图

Fig. 11 Single face testing result image

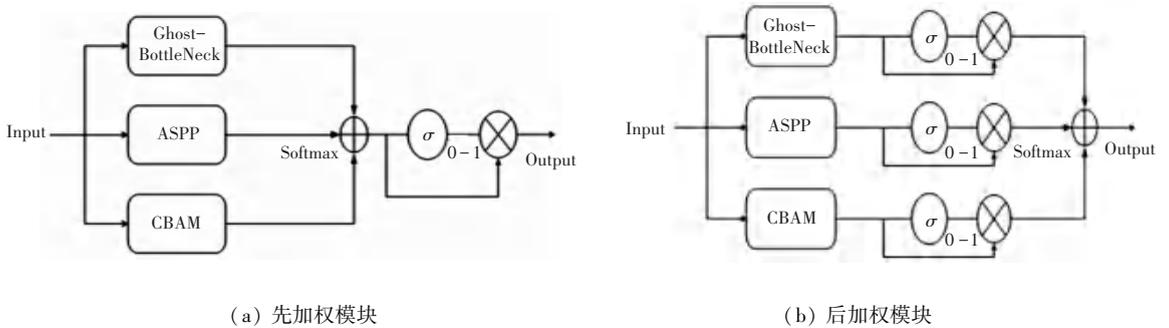


图 12 融合方式对比

Fig. 12 Comparison of order methods

在 BFLW 数据集上做测试实验,使用 ROC 曲线来衡量两种融合方式的性能。ROC 将不同阈值下的真正例率和假正例率相连,曲线下的面积(AUC 值)即为性能评价的指标,AUC 值越高则表明模型具有较好地区分正负样本的能力,即在尽可能少地

将负样本误判为正样本的同时,正确地将更多的正样本识别出来。两种融合方式在 BFLW 数据集中 ROC 曲线对比图如图 13 所示,可以看出将后加权模块串入网络中模型有更好的识别效果。



图 13 ROC 曲线对比

Fig. 13 ROC curve comparison

### 3 结束语

本文提出了一种改进的 FaceNet 算法,提出了多路径融合模块:3 条路径分别采用 Ghost 瓶颈模块加强对图像特征的提取;采用 ASPP 模块加强对多尺度信息的捕捉;采用 CBAM 注意力机制模块加强对网络重要信息的提取,进一步提高网络对关键信息的敏感性,随后将 3 个支路经 Sigmoid 门控单元处理后进行加权融合,在此基础上串入 FPN 模块,增强不同感受野下的特征提取。除此之外还优化了损失函数,在原有损失函数的基础上加入交叉熵分类损失,使模型可以尽快收敛。最后,还对比了两种多路径融合模块在对应评估集中的性能,表明后加权的方法有更好的表现。实验表明,本文算法有助于人脸识别准确率的提升,尤其是低光照、有遮挡的条件下采集到的人脸图片。但是在识别速率上,由于模型的参数增多,识别速率有所下降,有待后续研究优化。

### 参考文献

- [1] RANJAN R, SANKARANARAYANAN S, BANSAL A, et al. Deep learning for understanding faces: Machines may be just as good, or better, than humans[J]. IEEE Signal Processing Magazine, 2018, 35(1):66-83.
- [2] 冯婧, 顾梅花, 刘晓龙, 等. 课堂场景下改进 MobileNetv3 人脸识别算法[J]. 国外电子测量技术, 2022, 41(10):47-55.
- [3] 马博宇, 尉寅玮. 基于 AdaBoost 算法的人脸识别系统的研究与实现[J]. 仪器仪表学报, 2016, 37(S1):162-167.
- [4] 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. IEEE, 2015: 815-823.
- [5] 张晋婧, 刘双峰, 丰雷, 等. 融合注意力机制的人脸识别算法研究[J]. 国外电子测量技术, 2023, 42(2):107-113.
- [6] 张颖. 基于 FaceNet 改进的人脸识别算法[J]. 电子技术, 2022, 51(3):25-27.
- [7] LANG A H, VORA S, CAESAR H, et al. Pointpillars: Fast encoders for object detection from point clouds [C]// Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. IEEE, 2019: 12697-12705.
- [8] DENG J, GUO J, XUE N, et al. Arcface: Additive angular margin loss for deep face recognition [C]// Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. IEEE, 2019: 4690-4699.
- [9] KIM Y, PARK W, ROH M C, et al. Groupface: Learning latent groups and constructing group-based representations for face recognition [C]// Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. IEEE, 2020: 5621-5630.
- [10] LIU Y, QU Y. Construction of a smart face recognition model for university libraries based on FaceNet-MMAR algorithm [J]. Plos One, 2024, 19(1): e0296656.
- [11] WEI Q, MU T, HAN G, et al. Face recognition based on improved faceNet model [C]// Proceedings of the Fifth Euro-China Conference on Intelligent Data Analysis and Applications. IEEE, 2019: 614-624.
- [12] HAN K, WANG Y, TIAN Q, et al. Ghostnet: More features from cheap operations [C]// Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. IEEE, 2020: 1580-1589.
- [13] CHEN L C, PAPANDREOU G, SCHROFF F, et al. Rethinking atrous convolution for semantic image segmentation [J]. arXiv preprint arXiv:1706.05587, 2019.
- [14] WOO S, PARK J, LEE J Y, et al. CBAM: Convolutional block attention module [C]// Proceedings of the European Conference on Computer Vision (ECCV). IEEE, 2018: 3-19.