



联合多视角特征的口罩遮挡人脸识别算法

苏雪平¹, 孙丹丹¹, 李云红¹, 姚丽娜¹, 任颖萱², 王 灿¹

(1. 西安工程大学 电子信息学院, 陕西 西安 710048; 2. 西安电子科技大学 计算机科学与技术学院, 陕西 西安 710126)

摘要 佩戴口罩是预防呼吸道传染病最经济、最有效、最实用的防护措施,但也因此会降低人脸识别准确率。提出联合多视角特征的口罩遮挡人脸识别算法,首先,利用 BoTNet 作为骨干特征提取网络,提高识别准确率。其次,引入人脸注意力增强网络(face attention augmentation model, FAAM),生成未被口罩遮挡人脸区域(眼睛、眉毛、额头)的蒙版图,再精准提取蒙版图区域的特征,提高人脸识别的性能。此外,设计联合损失函数 L_{face} 提高模型的收敛速度和性能。在包括大约 10 000 人的 50×10^4 张人脸图片的公共遮挡人脸数据集上进行实验,相比其他算法,提出的算法识别准确率有显著提升,与经典算法 FaceNet 相比提升了 13.9%。

关键词 遮挡人脸识别, BoTNet, 注意力增强

中图分类号: TP391.41 **DOI**: 10.16152/j.cnki.xdxh.2025-02-006

Mask-occluded face recognition algorithm with integrated multi-view features

SU Xueping¹, SUN Dandan¹, LI Yunhong¹, YAO Lina¹,
REN Yingxuan², WANG Can¹

(1. School of Electronics and Information, Xi'an Polytechnic University, Xi'an 710048, China;

2. School of College of Computer Science and Technology, Xidian University, Xi'an 710126, China)

Abstract Wearing a mask is the most economical, effective and practical protective measure to prevent respiratory infectious diseases, but it also reduces the accuracy of face recognition. Therefore, a face recognition algorithm with mask occlusion combined with multi-view features is proposed. Firstly, BoTNet is used as the backbone feature extraction network to improve the recognition accuracy. Secondly, face attention augmentation model (FAAM) is introduced to generate a non-masked map of face areas (eyes, eyebrows and forehead) that are not covered by masks, and then the features of the masked map areas are accurately extracted to improve the performance of face recognition. In addition, the joint loss function L_{face} is designed to improve the convergence speed and performance of the model. Experiments are carried out on a public occlusion face data set with 500 000 face images of about 10 000 people. Compared with other algorithms, the recognition accuracy of this method is significantly improved, and it is 13.9% higher than that of the classic algorithm FaceNet.

Keywords masked face recognition; BoTNet; attention augmentation

收稿日期: 2024-10-20

基金项目: 国家自然科学基金(61902301); 陕西省自然科学基金基础研究计划(2024JC-YBMS-455); 陕西省哲学社会科学专项青年项目(2024QN107); 陕西高校青年创新团队。

第一作者: 苏雪平, 博士, 副教授, 从事新闻人物标志、图像标注、人脸识别研究, yifeichongtian120@163.com。

佩戴口罩在减缓新流行病的传播方面起着重要作用,研究证明,佩戴口罩可以减少65%的疾病传播^[1]。此外,发表在《美国国家科学院院刊》(PNAS)上的一项研究指出,在预防 COVID-19^[2]传播和感染方面,佩戴口罩比社交距离和家庭隔离政策重要得多。出行佩戴口罩成为常态,而口罩遮挡下的人脸特征信息丢失,增加了人脸识别的难度^[3]。人脸识别是指基于光学人脸图像的身份识别与验证,而遮挡人脸识别是基于眼睛和额头区域来识别佩戴口罩的人脸。遮挡的人脸特征不完整,因此遮挡人脸识别已成为人脸识别领域具有挑战性的任务,众多学者开展遮挡人脸识别研究,主要包含传统算法和深度学习算法^[4]。

1)传统算法。①基于稀疏表示的算法^[5]有 SRC^[6]、LASRC^[7]、CRC-RLS^[8]、RSC^[9]、RASR^[10]等,该类模型对噪声具有鲁棒性,但要求所有人脸图像必须对齐,反之的稀疏性很难满足,不适合推广;②基于协同表示的算法^[11]有 CRC^[12]、Pro-CRC^[13]、PK-PCRC^[14]、CCRC^[15]等,该类模型的分类方式以全局表示为基础,计算速度快,但局部信息发生变化会导致性能下降;③基于特征分析的算法^[16]有 LFA^[17]、AMM^[18]、LS-ICA^[19]、PS-VM^[20-21]等,该类算法适用于样本较少的情况,计算量小。

2)深度学习算法。基于卷积神经网络的算法^[22]有 CNN^[23]、PCANet^[24]、结合 CNN 与 LBP^[25]等,该类模型有效识别遮挡人脸,但参数量大、计算能耗高,不适用于移动设备,且 CNN 模型提取遮挡人脸特征时,被遮挡部分会被嵌入到潜在空间的表示中^[26];②基于生成对抗网络的算法^[27]有 GAN^[28-29]、WGAN^[30]、WGAN-GP^[31]、LS-GAN^[32]、DCGAN^[33]等,该类模型生成更加清晰、真实的样本,但存在训练稳定性差、梯度消失、模式崩溃等问题;③基于注意力机制的算法^[34]有 AM^[35]、SANs^[36]、GATs^[37]、CBAM^[38]等,该类模型通过建立全局依赖关系,扩大图像的感受野,提取更多的区别性视觉信息^[39],但需海量训练数据,计算量庞大。其中,2021年提出的 BoTNet^[40]模型整合了多种自我注意力,用全局自我注意力代替空间卷积,在实例分割和对象检测上有明显的改善,且减少了参数量。

综上,传统的遮挡人脸识别算法多采用浅层结构提取人脸图像特征,易忽略图像细节信息,且口罩遮挡人脸大部分特征使方法性能下降。而

现有基于深度学习的遮挡人脸识别算法以深层结构提取特征,应用更广泛,但参数量大、计算能耗高、无法同时兼顾全局信息和局部信息。因此,本文提出联合多视角特征的口罩遮挡人脸识别算法,该算法通过引入注意力机制将网络的关注点聚焦在口罩遮挡外的人脸区域并提取特征,通过输入去除口罩后的人脸区域进一步提高佩戴口罩人脸识别性能。相比于传统方法可适应大样本的情况,并提取深层结构特征,与现有深度学习方法相比,本文所提方法可提取人脸区域全局特征和 non-masked map 的局部特征,最终提高识别准确率。

1 人脸识别算法

本文算法的总体技术框架如图1所示,主要步骤如下。①人脸检测与 non-masked map 生成:若输入为佩戴口罩的人脸,则确定口罩区域再生成 non-masked map,若输入为非佩戴口罩的人脸,则模拟口罩遮挡再生成 non-masked map。②特征提取:使用 BoTNet 提取人脸全局区域的特征,使用人脸注意力增强网络(face attention augmentation model, FAAM)提取 non-masked map 区域的人脸局部特征。③口罩遮挡人脸识别:使用联合损失函数加速模型的收敛速度,并识别口罩遮挡人脸。

1.1 人脸检测与 non-masked map 生成

1.1.1 人脸检测

Dlib 是一个强大而流行的开源库,用于图像处理和机器学习任务。在人脸检测方面,Dlib 使用了基于学习的方法,使用了形状预测器和人脸描述器识别和区分人脸。轻量级 Dlib 使用 HOG + SVM^[41]的方法检测人脸,首先,通过候选窗口滑动技术,将图像分成多个不同大小的窗口,用于检测可能存在的人脸区域。接下来,使用支持向量机(SVM)分类器判断每个窗口是否包含人脸。SVM 是一种常见的机器学习算法,通过训练样本学习特征和类别之间的关系对新的样本进行分类。如果窗口被分类为包含人脸,则使用形状预测器确定人脸的精确位置。形状预测器是一种人脸关键点检测算法,可以检测面部特征,如眼睛、鼻子、嘴巴等,从而为后续任务提供更多信息。该算法对姿势变化具有强鲁棒性,可检测正脸、旋转角度较小的人脸及不同尺度的人脸,甚至可检测

遮挡下的人脸,因此,本文使用 Dlib 检测人脸。若输入为佩戴口罩的人脸,使用检测精度高、速度快、模型泛化能力强的 Efficient-Yolov3^[42] 进行口罩检测,再使用 Dlib 检测人脸的 68 个关键坐标

点(见图 2),并进行归一化处理。若输入为非佩戴口罩的人脸,使用 Dlib 检测人脸的 68 个关键坐标点(见图 3),并进行归一化处理。

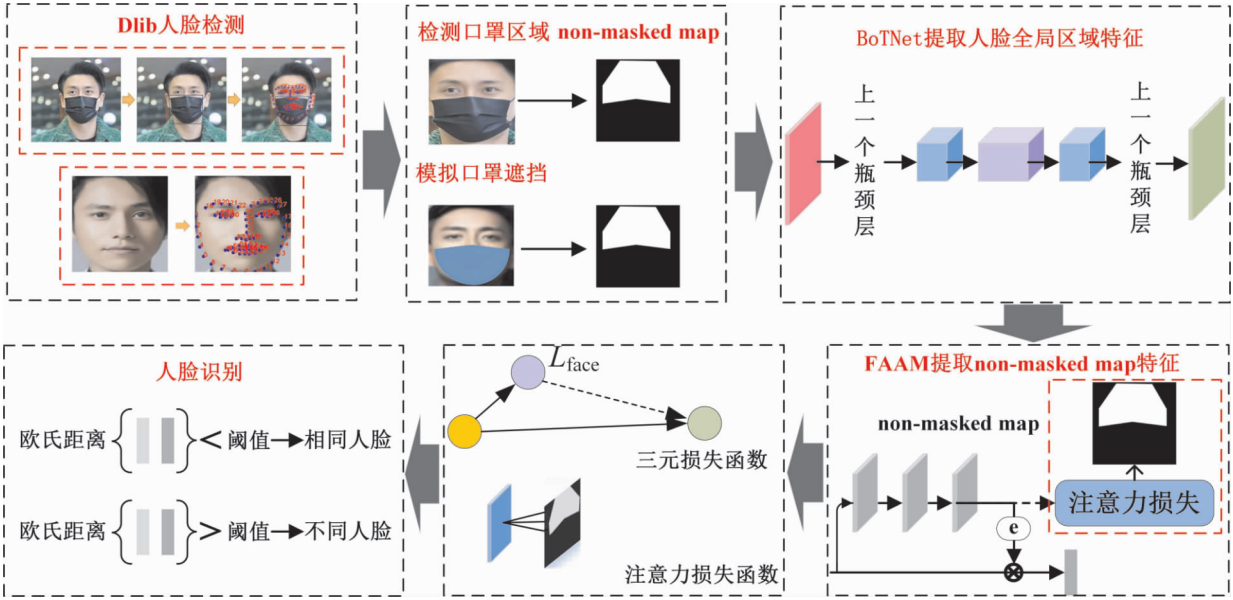


图 1 总体技术框架

Fig. 1 Overall technical framework diagram



图 2 口罩遮挡的人脸检测

Fig. 2 The detection of faces obstructed by masks

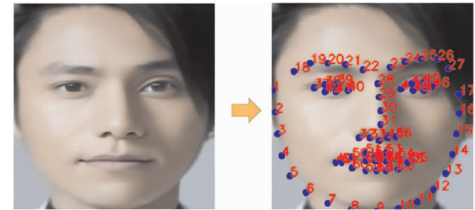


图 3 非口罩遮挡的人脸检测

Fig. 3 The detection of faces without obstruction by masks

1. 1. 2 生成 non-masked map

为了精准提取非口罩区域的人脸特征,本文生成非口罩区域的蒙版图(non-masked map)。输入为佩戴口罩的人脸,其生成步骤如下:①依次连接人脸关键坐标点 1、2、30、16、17,生成非口罩区域的下部边界;②利用眉毛关键点 20 和 25 计算出非口罩区域的上部边界;③依次连接坐标点 1、20、25 和 17,最终生成非口罩区域的 non-masked

map(主要包含眼睛、眉毛和额头周围区域),如图 4 所示。

输入为非佩戴口罩的人脸,使用 MaskThe-Face^[43] 模拟口罩遮挡,如图 5 所示。MaskThe-Face 是基于 Dlib 识别人脸倾斜度和应用口罩所需的人脸 68 个关键特征,再根据 68 个关键特征对模板口罩进行转换,该方法生成的口罩可以完美贴合面部。生成非口罩区域的 non-masked map 步骤与输入为佩戴口罩的人脸生成步骤一致。

1. 2 特征提取

1. 2. 1 BoTNet

FaceNet^[44] 使用端到端训练既简化了设置,又表明直接优化与手头任务相关的损失可以提高性能。FaceNet 采用残差网络(residual neural network, ResNet)作为主干网络,ResNet 解决网络深度增加导致的退化问题,实现高分类精度,但需要大容量的内存,训练时间长。文献[40]设计 BoT-Net,使用多头自我关注(multi-head self-attention, MHSA)替换 ResNet50 的 3 × 3 卷积,解决 Res-Net50 冗余大、计算能耗高、易过拟合的问题,且有效获取局部信息,参数少、速度快、效果好。因此,本文使用 BoTNet 作为主干网络,其网络结构如图 6 所示。

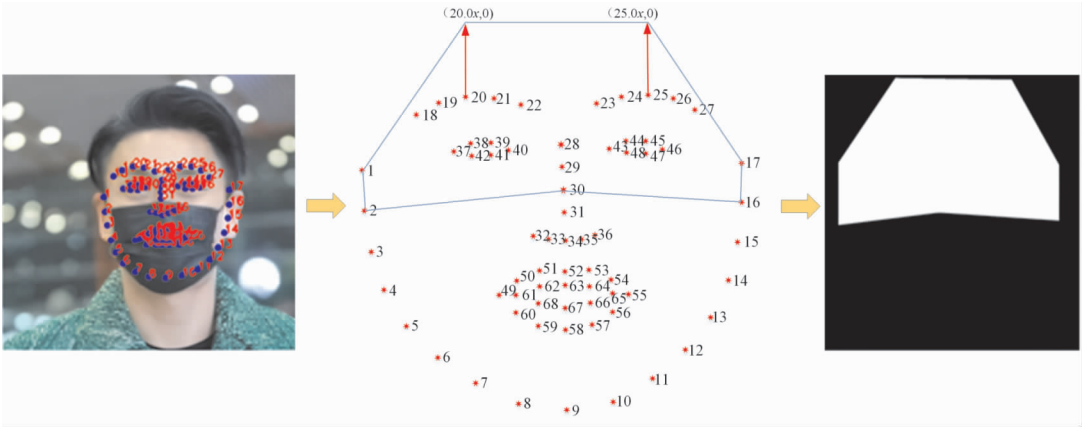


图 4 口罩遮挡人脸的 non-masked map 生成
Fig.4 Non-masked map generation of masked faces

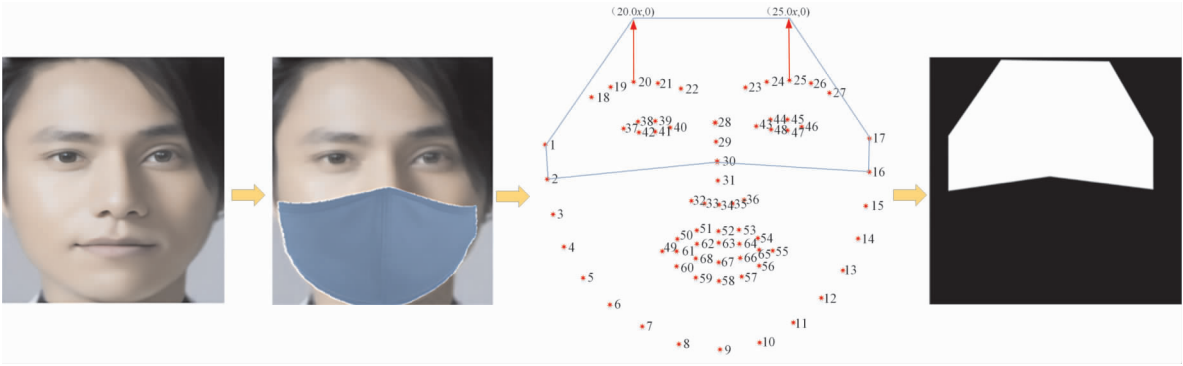


图 5 非口罩遮挡人脸的 non-masked map 生成
Fig.5 Non-masked map generation of non-masked faces

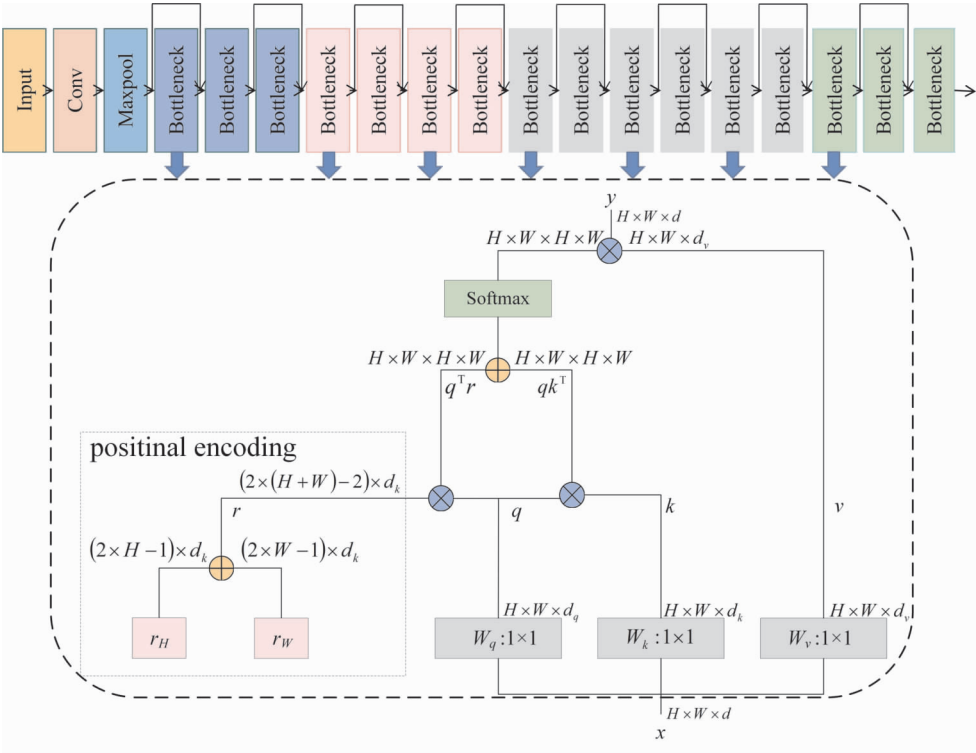


图 6 BoTNet 网络结构图^[40]
Fig.6 BoTNet architecture diagram

1.2.2 FAAM

现实场景中千人千面,但当人们佩戴口罩时,人脸部分具有高辨识度的五官(鼻子、嘴巴)被口罩遮挡,造成该区域特征的丢失。此外,同类型口罩区域的特征区分度小,降低了人脸识别的准确度。BoTNet 仅提取人脸区域的细粒度特征,未有效利用非口罩区域的人脸局部特征。因此,本文设计人脸注意力增强模块 FAAM,有效提取 non-masked map 区域的人脸局部特征,FAAM 结构如图 7 所示。FAAM 模块网络由 3 个卷积层级联组

成,对输入大小为 $W \times H \times d$ 的特征向量记为 y ,前 2 个卷积层使用步长为 1 的 512 个 3×3 尺寸的卷积核进行卷积操作,使用 ReLU 激活,得到 $W \times H \times 512$ 的特征图。第 3 个卷积层使用步长为 1 的 1 个 3×3 尺寸的卷积核进行降维,使用 Sigmoid 激活,将学习到的注意力特征图增强区域设置为 1,无用区域设置为 0。将降维特征记为 y_s 。进行 e 指数运算,将其结果与 y 进行点乘得到增强特征图 y' ,计算公式如下,

$$y' = y \otimes \exp(y_s)$$

(1)

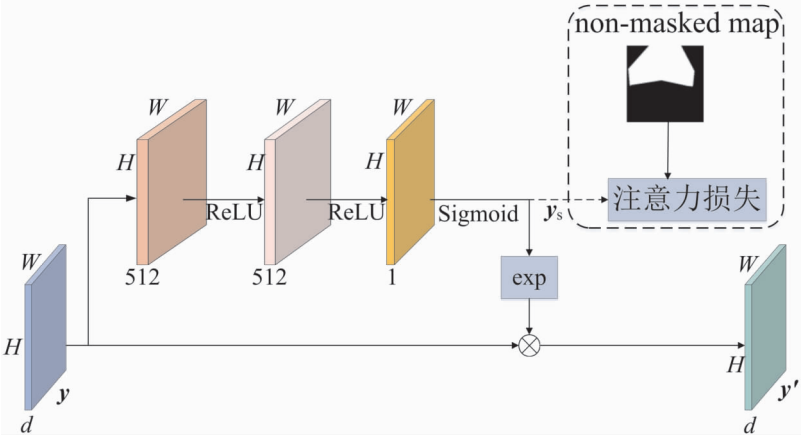


图 7 FAAM 模块结构图
Fig. 7 FAAM structure diagram

1.3 佩戴口罩人脸识别

1.3.1 联合损失函数

三元(triplet)损失函数能学习更好的细微特征,差异性较小的输入向量可学习到更好的表示,但其模型训练不稳定且收敛慢,需要不断调节参数,容易过拟合。而二元交叉熵(binary cross entropy, BCE)损失函数可衡量模型的预测准确性,让模型更快地收敛。于是本文使用联合损失函数用于人脸识别训练,帮助模型快速学习人脸特征。三元损失函数每次从人脸库中选取 3 张人脸样本图像,如图 8 所示,分别为①锚点(anchor, A),即参考人脸;②正例(positive, P),即与参考人脸同一类的人脸;③负例(negative, N),即与参考人脸不同类的人脸。

即 anchor 与 positive 和 anchor 与 negative 的距离大于设定阈值 m 则为同一类人脸;反之,则不是同一类人脸。三元损失函数 L_{triplet} 计算公式为

$$L_{\text{triplet}} = \sum_{i \in N} [\|f(x_i^a) - f(x_i^p)\|_2^2 - \|f(x_i^a) - f(x_i^n)\|_2^2 + \omega]$$

(2)

式中: N 为三元组总数; $f(x_i^a)$ 、 $f(x_i^p)$ 和 $f(x_i^n)$ 分别代表特征网络提取到的人脸锚点、人脸正例和人脸负例的特征向量。在创建人脸三元组样本时,可能会将同一类人的 3 张样本放在一起计算损失函数,造成计算到的损失函数为 0,为了避免此情况,本文引入权重系数 ω 判别是否为同一人, ω 值设置为 0.5。如表 1 所示, ω 的取值过大或过小都会影响最终的识别率。

表 1 不同 ω 下的识别率
Tab. 1 Recognition accuracy under different ω
单位: %

识别 指标	ω				
	0.00	0.25	0.50	0.75	1.00
Accuracy	51.6	62.5	74.7	63.2	62.6
Precision	59.6	66.5	72.5	62.5	60.7
Recall	52.6	55.6	81.1	63.8	74.3

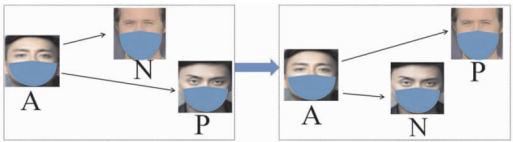


图 8 三元组选取示例
Fig. 8 An example of triplet selection

选取欧氏距离作为不同人脸图像的相似度,

同理, BCE Loss 每次选取 3 张人脸图像, 通过 BoTNet 提取到的热力特征图以及输入 FAAM 的 non-masked map 计算关键像素点交叉熵损失, 最后取 3 者均值作为 $L_{\text{attention}}$ 。其计算公式为

$$L_{\text{attention}} = \sum_{i \in N} \left\{ \frac{1}{3} [\text{BCE}(x_i^a) + \text{BCE}(x_i^p) + \text{BCE}(x_i^n)] \right\} \quad (3)$$

$$\text{BCE}(x) = - \sum_i^M \sum_j^M [m_{ij} \log(h_{ij}) + (1 - m_{ij}) \log(1 - h_{ij})] \quad (4)$$

式中: $\text{BCE}(x)$ 代表交叉熵损失函数; m_{ij} 为 non-masked map 的像素点值, 将 non-masked map 归一化成注意力特征热力图大小, 则口罩区域像素点为 0, 去除口罩区域后的人脸区域像素点为 1; h_{ij} 为提取到的注意力特征值; M 为特征图尺寸。通过式(3)和(4)计算后可去除口罩区域和人脸背景等无关特征, 定位并增强人脸非口罩区域特征。

综上, 本文最终的损失函数可表示为

$$L_{\text{face}} = L_{\text{triplet}} + L_{\text{attention}} \quad (5)$$

此外, 在训练阶段, 若样本对的参照人脸与正例极度相似或参照人脸与负例差异较大, 则会导致模型难以学习到人脸具有区分性的特征。因此, 本文将困难样本进行筛选后, 再次训练模型, 使模型学习到更加复杂且具有区分度的人脸特征。筛选的样本对主要包含 2 类: 参照人脸与正例差异较大的样本对和参照人脸与负例较为相似的样本对。其计算公式为

$$\|f(x_i^a) - f(x_i^n)\|_2^2 - \|f(x_i^a) - f(x_i^p)\|_2^2 < \omega \quad (6)$$

式中: ω 为权重系数, 本文取值为 0.5; 其他变量含义同式(2)所示。通过条件筛选后的样本对可避免模型出现梯度不下降的情况, 使得模型更稳定并且快速收敛。

1.3.2 识别

FaceNet 只需要裁剪脸部区域, 作为模型输入并提取特征后直接计算距离, 简单有效且对光照和姿态具有不变性。因此, 本文使用改进的 FaceNet 进行遮挡人脸识别, 步骤如下: ①输入 1 张人脸图像; ②通过上述 BoTNet 和 FAAM 网络分别提取人脸区域特征和 non-masked map 区域的局部特征; ③ L_2 标准化特征向量; ④与数据库中人脸图像进行对比并获得预测结果。

2 实验与分析

2.1 数据集

本文采用 CASIA-WebFace-mask^[45] 和 VGG-Face2^[46] 数据集作为训练集。其中, CASIA-WebFace-mask 数据集包含 445 466 张戴口罩的人脸图像, 所涉及到的口罩类型有 Surgical(白色医用外科口罩)、Surgical Blue(蓝色医用外科口罩)、N95、KN95 以及 Cloth(黑色布质口罩), 口罩的分布类型都是均匀分布随机生成的, 如图 9 所示。VGGFace2 数据集包含不同姿态、光照背景下约 8 000 人的 300×10^4 张人脸图像, 如图 10 所示。本文使用 Mask The Face 在 VGGFace2 数据集基础上生成模拟口罩遮挡的数据集, 如图 11 所示。在生成所需数据集的同时也对数据集中像素较低和人脸偏移角度过大的图像进行数据清洗, 保存并生成高质量的人脸数据集, 实现了数据增强, 并增加训练数据和提高模型泛化能力。



图 9 CASIA-WebFace-mask 数据集

Fig.9 CASIA-WebFace-mask dataset



图 10 VGGFace2 数据集

Fig.10 VGGFace2 dataset



图 11 模拟口罩遮挡数据集

Fig.11 Simulated mask occlusion dataset

测试集采用 SMFRD 数据集^[47], 该数据集包含 10 000 人的 50×10^4 张人脸图片, 它使用 GAN 网络对 LFW 数据集中的图像生成戴口罩的人脸图像, 如图 12 所示。



图 12 SMFRD 数据集

Fig.12 SMFRD dataset

2.2 评价指标

人脸识别评估指标使用正确率 (Accuracy, 式

中简记 A)、精度 (Precision, 式中简记 P)、召回率 (Recall, 式中简记 R)。其计算公式如下:

$$A = \frac{TP + TN}{(TP + FP) + (TN + FN)} \tag{7}$$

$$P = \frac{TP}{TP + FP} \tag{8}$$

$$R = \frac{TP}{TP + FN} \tag{9}$$

式中: TP 表示预测和实际均为真; TN 表示预测和实际均为假; FP 表示预测为真, 但是实际为假; FN 表示预测为假, 但是实际为真。

2.3 实验设置

本文在 NVIDIA GeForce RTX 3090 Ti 上训练该模型, 将数据集分为训练集、验证集和测试集, 占比分别为 70%、10% 和 20%。通过实验发现训练 145 个周期时, 受试者工作特征曲线 (receiver operating characteristic curve, ROC) 值达到最大, 为防止过拟合, 本文设置 epoch 为 145, 不同 epoch 下的 ROC 数据如表 2 所示。

表 2 不同 epoch 下 ROC 数据

Tab. 2 ROC data under different epochs

epoch	130	135	140	145	150
ROC	0.784	0.788	0.891	0.893	0.892

每次输入模拟人脸口罩遮挡的图片和 mask 人脸图片, 输入大小为 $256 \times 256 \times 3$, 训练共生成 10 万对三元组人脸样本, 从样本中分别取 30 对训练集人脸和测试集人脸用于训练。为了让本文设计的网络损失达到较快的收敛速度, 对于不同周期选取不同学习率, 学习率 r_1 的选择如式 (10) 所示,

$$r_1 = \begin{cases} 0.125 & 0 < \text{epoch} < 30 \\ 0.062 & 30 \leq \text{epoch} < 60 \\ 0.015 & 60 \leq \text{epoch} < 90 \\ 0.003 & 90 \leq \text{epoch} < 120 \\ 0.000 & 1 \leq \text{epoch} \leq 120 \end{cases} \tag{10}$$

2.4 结果及分析

为验证本文提出算法的有效性, 与算法 FaceNet、CosFace^[48]、ArcFace^[49]、PDSN^[50] 进行比较。其中, FaceNet 是经典人脸识别算法, 实现了端到端的训练, 本文算法是在 FaceNet 上进行的改进; ArcFace 是在继 Softmax Loss、Center Loss、A-Softmax Loss、Cosine Margin Loss 之后有更加卓越的表现; 与 ArcFace 类似, CosFace 引入一个余弦 Margin 可以在角度空间中增大决策 Margin, 因此

可以实现最小化类内差异, 最大化类间差异; PD-SN 与其他遮挡人脸识别算法不同, 该算法是利用遮挡和无遮挡人脸对的顶部卷积特征之间的差异建立字典进行识别的。不同算法的口罩遮挡人脸识别在 SMFRD 数据集上的实验结果如图 13、14 所示, 可以看出, 输入模拟口罩遮挡的图像后会进一步提升识别率。

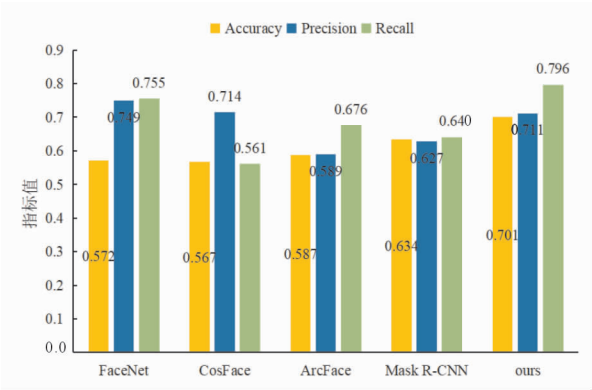


图 13 训练时输入非口罩遮挡图像的实验结果

Fig. 13 Experimental results after input non-masked images training

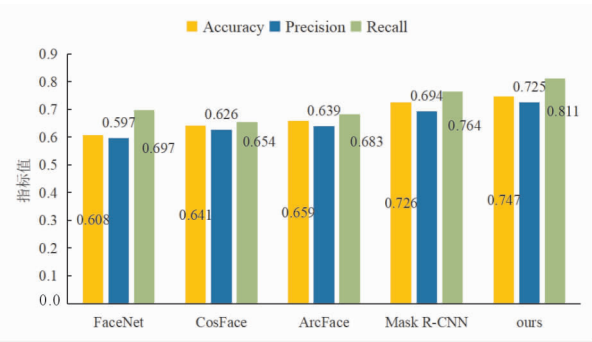


图 14 训练时输入模拟口罩遮挡图像的实验结果

Fig. 14 Experimental results after input simulated mask occlusion images training

对于口罩遮挡人脸无论是否输入模拟口罩遮挡的人脸, 本文所提算法的识别率均优于其他模型, 识别率可达 74.7%, 相比于其他 4 种算法的识别率, 本文算法分别提高了 13.9%、10.6%、8.8% 和 2.1%。其原因是本文提出的算法突出人脸信息较多的眼睛周围区域特征, 可以准确提取特征。综上, 本文算法可以达到遮挡人脸识别的最佳精度。

采用类似于控制变量法的方法深度剖析本文所提改进策略对口罩遮挡人脸识别的性能影响。同一张人脸的特征图可视化结果如图 15 所示。

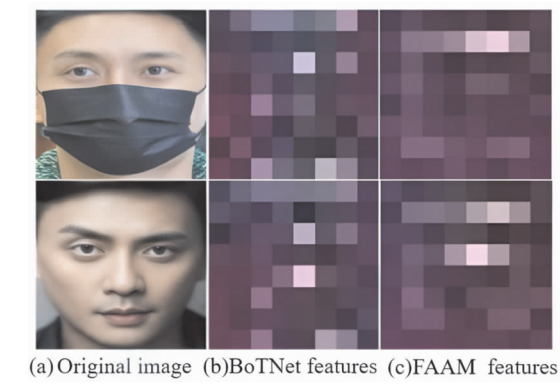


图 15 特征可视化结果
Fig. 15 Feature map visualization

图 15(a)为输入的人脸图像,15(b)为 BoTNet 网络提取的特征可视化结果,15(c)为 FAAM 网络提取的特征可视化结果。其中,第 1 行为输入的佩戴口罩人脸图像及其特征图,第 2 行为输入的不佩戴口罩人脸图像及其特征图。可以看出,BoTNet 网络提取人脸区域的细粒度特征,而 FAAM 精准定位眼部区域,将网络的关注度尽可能高地集中在眼睛、眉毛和额头周围区域的特征。

表 3 训练时输入非口罩遮挡的图像在 SMFRD 数据集上的消融实验结果

Tab. 3 Ablation experiment results on the SMFRD dataset when training with images not occluded by masks

ResNet50	BoTNet	FAAM	L_{face}	Accuracy/%	Precision/%	Recall/%
✓				57.2	74.9	75.5
	✓			58.6	63.6	61.5
✓		✓		65.4	71.2	69.3
✓			✓	59.3	66.6	71.7
	✓	✓	✓	70.1	71.1	79.6

表 4 训练时输入模拟口罩遮挡的图像在 SMFRD 数据集上的消融实验结果

Tab. 4 Ablation experiment results on the SMFRD dataset when training with images simulating mask occlusion

ResNet50	BoTNet	FAAM	L_{face}	Accuracy/%	Precision/%	Recall/%
✓				70.3	70.8	75.5
	✓			71.5	69.4	72.4
✓		✓		72.8	71.8	73.2
✓			✓	70.9	71.9	69.9
	✓	✓	✓	74.7	72.5	81.1

综上,BoTNet 帮助网络增强眼睛周围的人脸特征,提取人脸区域细粒度特征,FAAM 则精准定位人脸的非口罩遮挡区域,帮助网络提取 non-

在 SMFRD 数据集上的消融实验结果如表 3、4 所示。由表 3 可得,在同样选取 ResNet50 作为特征提取网络时,若训练时不输入模拟口罩遮挡人脸图像识别率仅为 57.2%,而训练时输入模拟口罩遮挡人脸图像识别率却可达到 70.3%,两者相差 13.1%。由表 4 可得,在选取 ResNet50 作为主特征提取网络时识别率仅为 70.3%,将 ResNet50 模型改为 BoTNet 后识别率提高了 1.2%,由此可以验证本文通过注意力机制增强非口罩区域的人脸特征的有效性。在 ResNet50 的基础上结合 FAAM 网络相比于仅使用 ResNet50 识别准确率提高了 2.5%,其原因是 FAAM 输入 non-masked map 后将口罩遮挡区域的特征丢弃,并通过计算与人脸原始输入照片的损失值精准定位去除口罩外的人脸面部轮廓,将网络的特征聚焦于非口罩区域人脸。使用联合损失函数后准确率提升了 0.6%,可见对于困难样本模型能学习到更加复杂且具有区分度的人脸特征。以上结果证明本文提出的每个改进均有可行性。

masked map 区域的局部特征, L_{face} 提高模型的收敛速度和性能,本文提出的算法对于口罩遮挡的人脸能够达到极佳的识别性能。

3 结语

本文提出联合多视角特征的口罩遮挡人脸识别算法,首先,将 BoTNet 网络作为主干网络,在每个残差卷积块中添加混合多头注意力增强人脸非口罩区域特征、提取人脸区域细粒度特征;其次,在 BoTNet 最后一个卷积块后添加 FAAM 进行人脸非口罩区域的定位、提取 non-masked map 区域的局部特征;最后,使用联合损失函数 L_{face} 提高模型的收敛速度和性能。通过实验验证本文算法能够有效识别口罩遮挡的人脸,为口罩遮挡人脸识别提供了有效的解决方案。但现实场景中口罩类别大不相同、对于人脸遮挡的范围也存在差异,本文仅模拟生成了一种类别的口罩遮挡,下一步工作将模拟生成不同类型的口罩区域,采用真实的口罩遮挡人脸进行测试,寻找口罩遮挡人脸识别的最优解决办法。遮挡环境下的人脸识别准确率相较于无遮挡环境仍相对偏低,因此还需研究新技术提高遮挡环境下的人脸识别准确率。

参考文献

- [1] ABBOAH-OFFEI M, SALIFU Y, ADEWALE B, et al. A rapid review of the use of face mask in preventing the spread of COVID-19[J]. International Journal of Nursing Studies Advances, 2021, 3: 100013.
- [2] HOWARD J, HUANG A, LI Z Y, et al. An evidence review of face masks against COVID-19[J]. Proceedings of the National Academy of Sciences of the United States of America, 2021, 118(4): e2014564118.
- [3] WALVEKAR S, SHINDE S. Efficient medical image segmentation of COVID-19 chest CT images based on deep learning techniques[C]//2021 International Conference on Emerging Smart Computing and Informatics (ESCI). Pune: IEEE, 2021: 203-206.
- [4] 刘瑞明, 徐春融, 周韬. 局部遮挡的人脸识别方法研究综述[J]. 江苏海洋大学学报(自然科学版), 2022, 31(3): 63-71.
LIU R M, XU C R, ZHOU T. Overview of face recognition methods with partial occlusion[J]. Journal of Jiangsu Ocean University (Natural Science Edition), 2022, 31(3): 63-71.
- [5] 刘显梅, 赵洪坡. 基于 Retinex 算法的光照变化条件下人脸识别研究[J]. 激光杂志, 2023, 44(7): 218-222.
LIU X M, ZHAO H P. Research on face recognition under illumination change based on Retinex algorithm[J]. Laser Journal, 2023, 44(7): 218-222.
- [6] 胡思佳, 赵志诚. 基于改进 SRC 的局部遮挡人脸识别方法[J]. 太原科技大学学报, 2024, 45(1): 7-12.
HU S J, ZHAO Z C. Local occlusion face recognition method based on improved SRC [J]. Journal of Taiyuan University of Science and Technology, 2024, 45(1): 7-12.
- [7] ORTIZ E G, BECKER B C. Face recognition for web-scale datasets[J]. Computer Vision and Image Understanding, 2014, 118: 153-170.
- [8] LEI Z, MENG Y, FENG X C. Sparse representation or collaborative representation: Which helps face recognition? [C]//2011 International Conference on Computer Vision. Barcelona: IEEE, 2011: 471-478.
- [9] 李小薪, 梁荣华. 有遮挡人脸识别综述: 从子空间回归到深度学习[J]. 计算机学报, 2018, 41(1): 177-207.
LI X X, LIANG R H. Overview of occluded face recognition: From subspace regression to deep learning[J]. Chinese Journal of Computers, 2018, 41(1): 177-207.
- [10] WAGNER A, WRIGHT J, GANESH A, et al. Toward a practical face recognition system: Robust alignment and illumination by sparse representation[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2012, 34(2): 372-386.
- [11] 雷燕, 李杰, 董博, 等. 基于深度学习特征的多模态人脸快速识别研究[J]. 电子设计工程, 2024, 32(3): 181-184.
LEI Y, LI J, DONG B, et al. Research on fast multimodal face recognition based on deep learning features[J]. Electronic Design Engineering, 2024, 32(3): 181-184.
- [12] PENG Y L, LI L J, LIU S G, et al. A new virtual samples-based CRC method for face recognition[J]. Neural Processing Letters, 2018, 48(1): 313-327.
- [13] CAI S J, ZHANG L, ZUO W M, et al. A probabilistic collaborative representation based approach for pattern classification[C]//2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Las Vegas, NV: IEEE, 2016: 2950-2959.
- [14] LAN R S, ZHOU Y C, LIU Z B, et al. Prior knowledge-based probabilistic collaborative representation for visual recognition[J]. IEEE Transactions on Cybernetics, 2020, 50(4): 1498-1508.
- [15] YUAN H L, LI X C, XU F Y, et al. A collaborative-

- competitive representation based classifier model[J]. *Neurocomputing*, 2018, 275: 627-635.
- [16] 刘加聪, 钟桂凤, 盛儒好. 基于 OpenCV 和 MTCNN 算法的遮蔽物影响下的人脸识别研究[J]. *电脑编程技巧与维护*, 2023 (11): 120-124.
- LIU J C, ZHONG G F, SHENG R Y. Research on face recognition under the influence of masks based on OpenCV and MTCNN algorithms [J]. *Computer Programming Skills and Maintenance*, 2023 (11): 120-124.
- [17] 张宣妮, 马秀霞, 鲁方莹, 等. 智慧小区的人脸识别门禁系统[J]. *信息系统工程*, 2021(4): 95-96.
- ZHANG X N, MA X X, LU F Y, et al. Face recognition access control system in smart community [J]. *Information System Engineering*, 2021(4): 95-96.
- [18] 靳若华, 白凡, 张洪豪. 基于多任务学习的特征融合跨年龄人脸识别研究[J]. *天津理工大学学报*, 2024, 40(1): 84-91.
- JIN R H, BAI F, ZHANG H H. Cross-age face recognition based on multi-task learning [J]. *Journal of Tianjin University of Technology*, 2024, 40(1): 84-91.
- [19] 雷耀花, 刘东宏. 基于 PCA + ICA 人脸识别技术的考勤系统设计[J]. *电脑与电信*, 2017(11): 55-57.
- LEI Y H, LIU D H. Design of attendance system based on PCA + ICA face recognition technology [J]. *Computer and Telecommunications*, 2017 (11): 55-57.
- [20] JIA H J, MARTINEZ A M. Support vector machines in face recognition with occlusions[C] // 2009 IEEE Conference on Computer Vision and Pattern Recognition. Miami, FL: IEEE, 2009: 136-141.
- [21] 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.
- LI Y H, LIU X R, XIE R R, et al. Low-resolution face recognition based on super-resolution reconstruction and common feature subspace [J]. *Journal of Northwest University (Natural Science Edition)*, 2023, 53(2): 241-247.
- [22] 罗亿. 浅谈部分遮挡下的人脸识别技术研究[J]. *中国新通信*, 2023, 25(2): 67-68.
- LUO Y. On the research of face recognition technology under partial occlusion [J]. *China New Communication*, 2023, 25(2): 67-68.
- [23] SIRADJUDDIN I A, REYNALDI, MUNTASA A. Faster region-based convolutional neural network for mask face detection[C] // 2021 5th International Conference on Informatics and Computational Sciences (ICICoS). Semarang: IEEE, 2021: 282-286.
- [24] CHENG X Y, ZHANG L. Research on occluded face recognition method based on deep learning[J]. *Electronic Technology*, 2022, 35(1): 35-39.
- [25] VU H N, NGUYEN M H, PHAM C. Masked face recognition with convolutional neural networks and local binary patterns[J]. *Applied Intelligence*, 2022, 52(5): 5497-5512.
- [26] ZENG D, VELDHUIS R, SPREEUWERS L, et al. Occlusion-invariant face recognition using simultaneous segmentation[J]. *IET Biometrics*, 2021, 10(6): 679-691.
- [27] 蒋文豪. 基于生成对抗网络的多姿态人脸识别算法[J]. *信息技术与信息化*, 2024(2): 188-191.
- JIANG W H. Multi-pose face recognition algorithm based on generated confrontation network [J]. *Information Technology and Informatization*, 2024 (2): 188-191.
- [28] 王跃进, 耿国华, 郭沛瑶, 等. 基于对抗生成网络的复原面貌真实感处理[J]. *西北大学学报(自然科学版)*, 2021, 51(5): 742-749.
- WANG Y J, GENG G H, GUO P Y, et al. Realistic processing of restored faces based on generative adversarial networks [J]. *Journal of Northwest University (Natural Science Edition)*, 2021, 51(5): 742-749.
- [29] DONG X B, MIAO Z H, MA L, et al. Reconstruct face from features based on genetic algorithm using GAN generator as a distribution constraint[J]. *Computers & Security*, 2023, 125: 103026.
- [30] ADLER J, LUNZ S. Banach Wasserstein Gan[EB/OL]. (2018-06-18) [2024-08-20]. <https://arxiv.org/abs/1806.06621>.
- [31] XU J B, XU D G, WAN K, et al. Alleviating sample imbalance in water quality assessment using the VAE-WGAN-GP model[J]. *Water Science & Technology*, 2023, 88(11): 2762-2778.
- [32] WANG C G, CAO Y, ZHANG S, et al. A reconstruction method for missing data in power system measurement based on LSGAN[J]. *Frontiers in Energy Research*, 2021, 9: 651807.
- [33] WU Q F, CHEN Y P, MENG J. DCGAN-based data augmentation for tomato leaf disease identification[J]. *IEEE Access*, 2020, 8: 98716-98728.
- [34] QIAN H J, ZHANG P P, JI S J, et al. Improving representation consistency with pairwise loss for masked face recognition[C] // 2021 IEEE/CVF International

- Conference on Computer Vision Workshops (IC-CVW). Montreal, BC: IEEE, 2021: 1462-1467.
- [35] GUO B W, LI X Y, YANG M M, et al. A robust and lightweight deep attention multiple instance learning algorithm for predicting genetic alterations[J]. Computerized Medical Imaging and Graphics, 2023, 105: 102189.
- [36] JIANG N, GAO L, DUAN F X, et al. SAN: Attention-based social aggregation neural networks for recommendation system[J]. International Journal of Intelligent Systems, 2022, 37(6): 3373-3393.
- [37] XIAO C, IMEL E A, PHAM N, et al. Patient-GAT: Sarcopenia prediction using multi-modal data fusion and weighted graph attention networks[J]. Proceedings of the Symposium on Applied Computing Symposium on Applied Computing, 2023, 2023: 614-617.
- [38] WOO S, PARK J, LEE J Y, et al. CBAM: Convolutional block attention module[C] // Computer Vision ECCV 2018. Cham: Springer International Publishing, 2018: 3-19.
- [39] WANG W X, HU H F. Image captioning using region-based attention joint with time-varying attention[J]. Neural Processing Letters, 2019, 50(1): 1005-1017.
- [40] SRINIVAS A, LIN T Y, PARMAR N, et al. Bottleneck transformers for visual recognition[C] // 2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Nashville, TN: IEEE, 2021: 16514-16524.
- [41] JACOB M P. Comparison of popular face detection and recognition techniques[J]. International Research Journal of Modernization in Engineering Technology and Science, 2021: 2582-5208.
- [42] SU X P, GAO M, REN J, et al. Face mask detection and classification via deep transfer learning[J]. Multimedia Tools and Applications, 2022, 81(3): 4475-4494.
- [43] HUANG Q Y, TANG C Y, ZHANG T S. Periocular biometric recognition for masked faces[J]. Wuhan University Journal of Natural Sciences, 2023, 28(2): 141-149.
- [44] 王传传, 高婕. 基于改进 FaceNet 算法的人脸智能识别方法[J]. 工业控制计算机, 2024, 37(2): 126-128.
- WANG C C, GAO J. Intelligent face recognition method based on improved FaceNet algorithm[J]. Industrial Control Computer, 2024, 37(2): 126-128.
- [45] DING F F, PENG P X, HUANG Y R, et al. Masked face recognition with latent part detection[C] // Proceedings of the 28th ACM International Conference on Multimedia. Seattle WA: ACM, 2020: 2281-2289.
- [46] CAO Q, SHEN L, XIE W D, et al. VGGFace2: A dataset for recognising faces across pose and age[C] // 2018 13th IEEE International Conference on Automatic Face & Gesture Recognition (FG 2018). Xi'an: IEEE, 2018: 67-74.
- [47] WANG Z Y, WANG G C, HUANG B J, et al. Masked face recognition dataset and application[EB/OL]. (2020-03-20) [2024-09-10]. <https://arxiv.org/abs/2003.09093>.
- [48] WANG H, WANG Y T, ZHOU Z, et al. CosFace: Large margin cosine loss for deep face recognition[C] // 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Salt Lake City: IEEE, 2018: 5265-5274.
- [49] DENG J K, GUO J, XUE N N, et al. ArcFace: Additive angular margin loss for deep face recognition[C] // 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Long Beach, CA: IEEE, 2019: 4690-4699.
- [50] SONG L X, GONG D H, LI Z F, et al. Occlusion robust face recognition based on mask learning with pairwise differential Siamese network[C] // 2019 IEEE/CVF International Conference on Computer Vision (ICCV). Seoul: IEEE, 2019: 773-782.

(编辑 李 静)