

细粒度深度特征掩码估计的遮挡人脸识别算法

王富平, 王定莎, 李藕, 刘卫华, 刘鸿玮

(西安邮电大学通信与信息工程学院, 710121, 西安)

摘要: 针对人脸遮挡产生面部结构信息丢失, 从而导致人脸识别准确率降低的问题, 提出了一种细粒度深度特征掩码估计的遮挡人脸识别算法。首先, 将人脸图像输入特征金字塔网络中, 从而得到多尺度深度语义特征; 其次, 将从特征金字塔网络提取的特征经过空洞卷积处理后, 与 MobileNetV3 网络提取的精细浅层特征进行融合, 并以像素级二值掩码为标签训练网络以获得细粒度特征掩码; 进而, 利用该深度特征掩码与深层特征相乘, 以抑制由遮挡产生的干扰特征, 获得更准确的人脸表征; 最后, 采用余弦损失和掩码估计损失联合训练网络, 提高遮挡人脸识别算法的性能。在 LFW 数据集基础上创建了口罩、围巾和中心遮挡 3 种类型的人脸遮挡数据集, 实验结果表明: 在不同的数据集上, 所提算法与现有算法相比均具有更高的识别准确率, 并在不同类型遮挡情况下均能获得十分稳定的人脸识别结果; 所提算法在数据集 LFW 和 LFW 口罩遮挡上的识别准确率分别达到了 99.38% 和 98.42%, 在数据集 LFW 围巾遮挡和 LFW 中心遮挡上的识别准确率分别达到了 98.72% 和 98.65%, 均优于对比算法。

关键词: 人脸识别; 细粒度; 掩码估计; 遮挡; 特征掩码

中图分类号: TP391 **文献标志码:** A

DOI: 10.7652/xjtuxb202502017 **文章编号:** 0253-987X(2025)02-0170-10

Occluded Face Recognition Algorithm Based on Fine-Grained Deep Feature Mask Estimation

WANG Fuping, WANG Dingsha, LI Ou, LIU Weihua, LIU Hongwei

(School of Communication and Information Engineering, Xi'an University of Posts & Telecommunications, Xi'an 710121, China)

Abstract: To solve the problem of accuracy decrease in facial recognition caused by the loss of facial structural information in the case of facial occlusion, an occluded face recognition algorithm based on fine-grained deep feature mask estimation was proposed. Firstly, the face image was fed into feature pyramid network (FPN) to obtain multi-scale deep semantic features. Next, the features extracted from the FPN were processed by the atrous convolution and fused with shallow features extracted using MobileNetV3; a pixel-wise binary mask was used as label to train the network to obtain the fine-grained deep feature mask; this deep feature mask was multiplied with the deep features to suppress the corrupted feature produced by occlusion and to obtain the better face representation. Finally, CosFace loss and mask estimation loss were jointly used to train the network to improve the performance of occluded face recognition algorithm. Three face occlusion

datasets with mask, scarf, and center occlusion were created based on the LFW dataset, respectively. The experimental results show that on four different datasets, the proposed algorithm was more accurate than existing algorithms and obtained very stable face recognition results under different occlusion situations. The recognition accuracy on LFW and LFW-mask-occlusion datasets reached 99.38% and 98.42%, respectively and that on LFW-scarf-occlusion and LFW-center-occlusion datasets reached 98.72% and 98.65%, respectively, outperforming the algorithms compared.

Keywords: face recognition; fine-grained; mask estimation; occluded; feature mask

人脸识别技术在多个领域得到广泛应用,如公安刑侦、公共安全、自动驾驶等^[1-4]。同时,基于深度卷积神经网络的人脸识别技术在非遮挡人脸识别方向取得十分显著的效果。然而,在日常生活当中,人们常常会配戴墨镜、帽子、围巾、口罩等遮挡物,这些遮挡物导致了人脸图像中的关键特征丢失,从而严重影响了人脸识别的准确性^[5]。

目前,遮挡人脸识别方法主要分为两大类:传统的人脸识别和基于深度学习的人脸识别方法。以下两类经典算法为传统的人脸识别算法:Wright等^[6]较早提出了稀疏表示分类模型,并将其应用于人脸识别领域,该算法将人脸图像表示为稀疏线性组合,并且能够在识别过程中对噪声表现出较强的鲁棒性,但人脸图像中的遮挡部分往往表现出不规则且随机的特征,使得部分算法容易忽略图像遮挡区域的误差分布;Zhou等^[7]提出了马尔可夫随机场模型,这种方法可以有效地识别出遮挡区域,并将其排除在稀疏表示之外。

随着人们对深度学习的不断探索和研究,提出了基于深度学习的人脸识别方法,这些方法主要分为两类:基于遮挡区域恢复和基于遮挡区域抑制的人脸识别方法^[8]。

基于遮挡区域恢复的人脸识别方法是通过恢复被遮挡的人脸部分来提升人脸识别性能。此类方法的核心在于提出一种高质量的遮挡人脸图像修复算法。Chen等^[9]提出了一种基于生成对抗网络的遮挡人脸识别模型,该模型通过利用在无遮挡图像上预训练的生成对抗网络,自动检测遮挡区域,进而修复出该区域的人脸图像像素,但由于生成对抗网络生成的结果的不确定性,导致仍会出现错误识别。为了解决人脸遮挡识别准确率较低的问题,Zhao等^[10]提出了鲁棒长短期记忆网络自编码模型,即使在野外遮挡情况下,仍然可以有效识别出人物身份,但该模型长短期记忆网络计算复杂度高,并需要大量的数据集进行训练。

基于遮挡区域抑制的人脸识别算法通过抑制遮

挡区域的干扰特征,利用不含干扰区域的特征进行识别^[8]。Song等^[11]提出了成对微分孪生网络(pairwise differential siamese network, PDSN)。该网络通过利用有干扰和无干扰人脸对特征之间的差异生成掩码字典,利用该字典生成遮挡部分的特征丢弃掩码,从而去除遮挡区域对人脸图像特征的干扰。但是,由于需要训练 K^2 个掩码生成模型,导致模型的计算量较大。Qiu等^[12]提出了一种基于端到端深度神经网络的人脸识别方法(face recognition with occlusion masks, FROM),该方法利用解码器从图像金字塔特征中提取遮挡区域的掩码信息,使其与深度卷积特征相乘以去除遮挡区域的干扰特征,从而获得更高的准确率。但是,由于该模型预测出的特征掩码均为矩形区域,使得在预测过程中无法准确识别出所有的掩码特征,同时,难以适应任意形状的随机遮挡。Hao等^[13]提出了人脸特征修正网络(face feature rectification network, FFR-Net),该网络通过引入修正模块,对提取的人脸特征进行处理,从而提高在有遮挡和无遮挡条件下人脸图像的识别性能。

上述算法预测出的特征掩码均为矩形或块状区域,无法准确描述遮挡边界区域。针对上述问题,本文算法采用遮挡区域抑制的方法进行人脸识别,提出了细粒度深度特征掩码估计的遮挡人脸识别算法。与此同时,本文所提出的算法在保证识别准确率高的前提下,降低了参数规模,并提高了运行速率。本文主要贡献如下:

(1) 提出了掩码估计模块(mask estimation module, MEM),通过对深层特征和浅层特征进行融合,并以像素级二值掩码为监督,可以更准确地估计特征掩码,从而进一步提高了遮挡人脸识别算法的识别准确率;

(2) 采用了融合余弦损失(CosFace Loss)函数^[14]与掩码估计损失函数的联合损失函数,其中掩码估计损失函数通过遮挡区域的特征掩码预测与真实分割图之间相似性,保证提取更精确的特征掩码,

有助于提高遮挡人脸识别网络的性能。

1 细粒度深度特征掩码估计的遮挡人脸识别算法

为了解决遮挡人脸图像识别精确度不高的问题,构建了一个细粒度深度特征掩码估计的遮挡人脸识别算法。该算法通过特征提取模块(feature extraction module, FEM)生成人脸图像的多尺度深度语义特征,进而利用掩码深层语义特征(mask deep semantic feature, MDSF)增强特征来生成深层特征 x_{deep} 。同时,将原图输入轻量级且性能高效的 MobileNetV3 网络,提取浅层特征 x_{low} ,并与深层特征 x_{deep} 进行融合,并以像素级二值掩码为监督得到精确的掩码 x_{cls} 。最终,将其与 FEM 模块中的深层特征 x_3 相乘,从而抑制遮挡情况下产生的干扰特征。最后,利用得到的有效特征进行人脸识别。

1.1 特征提取模块

特征提取模块采用特征金字塔来进行多尺度的特征提取。该模块采用自顶向下和横向连接的结构,并以 LResnet50E-IR 作为主干网络^[15],分别得到融合深层和浅层的融合特征 x_1 和深层特征 x_3 。特征提取模块能够提取不同尺度下的深度特征,从而更好地适应不同尺寸目标的识别需求。

图 1 为特征提取模块原理。利用特征提取模块处理随机遮挡的人脸图像后,可以从浅层到深层逐步提取特征,生成 3 种不同尺度的原始特征 C_1 、 C_2 、 C_3 。进而,采用逐级向浅层融合的策略,得到 3 种尺度下的融合特征 x_3 、 x_2 、 x_1 ,融合特征的表达式如下

$$x_3 = \text{Conv}_{3 \times 3}(\text{Conv}_{1 \times 1}(C_3)) \quad (1)$$

$$x_2 = \text{Conv}_{3 \times 3}(\text{Conv}_{1 \times 1}(C_2) + \text{Upsample}(x_3)) \quad (2)$$

$$x_1 = \text{Conv}_{3 \times 3}(\text{Conv}_{1 \times 1}(C_1) + \text{Upsample}(x_2)) \quad (3)$$

式中: $\text{Conv}_{1 \times 1}$ 为 1×1 卷积处理; $\text{Conv}_{3 \times 3}$ 为 3×3 卷积处理; Upsample 为上采样处理。

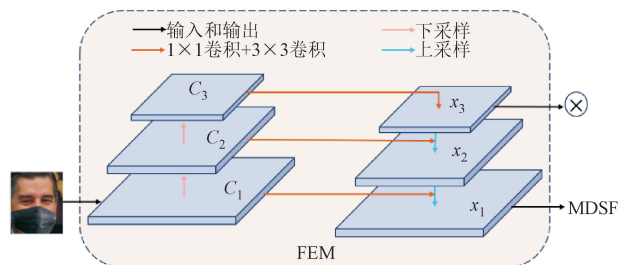


图 1 特征提取模块原理

Fig. 1 Principle of feature extraction module

由于不同尺度特征的尺寸不同,为此采用最邻近插值法对不同层特征进行上采样,以保持相邻层特征的尺寸一致,并通过 1×1 卷积调整通道数。这一流程展示了多层次的特征提取和融合过程,确保了最终人脸特征包含多尺度信息。

1.2 掩码估计模块

掩码估计模块由掩码深层语义特征提取、MobileNetV3 网络浅层特征提取和掩码特征融合 3 部分组成。掩码深度语义层包含 4 个不同膨胀率 (R) 的空洞卷积层和 1 个全局池化层。这些空洞卷积具体为 R 为 1 的 1×1 卷积、 R 为 6 的 3×3 卷积、 R 为 12 的 3×3 卷积、 R 为 18 的 3×3 卷积。空洞卷积通过增大感受野,可实现对大面积遮挡区域的掩码估计。相较于传统卷积而言,空洞卷积可以通过增大感受野来提高信息融合能力,有助于分割掩码的同时保持区域边缘细节。掩码深层语义特征模块原理如图 2 所示,其中 x_{concat} 为拼接后得到的拼接特征, x_{Conv} 为通过 1×1 卷积调整通道数后得到的中间特征。

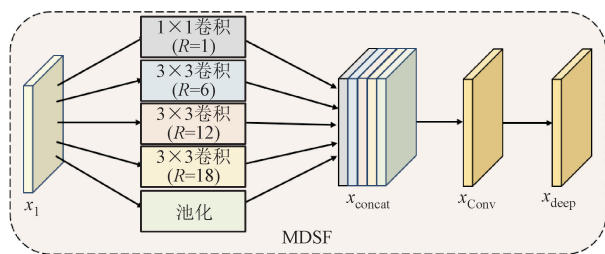


图 2 掩码深层语义特征模块原理

Fig. 2 Principle of mask deep semantic feature module

为了减少网络参数提高网络推理效率,在掩码深层语义特征和 MobileNetV3 网络浅层特征提取过程中都使用了深度可分离卷积,其将传统卷积分解为两个单独的卷积核:深度卷积和逐点卷积。深度卷积用于提取空间特征,而逐点卷积用于提取通道特征,这种方式可以显著减少网络的参数数量,还降低了计算复杂度。如图 3 所示,掩码估计模块具体流程如下。

(1) 将 FEM 模块中的融合特征 x_1 输入掩码深层语义特征提取模块得到 5 个特征图进行融合,并通过 1×1 卷积对其通道数进行调整得到中间特征 x_{Conv} 。然后,对 x_{Conv} 进行上采样得到深层特征 x_{deep} ,使其与浅层特征 x_{low} 保持尺寸一致。

(2) 将原图输入到 MobileNetV3 网络中,提取原图浅层特征 x_{low} ,并采用 1×1 卷积调整浅层特征 x_{low} 的通道数。然后,将深层特征 x_{deep} 和浅层特征 x_{low} 进行拼接得到融合特征 x_{cat} ,进而通过两次深度

可分离卷积和 Sigmoid 函数得到特征掩码 x_{cls} , 其值在 $0 \sim 1$ 之间, 代表对应特征空间位置上人脸深度特征的遮挡概率或分割权重。此时, 所得的深层细粒

度特征掩码代表了整张图像的深层细粒度特征掩码, 再通过上采样调整调整人脸特征 x_{face} 的尺寸, 使其与原始图像保持一致。

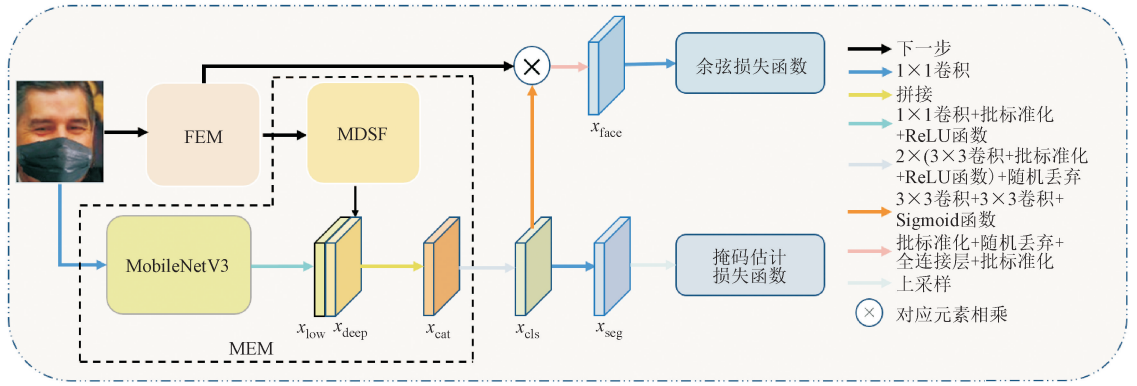


图3 细粒度深度特征掩码估计的遮挡人脸识别算法框架

Fig. 3 The framework for occluded face recognition algorithm based on fine-grained deep feature mask estimation

1.3 损失函数

本文所提算法的损失函数是由余弦损失函数和掩码估计损失函数组成的联合损失函数, 如下式所示

$$L_{\text{total}} = L_{\text{cos}} + \lambda L_{\text{seg}} \quad (4)$$

式中: L_{cos} 为评估人脸相似性的余弦损失函数; L_{seg} 为评估遮挡区域分割精度的掩码估计损失函数; λ 为损失因子。经过实验可知, 当 $\lambda = 1$ 时识别精度最高, 为此在本文所提算法中设置 $\lambda = 1$ 。

1.3.1 余弦损失函数

余弦损失函数是用于评估两个向量之间余弦相似性的损失函数, 以训练网络将相似样本映射到嵌入空间的临近位置^[14], 通常被用于监督学习任务, 例如人脸识别、文本相似性匹配等。两个向量的余弦损失函数值越高, 则相似性越高。余弦损失函数如下式所示

$$l_{\text{cos}} = -\frac{1}{M} \sum_{i=1}^M \ln[\exp(s(\cos(\theta_{y_i}, i) - m)) / (\exp(s(\cos(\theta_{y_i}, i) - m)) + \sum_{j \neq y_i} \exp(s(\cos(\theta_j, i))))] \quad (5)$$

式中: $\cos(\theta_j, i) = \mathbf{w}_j^T \mathbf{x}_i$, θ_j 是 $\mathbf{x}_i$ 与 $\mathbf{w}_j$ 之间的夹角; $\mathbf{w}_j$ 为类别 j 对应的权重向量, $\mathbf{w}_j = \frac{\mathbf{W}_j^*}{\|\mathbf{W}_j^*\|}$, $\mathbf{w}_j^T$ 是权重向量 $\mathbf{w}_j$ 的转置, $\mathbf{w}_j^*$ 是权重向量 $\mathbf{w}_j$ 的共轭; $\mathbf{x}_i = \frac{\mathbf{x}_i^*}{\|\mathbf{x}_i^*\|}$ 为本文所提算法的人脸有效特征向量, $\mathbf{x}_i^*$ 是特征向量 $\mathbf{x}_i$ 的共轭, 其对应的标签为 y_i ; M 为样本数; s 为缩放参数, 控制余弦值的分布范围; m 是

余弦间隔, 用于增加类别之间的角度间隔。

1.3.2 掩码估计损失函数

掩码估计损失函数通过二进制交叉熵损失评估掩码估计和真实掩码的相似性。二进制交叉熵损失也被称为对数损失或逻辑损失, 是一种用于度量二元分类模型性能的损失函数。它测量了模型的二元分类输出与实际标签之间的差异, 通常用于训练和评估二元分类模型, 其中每个样本属于两个可能的类别之一(通常表示为类别 0 和类别 1)。

二进制交叉熵损失能够有效地激励模型输出的概率分布接近真实标签的分布, 且在训练中可以更快收敛, 并在分割任务中得到掩码估计结果。由于本文算法掩码特征 x_{cls} 的尺寸小于原始图像, 为了利用像素级二值掩码监督模型, 本文算法通过将掩码特征 x_{cls} 进行去卷积得到原始图像大小的细粒度掩码 x_{seg} , 从而与原始遮挡标签一起进行网络训练。掩码估计损失函数表示如下

$$L_{\text{seg}} = -\frac{1}{N} \sum_{i=1}^N [p(x_i) \ln q(x_i) + (1 - p(x_i)) \ln(1 - q(x_i))] \quad (6)$$

式中: N 为样本数; $p(x_i)$ 表示真实标签的概率分布, 通常是 0 或 1, 表示样本属于类别 0 或类别 1; $q(x_i)$ 是网络的预测值, 通常是一个介于 0 和 1 之间的概率, 表示样本属于类别 1 的概率。

2 实验结果与分析

2.1 实验设置

本文算法的实验操作系统为 Ubuntu 20.04 LTS,

显卡为英伟达 RTX 3090。以式(4)中的联合损失作为损失函数,并使用随机梯度下降法^[16]作为优化算法。

网络的训练策略如下:初始学习率 L 设置为 0.01,经过 15 轮训练后,将 L 调整为 0.001,继续训练 10 轮后,将 L 设置为 0.000 1,直至训练结束。本文算法在以下所有实验中均采用相同的实验参数设置。

2.2 评价指标

本文算法采用以下 2 个指标对算法的性能进行评估:识别准确率(recognition accuracy,用符号 A_{acc} 表示)和正确接受比例(true accept rate,用符号 A_{tar} 表示),其计算公式如下

$$A_{acc} = \frac{T_p + T_N}{T_p + T_N + F_p + F_N} \quad (7)$$

$$A_{tar} = \frac{T_p}{T_p + F_N} \quad (8)$$

式中: A_{acc} 为识别准确率,表示所有正确分类的人脸对图像占全部测试人脸对图像的比例; A_{tar} 为正确接受比例,表示对属于同一人脸的图片对进行多次识别时,网络判断正确的比例; T_p 表示正确地将同一人的两张人脸图像识别为同一人的个数; F_p 表示错误地将不同人的两张图像识别为同一人的个数; T_N 表示正确地将不同人的两张人脸图像识别为不同人的个数; F_N 表示错误地将同一个人的两张人脸图像识别为不同人的个数。

2.3 数据集

与文献[11-12]相同,本文算法采用了 CASIA-WebFace^[17]数据集进行网络训练。该数据集包含了 10 575 个不同身份的人物,总共收录了 494 144 张人脸图像。为了增加数据集的多样性,本文算法在 CASIA-WebFace 数据集中引入了一些常见的遮挡,以模拟生活中可能会遇到的情况。并将处理后的数据集命名为 CASIA-WebFace-Occ,其中有遮挡和无遮挡图像的比例为 2 : 1。部分样本的示例如图 4 所示。



图 4 CASIA-WebFace-Occ 数据集示例图
Fig. 4 Samples of CASIA-WebFace-Occ dataset

为了充分验证所提算法性能,分别在以下 4 类数据集上进行实验。

(1)LFW 数据集。LFW 数据集^[18]是开源的人脸数据集,数据采集过程不受限制。该数据集共包含 5 749 个不同人物的图像,总共收录了 13 233 张人脸图像。这些图像展示了同一人物在各种背景下的面孔。该数据集包含了 6 000 对人脸图像,其中 3 000 对属于匹配的同一个人脸图像,另外 3 000 对则是来自不同身份的人脸图像。图 5 展示了数据集中的部分图像,第 1~3 列(图 5(a))为不同身份人脸对,第 4~6 列图 5(b)为相同身份人脸对。



(a)同一列上下 2 张图为不同身份人脸对
(b)同一列上下 2 张图为相同身份人脸对

图 5 LFW 数据集示例图
Fig. 5 Samples of LFW dataset

(2)LFW 口罩遮挡数据集。在日常生活中,口罩是最常见的遮挡物。本文算法在 LFW 数据集基础上创建了一个模拟口罩遮挡的数据集,该数据集包含医用外科口罩、N95 口罩、KN95 口罩以及各种不同颜色的布料口罩。图 6 展示了数据集的部分样本图像。



图 6 LFW 口罩遮挡数据集示例图
Fig. 6 Samples of LFW-mask-occlusion dataset

(3)LFW 围巾遮挡数据集。佩戴围巾在日常生活中也十分常见,为此,本文算法在 LFW 数据集上创建了一个模拟围巾遮挡的数据集。图 7 展示了 LFW 围巾遮挡数据集的部分图像。

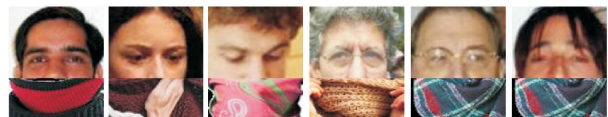


图 7 LFW 围巾遮挡数据集示例图
Fig. 7 Samples of LFW-scarf-occlusion dataset

(4)LFW 中心遮挡数据集。由于人脸的五官大多集中在中心区域,这些特征对身份认证至关重要。

因此,所提算法在人脸的中心区域添加随机的遮挡物,从而利用 LFW 数据集创建了一个模拟中心遮挡的数据集。图 8 展示了数据集的部分图像。



图 8 LFW 中心遮挡数据集示例图

Fig. 8 Samples of LFW-center-occlusion dataset

2.4 消融实验

为了验证本文所提细粒度深度特征掩码估计的遮挡人脸识别算法中不同模块的有效性,采用本文 5 种算法在 LFW 数据集上进行了消融实验^[12],实验结果如表 1 所示。

Baseline:该算法仅由主干网络 LResnet50E-IR 组

表 1 LFW 数据集消融实验结果

Table 1 Ablation experiments on the LFW dataset

算法	主干网络	特征提取模块	掩码估计模块	掩码估计损失函数	识别准确率/%
Baseline	有	无	无	无	99.08
Baseline-FEM	有	有	无	无	99.18
Baseline-MEM	有	无	有	无	99.10
Baseline-FEM-MEM	有	有	有	无	99.32
本文 BFMM	有	有	有	有	99.38

注:表中加粗数据为最优值。

由表 1 可得,在 LFW 数据集上,引入特征提取模块的 Baseline-FEM 算法显著提升了人脸识别性能,相较于 Baseline 算法,识别准确率有所提高。然而,仅包含主干网络和掩码估计模块的算法(Baseline-MEM)由于无法提取多尺度深度语义特征,并且缺乏掩码估计损失函数的监督,从而无法生成准确的特征掩码,导致识别准确率较低。Baseline-FEM-MEM 算法相较于 Baseline 算法的识别准确率提升了 0.24%,其主要原因是掩码估计模块将深层特征中的有效信息与浅层特征进行融合,生成了细粒度特征掩码有效地降低了深层损坏特征对识别的干扰,从而增强了识别性能。

相较于 Baseline-FEM-MEM 算法,本文 BFMM 算法的识别准确率提高了 0.06%。这说明引入掩码估计损失函数能够有效地激励网络输出接近真实掩码标签,在训练过程中促进网络更快地收敛,并在分类任务中取得更出色的结果,从而提升了遮挡人脸识别算法的性能。

2.5 对比实验

本文使用 4 类不同测试数据集对各种算法进行

成,损失函数为余弦损失函数。

Baseline-FEM:该算法由主干网络和特征提取模块组成,使用特征提取模块提取遮挡图片的特征,经过全连接层后计算其损失,损失函数为余弦损失函数。

Baseline-MEM:该算法由主干网络和掩码估计模块组成,将原图直接输入掩码估计模块,增强原图特征并得到掩码特征,并使用余弦损失函数作为损失函数。

Baseline-FEM-MEM:该算法由主干网络、特征提取模块以及掩码估计模块构成。将遮挡图像输入到网络中,提取出有效的特征,并使用余弦损失函数作为损失函数。

本文 BFMM 算法:该算法由主干网络、特征提取模块和掩码估计模块组成,损失函数是由余弦损失函数和掩码估计损失函数组成的联合损失函数。

性能评估,分别为:LFW 数据集、LFW 口罩遮挡数据集、LFW 围巾遮挡数据集和 LFW 中心遮挡数据集。

2.5.1 LFW 数据集对比实验

该部分对比实验基于 LFW 数据集,对比结果如表 2 所示。

表 2 LFW 数据集实验结果

Table 2 Results of LFW dataset

算法	识别准确率/%	图像训练集大小
PDSN ^[11]	99.20	49×10 ⁴
FROM ^[12]	99.38	49×10 ⁴
Center Face ^[22]	99.28	70×10 ⁴
FFR-Net ^[13]	99.43	98×10 ⁴
AdaFace ^[26]	99.83	580×10 ⁴
Baidu ^[21]	99.13	130×10 ⁴
MobileFacenet ^[20]	99.28	49×10 ⁴
Deep Face ^[23]	97.35	440×10 ⁴
VGG Face ^[24]	98.95	260×10 ⁴
Baseline	99.08	49×10 ⁴
本文 BFMM	99.38	49×10 ⁴

注:表中加粗数据为最优值。

由表 2 中的结果可知,本文 BFMM 算法的识别准确率达到 99.38%。相比与在大规模数据集上训练获得的 Center Face 算法、Baidu 算法以及 Deep Face 算法而言,本文 BFMM 算法性能更好。Baseline 在识别准确率上略优于 Deep Face 算法和 VGG Face 算法,这主要归因于主干网络的特征提取能力。由于本文 BFMM 算法采用随机从数据集中选择一定比例样本进行遮挡的方式产生训练集,其充分包含遮挡和非遮挡人脸样本。这使得算法同时对遮挡和非遮挡情况下的人脸识别准确率高且性能稳定。因此,本文 BFMM 算法在无遮挡人脸的识别上也取得了较高的精度。

2.5.2 LFW 口罩遮挡数据集对比实验

该部分对比实验基于 LFW 口罩遮挡数据集,对比结果如表 3 所示。

表 3 LFW 口罩遮挡数据集实验结果
Table 3 Results of LFW-mask-occlusion dataset

算法	识别准确率/%	正确接受比例/%
SphereFace ^[19]	96.12	52.10
FFR-Net ^[13]	97.26	85.23
ArcFace ^[25]	98.12	79.97
CosFace ^[14]	98.25	81.63
FROM ^[12]	98.15	83.07
Baseline	98.13	86.27
本文 BFMM	98.42	88.43

注:表中加粗数据为最优值。

从表 3 中可以看出,在 LFW 口罩遮挡数据集上,本文 BFMM 算法能够进行有效的识别,识别准确率达到 98.42%,正确接受比例达到 88.43%。而相比之下,SphereFace 特征提取模块结构简单,特征表征能力有限,导致其识别准确率仅为 96.12%,正确接受比例为 52.10%。相比之下,FROM 算法引入了掩码预测约束,有效提升了人脸识别准确率,但由于该算法预测出来的掩码特征为矩形区域,会导致无法识别具有曲线边缘的部分口罩区域,导致无法精确提取口罩遮挡的掩码特征,影响识别精度。同时,经典算法 ArcFace 和 CosFace 的识别准确率略低。相比于 FFR-Net 算法,本文 BFMM 算法的识别准确率提高了 1.16%。这表明本文 BFMM 算法综合性能上优于其他对比算法。

2.5.3 LFW 围巾遮挡数据集对比实验

该部分对比实验基于 LFW 围巾遮挡数据集,

对比结果如表 4 所示。

表 4 LFW 围巾遮挡数据集实验结果
Table 4 Results of LFW-scarf-occlusion dataset

算法	识别准确率/%	正确接受比例/%
SphereFace ^[19]	96.42	59.03
FFR-Net ^[13]	97.58	92.43
ArcFace ^[25]	98.48	81.17
CosFace ^[14]	98.45	81.47
FROM ^[12]	98.60	80.80
Baseline	98.37	84.40
本文 BFMM	98.72	91.57

注:表中加粗数据为最优值。

表 4 中的结果显示,SphereFace 算法识别准确率和正确接受比例均最低。与之相比,FFR-Net 算法的识别准确率提升了 1.16%,而正确接受比例最高。而 ArcFace、CosFace 和 FROM 算法的识别准确率进一步显著提升,但其正确接受比例有所下降。相比之下,本文 BFMM 算法在 LFW 围巾遮挡数据集上的识别准确率最高,达到了 98.72%,正确接受比例达到了 91.57%。结果表明,本文 BFMM 算法在围巾遮挡情况下表现出了很好的性能,具有更高的识别准确率。

2.5.4 LFW 中心遮挡数据集对比实验

该部分对比实验基于 LFW 中心遮挡数据集,对比结果如表 5 所示。

表 5 LFW 中心遮挡数据集实验结果
Table 5 Results of LFW-center-occlusion dataset

算法	识别准确率/%	正确接受比例/%
SphereFace ^[19]	96.83	53.83
FFR-Net ^[13]	96.86	77.83
ArcFace ^[25]	98.48	84.83
CosFace ^[14]	98.33	79.60
FROM ^[12]	98.28	78.90
Baseline	98.33	82.50
本文 BFMM	98.65	84.00

注:表中加粗数据为最优值。

根据表 5 中的数据,在 LFW 中心遮挡数据集上,本文 BFMM 算法的识别准确率为 98.65%,正确接受比例为 84.00%。相比之下,SphereFace 算法的识别准确率较低,为 96.83%。本文 BFMM 算法相较于经典算法 FROM,识别准确率提高了

0.37%,正确接受比例提升了 5.10%。相较于经典算法 ArcFace、CosFace 以及 FFR-Net,识别准确率和正确接受比例都有所提升。这些结果表明本文 BFMM 算法在处理中心遮挡情况下表现出了较高的精度。虽然中心遮挡导致大量人脸细节信息的丢失,但本文 BFMM 算法通过特征抑制,有效避免了遮挡区域无效特征的干扰,同时有效利用了非遮挡区域的有限人脸特征,从而取得了更高性能。

2.6 不同遮挡比例对人脸识别性能的影响

为了评估不同遮挡面积对识别性能的影响,本文在 LFW 数据集的中心区域采用了随机遮挡策略,分别设置遮挡面积占整张人脸图像的比例为 1%、5%、10%、15%、20%、30%、40% 和 50%,图 9 展示了在不同遮挡比例下同一人物面部图像的效果。由于遮挡比例超过 50%的人脸图像(如图 9(h)所示)几乎不含有识别所需的有效人脸区域,因此对遮挡比例大于 50%的人脸图像不进行比较。图 10 分别展示了这些遮挡比例对识别准确率和正确接受比例的影响。

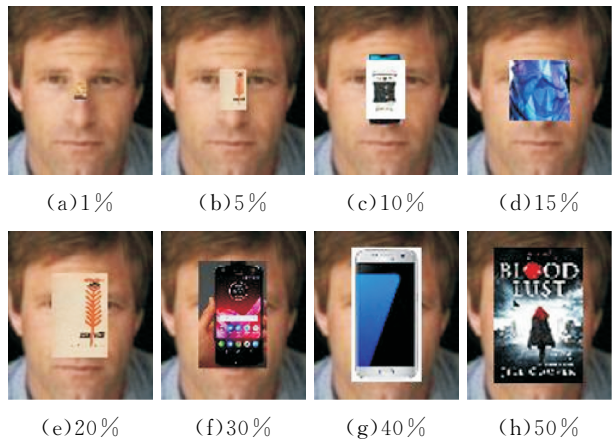
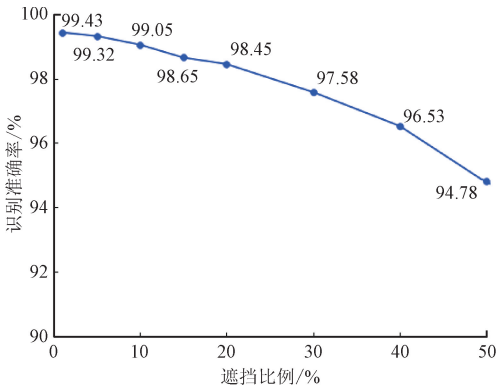
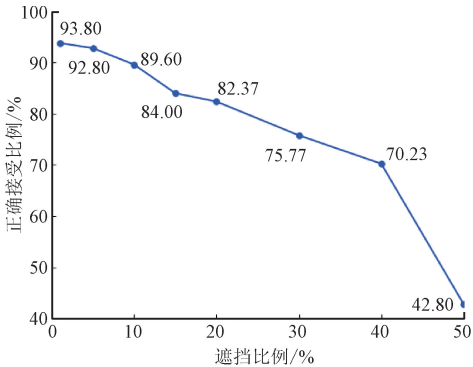


图 9 同一人脸在不同遮挡比例下的图像

Fig. 9 The image of the same face under different occlusion ratio



(a)不同遮挡面积下的识别准确率



(b)不同遮挡面积下的正确接受比例

图 10 不同遮挡面积下的识别准确率和正确接受比例

Fig. 10 Recognition accuracy and true accept rate of different occluded areas

如图 9 所示,当遮挡面积超过 40%时,人脸五官几乎完全被遮挡,从而导致识别错误率显著提升。图 10 中的结果显示,当遮挡面积超过总面积的 40%时,识别准确率和正确接受比例急剧下降。因此,遮挡对人脸识别算法的影响是非线性的,且当遮挡比例达到一定程度时,算法难以提取有效的识别特征,从而降低了识别性能。

3 结 论

针对因人脸受到遮挡使得部分面部结构信息丢失,从而影响识别准确率的问题,本文提出了一种细粒度深度特征掩码估计的遮挡人脸识别算法。本文的主要结论如下。

- (1)本文算法利用特征金字塔网络提取人脸图像的多尺度深度语义特征,确保了最终人脸特征包含多尺度信息。
- (2)利用掩码深层语义特征处理获得深层特征,并与 MobileNetV3 网络所提取的浅层特征进行融合,进而以像素级二值掩码为标签训练网络以获得细粒度特征掩码。利用该深度特征掩码与深层特征相乘,以抑制遮挡干扰特征,提高遮挡人脸识别的准确率。最后,采用余弦损失函数和掩码估计损失函数联合训练网络,从而提高了人脸识别算法对任意遮挡的鲁棒性。同时,在识别准确率高的前提下,本文算法的参数规模更小,运行速率更高。
- (3)尽管本文提出的细粒度深度特征掩码估计算法在遮挡人脸识别中表现出较好的性能,但该算法仍存在一些局限性,特别是在识别非正面人脸遮挡或复杂遮挡情况时。首先,本文算法主要针对正面人脸设计,因而在非正面人脸的遮挡情况下,其性

能可能会受到限制。非正面视角下的人脸特征与正面视角存在显著差异,这可能导致生成的特征掩码无法准确反映遮挡区域的实际形状,从而影响识别效果。此外,尽管细粒度掩码能够提供更精细的遮挡边界描述,但在处理复杂遮挡情况下,例如遮挡区域具有高度遮挡的细节或遮挡物的纹理与人脸特征高度相似时,当前算法仍可能表现不佳。这些复杂遮挡场景可能导致网络难以充分学习并准确估计遮挡区域的细节,从而在实际应用中导致识别准确率降低。

(4)实验结果表明,本文所提算法在不同类型遮挡情况下均获得十分稳定的人脸识别结果。但是,仍难以识别部分非正面遮挡人脸。针对这个问题,下一步会继续探索基于遮挡区域恢复的研究方法,提高非正面遮挡人脸图像的识别准确率。

参考文献:

- [1] 李明, 常青霞. 融合 Transformer 和 CNN 的轻量级人脸识别算法 [J]. 计算机工程与应用, 2024, 60(14): 96-104.
LI Ming, DANG Qingxia. Lightweight face recognition algorithm combining transformer and CNN [J]. Computer Engineering and Applications, 2024, 60(14): 96-104.
- [2] 岳也, 温瑞萍, 王川龙. 带有特征信息卷积神经网络的人脸识别算法 [J]. 工程数学学报, 2024, 41(3): 410-420.
YUE Ye, WEN Ruiping, WANG Chuanlong. Face recognition algorithm based on convolutional neural network with feature information [J]. Chinese Journal of Engineering Mathematics, 2024, 41(3): 410-420.
- [3] 余璀璨, 李慧斌. 基于深度学习的人脸识别方法综述 [J]. 工程数学学报, 2021, 38(4): 451-469.
YU Cuican, LI Huibin. Deep learning based 2D face recognition: a survey [J]. Chinese Journal of Engineering Mathematics, 2021, 38(4): 451-469.
- [4] TOMAR V, KUMAR N, SRIVASTAVA A R. Single sample face recognition using deep learning: a survey [J]. Artificial Intelligence Review, 2023, 56(1): 1063-1111.
- [5] JEEVAN G, ZACHARIAS G C, NAIR M S, et al. An empirical study of the impact of masks on face recognition [J]. Pattern Recognition, 2022, 122: 108308.
- [6] WRIGHT J, YANG A Y, GANESH A, et al. Robust face recognition via sparse representation [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2009, 31(2): 210-227.
- [7] ZHOU Zihan, WAGNER A, MOBAHI H, et al. Face recognition with contiguous occlusion using Markov random fields [C]//2009 IEEE 12th International Conference on Computer Vision. Piscataway, NJ, USA: IEEE, 2009: 1050-1057.
- [8] 张铮, 芦天亮, 曹金璇. 基于分割和多级掩膜学习的遮挡人脸识别方法 [J]. 计算机科学与探索, 2024, 18(7): 1814-1825.
ZHANG Zheng, LU Tianliang, CAO Jinxuan. Occluded face recognition based on segmentation and multi-stage mask learning [J]. Journal of Frontiers of Computer Science and Technology, 2024, 18(7): 1814-1825.
- [9] CHEN Yuan, CHEN Weiche, WEI C P, et al. Occlusion-aware face inpainting via generative adversarial networks [C]//2017 IEEE International Conference on Image Processing (ICIP). Piscataway, NJ, USA: IEEE, 2017: 1202-1206.
- [10] ZHAO Fang, FENG Jiashi, ZHAO Jian, et al. Robust LSTM-autoencoders for face de-occlusion in the wild [J]. IEEE Transactions on Image Processing, 2018, 27(2): 778-790.
- [11] SONG Lingxue, GONG Dihong, LI Zhifeng, 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). Piscataway, NJ, USA: IEEE, 2019: 773-782.
- [12] QIU Haibo, GONG Dihong, LI Zhifeng, et al. End2end occluded face recognition by masking corrupted features [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2022, 44(10): 6939-6952.
- [13] HAO Shaozhe, CHEN Chaofeng, CHEN Zhenfang, et al. A unified framework for masked and mask-free face recognition via feature rectification [C]//2022 IEEE International Conference on Image Processing (ICIP). Piscataway, NJ, USA: IEEE, 2022: 726-730.
- [14] WANG Hao, WANG Yitong, ZHOU Zheng, et al. CosFace: large margin cosine loss for deep face recognition [C]//2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Piscataway, NJ, USA: IEEE, 2018: 5265-5274.
- [15] LIN T Y, DOLLÁR P, GIRSHICK R, et al. Feature pyramid networks for object detection [C]//2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Piscataway, NJ, USA: IEEE, 2017: 936-944.

[16] 史加荣, 王丹, 尚凡华, 等. 随机梯度下降算法研究进展 [J]. 自动化学报, 2021, 47(9): 2103-2119.
SHI Jiarong, WANG Dan, SHANG Fanhua, et al. Research advances on stochastic gradient descent algorithms [J]. Acta Automatica Sinica, 2021, 47(9): 2103-2119.

[17] YI Dong, LEI Zhen, LIAO Shengcai, et al. Learning face representation from scratch [EB/OL]. (2014-11-28)[2024-08-01]. <https://arxiv.org/abs/1411.7923>.

[18] HUANG G B, MATTAR M, BERG T, et al. Labeled faces in the wild: a database for studying face recognition in unconstrained environments [C]//Workshop on Faces in ‘Real-Life’ Images: Detection, Alignment, and Recognition. Marseille, France: Inria, 2008: 617-624.

[19] LIU Weiyang, WEN Yandong, YU Zhiding, et al. SphereFace: deep hypersphere embedding for face recognition [C]//2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Piscataway, NJ, USA: IEEE, 2017: 6738-6746.

[20] CHEN Sheng, LIU Yang, GAO Xiang, et al. MobileFaceNets: efficient CNNs for accurate real-time face verification on mobile devices [C]//Biometric Recognition. Cham: Springer International Publishing, 2018: 428-438.

[21] LIU Jingtuo, DENG Yafeng, BAI Tao, et al. Targeting ultimate accuracy: face recognition via deep embedding [EB/OL]. (2015-07-23)[2024-08-01]. <https://arxiv.org/abs/1506.07310>.

[22] WEN Yandong, ZHANG Kaipeng, LI Zhifeng, et al. A discriminative feature learning approach for deep face recognition [C]//Computer Vision-ECCV 2016. Cham: Springer International Publishing, 2016: 499-515.

[23] TAIGMAN Y, YANG Ming, RANZATO M, et al. DeepFace: closing the gap to human-level performance in face verification [C]//2014 IEEE Conference on Computer Vision and Pattern Recognition. Piscataway, NJ, USA: IEEE, 2014: 1701-1708.

[24] PARKHI O M, VEDALDI A, ZISSERMAN A. Deep face recognition [C]//BMVC 2015-Proceedings of the British Machine Vision Conference 2015. Swansea, UK: BMVA, 2015: 1-12.

[25] DENG Jiankang, GUO Jia, XUE Niannan, et al. ArcFace: additive angular margin loss for deep face recognition [C]//2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Piscataway, NJ, USA: IEEE, 2019: 4685-4694.

[26] KIM M, JAIN A K, LIU Xiaoming. AdaFace: quality adaptive margin for face recognition [C]//2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Piscataway, NJ, USA: IEEE, 2022: 18729-18738.

(编辑 刘杨 陶晴)