

基于深度学习的跨年龄人脸识别

孙文斌¹, 王荣^{2,3*}, 孙连烛⁴, 林源松¹

¹广东工业大学信息工程学院, 广东 广州 510006;

²西北农林科技大学信息工程学院, 陕西 咸阳 712100;

³国家农业信息化工程技术研究中心, 北京 100097;

⁴北京交通大学交通运输学院, 北京 100044

摘要 年龄变化是影响人脸识别模型性能的主要原因之一,为解决年龄变化所带来的模型识别率低的问题,提出了一种基于深度学习的跨年龄卷积神经网络模型(CA-CNN)用于跨年龄人脸识别。首先,利用卷积神经网络提取人脸图像中的深度人脸特征;然后,提出一种高效的卷积注意力模块从深度人脸特征中获取年龄特征,并结合多层感知机和多任务监督学习,将深度人脸特征非线性分解为年龄特征和身份特征;最后,为了更好地区分身份特征和年龄特征,提出了一种批核典型相关性分析模块对分解后的身份特征和年龄特征进行相关性分析。经过对抗性学习训练后,相关性最小化,实现了跨年龄人脸识别。所提模型在MORPH Album 2数据集上的rank-1识别准确率达到99.03%,在CALFM数据集上的人脸验证等错率为9.8%,表明了所提模型的有效性。

关键词 机器视觉; 人脸识别; 深度学习; 注意力模型; 相关性分析; 跨年龄

中图分类号 TP751

文献标志码 A

doi: 10.3788/LOP202259.0215001

Cross-Age Face Recognition Based on Deep Learning

Sun Wenbin¹, Wang Rong^{2,3*}, Sun Lianzhu⁴, Lin Yuansong¹

¹School of Information Engineering, Guangdong University of Technology,
Guangzhou, Guangdong 510006, China;

²College of Information Engineering, Northwest A&F University, Xianyang, Shaanxi 712100, China;

³National Engineering Research Center for Information Technology in Agriculture, Beijing 100097, China;

⁴School of Traffic and Transportation, Beijing Jiaotong University, Beijing 100044, China

Abstract Age change is one of the main reasons that affect the performance of face recognition. In order to solve the problem of low face recognition rate caused by the change of age, a cross-age face recognition model (CA-CNN) based on deep learning is proposed for cross-age face recognition. First, the overall face features are extracted from the face image by the convolutional neural network; then, an efficient convolutional attention module is proposed to obtain age features from overall face features, and combined with multi-layer perceptrons and multi-task supervised learning, the overall face features are non-linearly decomposed into age features and identity features; finally, for better distinguish between identity features and age features, an approved batch kernel canonical correlation analysis module is put forward to analyze the correlation between the decomposed identity features and age features. After the training of adversarial learning, the correlation is minimized and cross-age face recognition is realized. The proposed model achieves the recognition accuracy up to 99.03% of rank-1 on the MORPH Album 2 dataset, and the face verification of equal error rate of 9.8% on the CALFM dataset, which indicates the effectiveness of the proposed model.

Key words machine vision; face recognition; deep learning; attention model; correlation analysis; cross-age

收稿日期: 2020-12-31; 修回日期: 2021-02-03; 录用日期: 2021-03-09

通信作者: *rongw@nwafu.edu.cn

1 引言

人脸识别作为计算机视觉领域应用十分广泛的技术,数十年来一直受到学术界和工业界的关注。人脸识别一般包括人脸检测、特征提取和特征比对 3 个阶段,准确率受年龄、角度、光照、遮挡和面部表情等因素影响较大。目前人脸识别算法虽然取得了很好的识别效果,但这些算法都是针对相同年龄的人脸图像进行识别的。Cheng 等^[1]和 Best-Rowden L 等^[2]的研究结果表明,年龄的变化对人脸识别的影响较大。跨年龄的人脸识别仍然是一项比较艰巨的任务^[3-5],因此探索一种抗年龄干扰的人脸识别方法具有重要意义。

现有的跨年龄人脸识别模型主要为生成式模型、判别式模型及深度学习模型^[6-7]。生成式模型主要针对年龄变化过程中的人脸老化过程进行建模,通过已标注年龄信息的人脸图像生成其他年龄段的人脸图像,对数据集的要求比较苛刻,需要大量带有年龄标签的人脸图像。判别式模型的主要思想则是寻找人脸图像中对年龄具有鲁棒性的特征,从而减小年龄对模型的影响。基于梯度方向金字塔(GOP)、局部特征描述子(LBP)及旋转不变特征(SIFT)相结合的模型^[8-9]先后被提出。Gong 等^[10-11]提出了隐变量分析(HFA)和增强的隐变量分析(MHFA)特征分解模型,MHFA 利用隐性因子分析算法将提取到的特征线性分解为平均脸特征、身份因子和隐性年龄因子,从而将年龄特征从人脸特征中分离,显著提高了跨年龄人脸识别模型的性能。但上述方法在一定程度上受限于人工特征描述子的设计和选择,存在着一定的局限性。

近年来,深度学习发展迅速,由于强大的非线性表达能力,在计算机视觉、语音识别及自然语言处理等领域都取得了显著成效。Wen 等^[12]将卷积神经网络(CNN)应用于跨年龄人脸识别模型,提出了隐性因子引导的 CNN 模型(LF-CNN)。Xu 等^[13]提出了耦合自编码网络(CAN)来模拟人脸老化过程,该网络通过自编码网络学习到身份因子、年龄因子及噪声因子的非线性表示。Zheng 等^[14]受文献^[10-11]的启发,去掉了平均脸特征,认为年龄因素不是隐性的因子,年龄特征也可以通过年龄回归任务得到,提出了年龄引导下的跨年龄人脸识别模型(AE-CNN)。为实现更好的识别效果,Wang 等^[15]利用正交分解将人脸特征分解为互相正交的

身份信息 and 年龄信息;Wang 等^[16]借用典型相关性分析(CCA)^[17]的思想,提出了适用于深度学习的批典型相关性分析方法(BCCA),该方法通过去相关对抗学习将人脸特征分离为两个不相关的身份信息和年龄信息,在跨年龄人脸数据集上取得了较好的效果,表明了身份和年龄信息的非线性关系不可忽略。

本文提出了一种基于深度学习的跨年龄卷积神经网络模型(CA-CNN)用于跨年龄人脸识别。采用一种高效的卷积注意力模块(ECBAM)对深度人脸特征进行非线性分解,并利用多层感知机(MLP)^[18]去除身份特征中的年龄影响因子;同时,提出一种批核典型相关性分析(BKCCA)方法,对分解后的身份特征和年龄特征进行相关性约束,以期实现更好的跨年龄人脸识别效果。

2 所提算法

2.1 基于深度学习的跨年龄人脸识别模型

所提模型主要由 CNN 特征提取模块、深度人脸特征分解模块和 BKCCA 约束模块 3 部分组成,如图 1 所示。为使用多任务学习的方式提取人脸图像中的身份信息和年龄信息,首先由 CNN 特征提取模块提取深度人脸特征图(feature map),利用 ECBAM、MLP 和多任务监督学习模块对深度人脸特征进行非线性分解,得到年龄特征向量 \mathbf{X}_{age} 和身份特征向量 \mathbf{X}_{id} 。然后对 \mathbf{X}_{age} 和 \mathbf{X}_{id} 进行 BKCCA,通过反复训练,调整权重,降低两者之间的相关性。最后将模型优化后的身份特征 \mathbf{X}_{id} 与人脸库中的特征向量进行余弦距离比较,达到跨年龄人脸识别的目的。

为提高跨年龄人脸识别模型的准确率,寻找一种有效的 CNN 提取深度人脸特征至关重要。主流的 CNN 网络如 AlexNet^[19]、VGG^[20]等都通过增加网络的层数来获得更好的效果,但通过增加网络的深度来提升性能,网络容易出现过拟合、梯度消失和梯度爆炸等问题。针对以上问题,文献^[21]和文献^[22]分别从拓展网络宽度和跨层连接的角度提出了 GoogLeNet 和 ResNet。Deng 等^[23]提出了一种更适合人脸特征提取的 improved residual unit (IR) 网络结构,该结构由批量归一化层(BN)^[24]、 3×3 卷积、非线性映射单元 PReLU^[25]、 3×3 卷积、BN 层串联组成。

受 InsightFace、ResNet、OE-CNN^[15]和文献^[26]

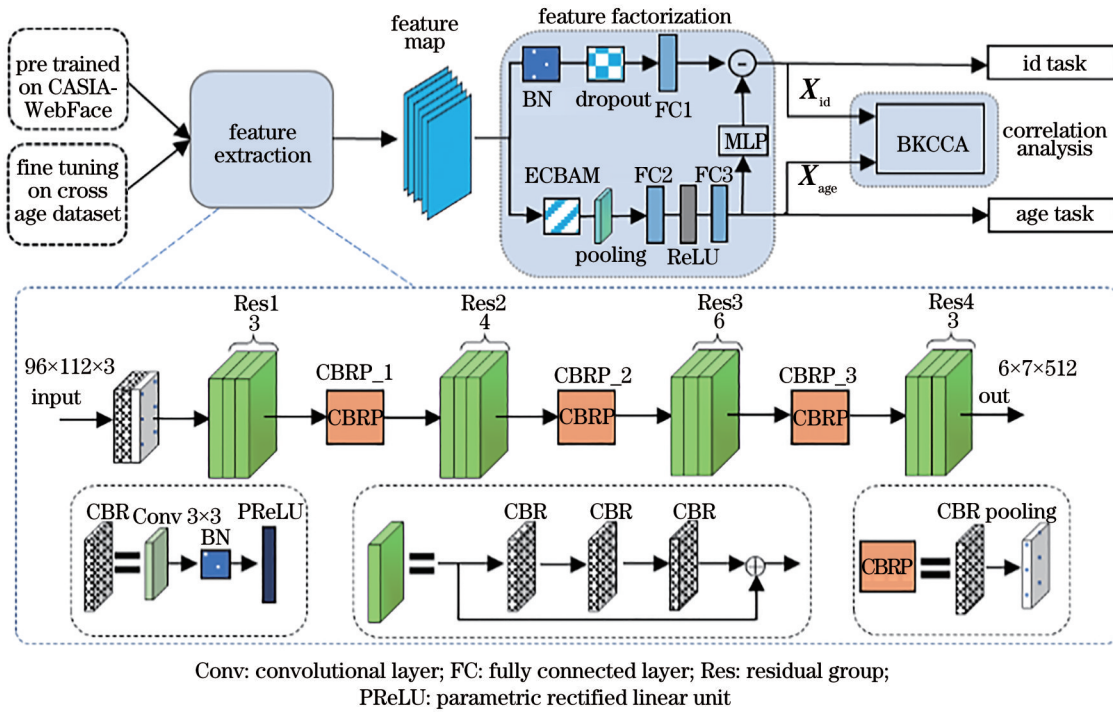


图 1 CA-CNN 模型结构图

Fig. 1 Structure of CA-CNN model

的启发,设计了图 1 中的 CNN 特征提取模块。该特征提取模块主要由 2 组 CBR 模块、4 个残差组 (Res1、Res2、Res3、Res4)、3 个 CBRP 模块和 1 个池化层组成,每个残差组分别包含 3、4、6、3 个残差块,每个残差块由 3 组 CBR 组成。每个残差组输入特征通道数和输出特征通道数相同,分别为 64、128、256、512,残差块之间通过 CBRP 模块进行下采样及通道维度的扩充。预处理之后的人脸图像大小为 $96 \times 112 \times 3$,将其输入到 CNN 特征提取模块,经过两组 CBR 模块及最大池化后,特征图大小为 $48 \times 56 \times 64$,残差组 Res1 不改变特征图的大小,经过 CBRP_1 模块后,特征图大小为 $24 \times 28 \times 128$,同理可知经过残差组 Res4 后的特征图大小为 $6 \times 7 \times 512$ 。

2.2 深度人脸特征分解模块

为减少从人脸特征中分离年龄特征时对身份特征造成的损失,提出了一种非线性深度人脸特征分解模块 (feature factorization),结构如图 1 所示。首先利用 ECBAM 对年龄子网进行监督学习,增强显著的年龄特征并抑制其他冗余特征,从而得到更加充分和准确的年龄特征。然后将身份因子与 MLP 提取的年龄因子相减,实现非线性人脸特征分解,得到不含年龄影响的身份特征,去除年龄因素对身份识别的影响。

注意力机制是仿照人类视觉的观察模型,使网

络更关注目标的显著信息,抑制冗余信息,具有对特定区域进行特征提取的作用。卷积注意力模块 (CBAM)^[27] 依次利用通道注意力模块 (CAM) 和空间注意力模块 (SAM) 来增强特征的表达能力,结构如图 2(a) 所示。CAM 模块中,分别将全局平均池化 (GAP) 和全局最大池化 (GMP) 后的特征输入至 MLP 进行通道间的信息交互;而 CAM 模块中 MLP 先压缩后扩展的操作会造成信息的部分损失。ECA-Net^[28] 证实了适当的跨通道交互能够提高通道注意力模块特征提取的有效性。所使用的 ECBAM 由高效空间注意力模块 (ECAM) 和 SAM 模块串联组成,通过引入可训练参数 α 对不同通道特征加权,结构如图 2(b) 所示。

ECAM 利用 2 个一维卷积代替 CAM 中的共享 MLP 对 GAP 和 GMP 两种池化通道上的特征实现相邻 K 个通道间的信息交互,2 个通道的卷积核参数不共享;为提高模型的表达能力,引入可训练参数 α 对双通道信息进行加权求和,通过反向传播训练,实现模型自主选择通道中的有效特征;最后经过 Sigmoid 函数得到每个通道的权重。假设经过 CNN 特征提取模块得到的 feature map 为 $t \in \mathbb{R}^{C \times W \times H}$, C, W, H 分别为特征图的通道数、宽和高,特征图经过 ECAM 模块和 SAM 模块的输出分别为 W'_{age} 和 W''_{age} 。 W'_{age} 的表达式为

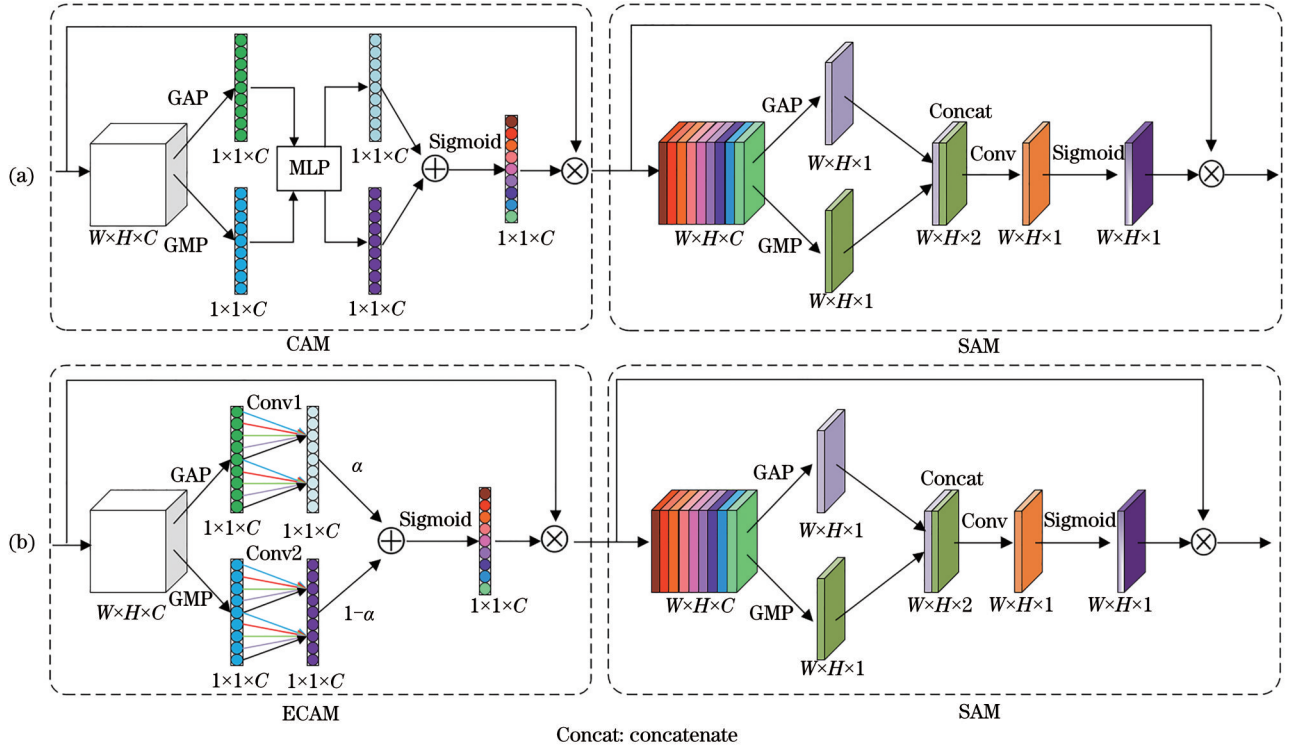


图2 CBAM和ECBAM的结构。(a) CBAM; (b) ECBAM

Fig. 2 Structure of CBAM and ECBAM. (a) CBAM; (b) ECBAM

$$W'_{age} = t \otimes \sigma \left\{ \text{Conv1} \left[F^1_{GAP}(t), F^2_{GAP}(t), \dots, F^C_{GAP}(t) \right] \times \alpha + \text{Conv2} \left[F^1_{GMP}(t), F^2_{GMP}(t), \dots, F^C_{GMP}(t) \right] \times (1 - \alpha) \right\}, \quad (1)$$

$$F^k_{GAP}(t) = \frac{1}{W \times H} \left[\sum_{i=1}^W \sum_{j=1}^H v_k(i, j) \right], \quad i \in [1, W], j \in [1, H], k \in [1, C], \quad (2)$$

$$F^k_{GMP}(t) = \text{Max} \left[v_k(i, j) \right], \quad i \in [1, W], j \in [1, H], k \in [1, C], \quad (3)$$

式中: F^k_{GAP} 和 F^k_{GMP} 分别表示对第 k 个通道上的特征图进行 GAP 和 GMP 操作; $v_k(i, j)$ 表示第 k 个通道上的特征图中坐标为 (i, j) 的值; Conv1 和 Conv2 表示两个一维卷积操作, 卷积核大小设为 5, $\alpha \in [0, 1]$, 参与训练; σ 表示 Sigmoid 激活函数; \otimes 表示矩阵张量元素对应位置相乘。

SAM 模块利用特征的空间关系生成空间注意力图, 突出特征图中的有效信息, 是通道注意力的补充, 结构如图 2(b) 所示。SAM 模块对 GAP 和 GMP 生成的 2 个二维特征进行通道上的级联 (Concat), 经过卷积操作后送入 Sigmoid 中得到空间维度上的注意力权重, 然后将其与输入特征相乘, 实现特征重校准。 W''_{age} 的表达式为

$$W''_{age} = W'_{age} \otimes \sigma \left[f^{3 \times 3} (F^s_{GAP}, F^s_{GMP}) \right] \quad (4)$$

$$F^s_{GAP}(i, j) = \frac{1}{C} \sum_{k=1}^C W'^k_{age}(i, j), \quad i \in [1, W], j \in [1, H], k \in [1, C], \quad (5)$$

$$F^s_{GMP}(i, j) = \text{Max} \left[W'^1_{age}(i, j), W'^2_{age}(i, j), \dots, W'^C_{age}(i, j) \right], \quad i \in [1, W], j \in [1, H], \quad (6)$$

式中: $F^s_{GAP}(i, j)$ 和 $F^s_{GMP}(i, j)$ 分别表示对特征图通道上位于坐标 (i, j) 处的所有元素进行 GAP 和 GMP 操作; $f^{3 \times 3}(\cdot)$ 表示 3×3 的卷积操作; \otimes 表示矩阵

张量元素对应位置相乘。

2.3 相关性分析模块

多数深度人脸特征分解模块仅通过年龄任务加身份识别任务进行联合监督学习, 分离年龄特征

和身份特征,并未增加分析模块来约束年龄特征和身份特征。文献[26]利用余弦相似性度量引导年龄任务与身份识别任务之间的特征学习,但该方法只能对两个特征向量对应位置的元素进行分析。因此,DAL模型^[16]利用CCA的思想,提出了适用于深度学习的BCCA,BCCA将年龄和身份的高维特征通过线性变换映射到低维空间上进行相关性分析,但忽略了特征向量内部元素之间的非线性关系。核典型相关性分析(KCCA)^[29]的思想是将低维特征的内部元素通过非线性核函数映射到高维空间,然后再将高维特征线性映射到低维空间进行相关性分析的。基于KCCA、MLP和DAL模型^[16]的思想,所提BKCCA模块如图3所示,该模块对特征分解模块输出的身份特征向量 \mathbf{X}_{id} 和年龄特征向量 \mathbf{X}_{age} 的非线性关系进行相关性分析。

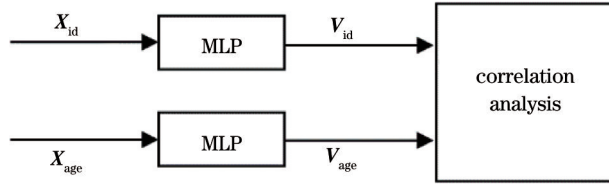


图3 BKCCA 模块

Fig. 3 BKCCA module

首先将 \mathbf{X}_{age} 和 \mathbf{X}_{id} 利用MLP非线性映射到高维空间,然后再线性映射到一维空间上进行分析,并使用斯皮尔曼相关系数 ρ 分析一维空间上的关系。

$$\forall i \in \{id, age\} : \mathbf{V}_i = \mathfrak{R}_i(\mathbf{B}_i), \quad (7)$$

$$\rho(\mathbf{V}_{id}, \mathbf{V}_{age}) = \left[\frac{\text{Cov}(\mathbf{V}_{id}, \mathbf{V}_{age})}{\sqrt{\text{Var}(\mathbf{V}_{id})} \cdot \sqrt{\text{Var}(\mathbf{V}_{age})}} \right], \quad (8)$$

式中: \mathfrak{R} 表示MLP函数; $\text{Cov}(\cdot)$ 为协方差; $\text{Var}(\cdot)$ 为方差。设 $\mathbf{B}_{id} = \{\mathbf{X}_{id}^1, \dots, \mathbf{X}_{id}^m\}$, $\mathbf{B}_{age} = \{\mathbf{X}_{age}^1, \dots, \mathbf{X}_{age}^m\}$ 为一个batch, m 为batch的大小,设为32,则 ρ 可化简为

$$\rho(\mathbf{V}_{id}, \mathbf{V}_{age}) = \frac{\frac{1}{m} \sum_{i=1}^m (v_{id}^i - \bar{v}_{id})(v_{age}^i - \bar{v}_{age})}{\sqrt{\sigma_{id}^2 + \varepsilon} \cdot \sqrt{\sigma_{age}^2 + \varepsilon}}, \quad (9)$$

式中: \bar{v} 和 σ^2 分别为均值和方差; ε 是一个避免分母为零的极小值常量。由(9)式可知, $\rho \in [-1, 1]$,因此,定义优化的相关性约束损失函数为 $|\rho|$ 。

设输入图像为 \mathbf{P} , $F(\cdot)$ 表示模型的CNN特征提取模块,则大小为 $6 \times 7 \times 512$ 的深度人脸特征图(feature map) \mathbf{t} 的计算过程为

$$\mathbf{t} = F(\mathbf{P}). \quad (10)$$

\mathbf{t} 经过特征分解模块后,得到512维的身份特征向量 \mathbf{X}_{id} 和年龄特征向量 \mathbf{X}_{age} 。

$$\mathbf{X}_{age} = F_3(\mathbf{t}), \quad (11)$$

$$\mathbf{X}_{id} = F_1(\mathbf{t}) - F_2(\mathbf{X}_{age}), \quad (12)$$

式中: $F_3(\cdot)$ 表示 \mathbf{t} 到 \mathbf{X}_{age} 的映射; $F_2(\cdot)$ 为MLP; $F_1(\cdot)$ 表示 \mathbf{t} 到全连接层FC1输出的映射。

采用对抗学习的方式进行网络参数的更新训练。首先,保持 $F, F_1, F_2, F_3, F_{id}, F_{age}$ 的参数不变,通过最大化 $|\rho|$ 来训练 \mathfrak{R} ; \mathfrak{R} 参数固定后,再训练参数 $F, F_1, F_2, F_3, F_{id}, F_{age}$,计算 \mathbf{X}_{id} 和 \mathbf{X}_{age} 之间的相关性。重复上述步骤,通过不断迭代更新网络参数,使得最后两特征向量之间相关性最小。

$$L_{(id, age)} = \min_{F, F_1, F_2, F_3, F_{id}, F_{age}} \max_{\mathfrak{R}} \left| \rho \left\{ \mathfrak{R} \left\{ F_1[F(\mathbf{P})] \right\} \right\} - F_2 \left\{ F_3[F(\mathbf{P})] \right\}, \mathfrak{R} \left\{ F_3[F(\mathbf{P})] \right\} \right|, \quad (13)$$

式中: F_{id} 和 F_{age} 分别为身份任务和年龄任务的映射。

2.4 模型的多任务训练

3个模块联合监督学习身份任务、年龄任务和相关性约束任务。进行身份任务判别时,使用CosFace损失^[30]对身份特征 \mathbf{X}_{id} 进行监督学习,将分类边界投影到余弦角度空间上,以期增大类间距离,减少类内距离。CosFace损失的表达式为

$$L_{id} = \frac{1}{N} \sum_{i=1}^N - \log \frac{e^{s[\cos(\theta_{y_i}, i) - b]}}{e^{s[\cos(\theta_{y_i}, i) - b]} + \sum_{j \neq y_i} e^{s \times \cos(\theta_j, i)}}, \quad (14)$$

式中: N 为身份个体的数量; y_i 为第 i 个样本的身份标签; $\cos(\theta_{y_i}, i) = \frac{\mathbf{W}_j^T \cdot \mathbf{X}_{id}^i}{\|\mathbf{W}_j\| \cdot \|\mathbf{X}_{id}^i\|}$ 为第 i 个样本特征 \mathbf{X}_{id}^i

和类别 y_i 之间的余弦值。 $s(s > 0)$ 为缩放因子; b 为控制余弦边界的常量。

对于年龄任务,采用线性回归的方式估计年龄,以均方差作为损失函数。

$$L_{age} = \frac{1}{2m} \sum_{i=1}^m [F_{age}(\mathbf{X}_{age}^i) - z^i]^2, \quad (15)$$

式中: \mathbf{X}_{age}^i 为第 i 个样本的年龄特征向量; z^i 为第 i 个样本的年龄标签。

综合(13)~(15)式,整个网络的联合损失函数为

$$L = L_{id} + \lambda_1 L_{age} + \lambda_2 L_{(id, age)}, \quad (16)$$

式中: λ_1 和 λ_2 为平衡3个损失函数的权重超参数。

3 实验分析与讨论

3.1 数据预处理与实验方案

为避免角度对跨年龄人脸识别模型的干扰,在进行模型训练、测试之前,首先用MTCNN算法^[31]对人脸图像进行面部和面部关键点检测,然后根据5个面部关键点进行相似性变换,实现人脸对齐。对齐后的人脸图像大小为 96×112 ,将人脸图像像素值归一化到 $[-1, 1]$ 。对齐效果如图4所示。

由于跨年龄人脸数据集采集困难、涉及隐私等问题,可用于跨年龄人脸识别的公开数据非常有限。因此,按照文献[10-12,32]的方式,由CNN特征提取模块加上FC1全连接层组成Baseline模型,利用Softmax损失函数在CASIA-WebFace数据集上预训练,得到Baseline模型参数,然后加入所提特征分解模块、相关性分析模块及多任务损失函数,利用跨年龄人脸数据集进行模型的微调,从而实现对CA-CNN模型的训练。

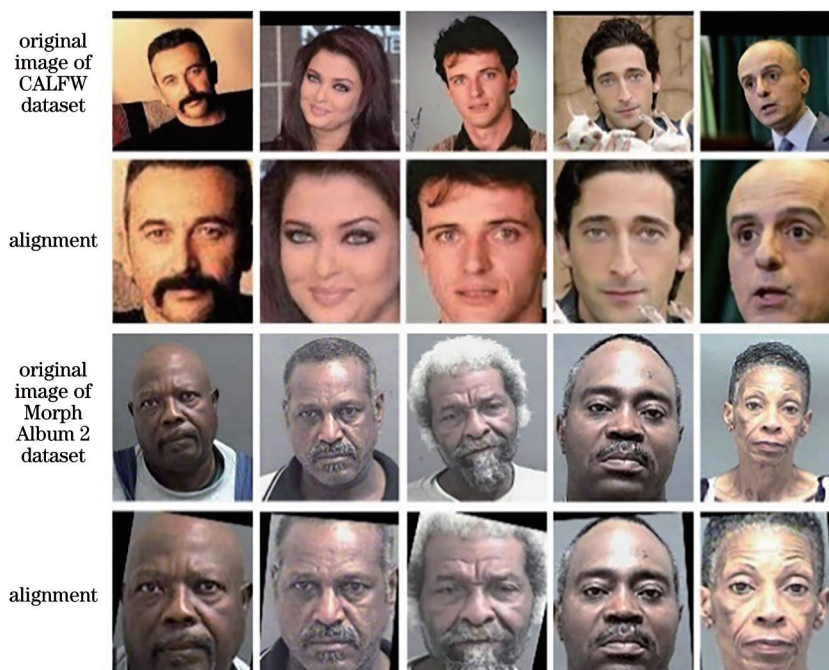


图4 人脸对齐效果

Fig. 4 Face alignment effect

模型预训练时,在Ubuntu16.04 64位操作系统和Pytorch 1.3.0版本上搭建深度学习算法训练平台,采用1块GPU(NVIDIA Tesla P100)进行加速,batch size设置为128,采用随机梯度下降法(SGD)进行模型优化,初始学习率设置为0.01,动量因子为0.9,迭代训练轮数设置为70,在第5、15、30和45轮,学习率降至上轮的十分之一。模型微调与测试时,在Windows 10 64位操作系统和Pytorch 1.2.0上搭建训练测试平台,采用1块GPU(NVIDIA 2080TI)进行加速。

3.2 Morph Album 2数据集上的实验

Morph Album 2数据集包含55134张人脸图像,13617个个体,单个个体的年龄跨度最大为12周岁,其中457个个体只有一张人脸图片,因此可用的个体数为13160。经过图像预处理后,按照文献[33]的设置将整个数据集划分为训练集和测试

集,随机抽取10000个个体内年龄跨度最大的两张图片作为训练集,共20000张图片;剩下的抽取3000个个体内年龄跨度最大的两张图片作为测试集,共6000张图片,其中年龄较小的作为gallery集,年龄较大的作为probe集,训练集和测试集无交集。在模型微调阶段,batch大小设置为32,训练轮数为25,初始学习率设置为0.001,学习率在第6、10和18轮衰减为上一轮学习率的十分之一。(14)式中的 b 和 s 按照文献[15-16,32]设置为 $b=0.35, s=64$ 。模型测试时,将probe原始图像与水平翻转之后图像输入到模型中得到两个512维身份特征向量,进行级联后得到1024维的最终身份特征向量。gallery采用相同的处理方式,最后将probe中每一张图片与gallery集中所有图像进行余弦距离比较,距离从小到大,即top-1到top- n 排序,若gallery集中最近的那个与probe身份相同,则为Rank-1识别准确。

针对(16)式中的超参数,设置了不同的值,对应的 Rank-1 准确率如表 1 所示。

表 1 不同的 λ_1 和 λ_2 所对应的 Rank-1 准确率

Table 1 Recognition rate of Rank-1 for different values of λ_1 and λ_2

λ_1	λ_2	Accuracy of Rank-1 / %
0.01	1	99.03
0.005	1	99.00
0.015	1	98.80
0.01	0.5	98.56
0.02	1	98.67

由表 1 可知,当 λ_1 为 0.01, λ_2 为 1 时, Rank-1 的准确率最高,为 99.03%。当 λ_1 为 0.01, λ_2 为 1 时,训练过程中的曲线变化如图 5 所示,其中 age loss、id loss、total loss 均为训练时的损失值。由测试时的年龄损失值曲线 test age loss 可知,测试时,年龄估计任务的年龄误差在 5 周岁左右。

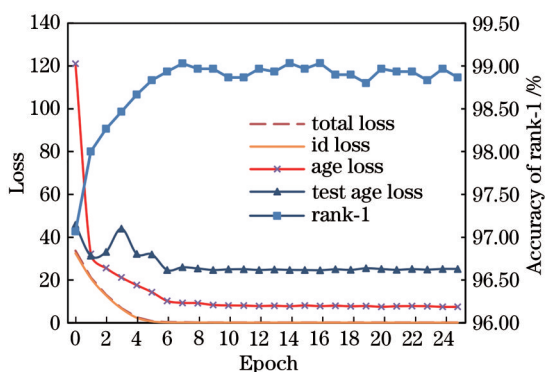


图 5 损失值和 rank-1 准确率在训练过程中的变化曲线
Fig 5 Variation curve of loss value and rank-1 accuracy during training

为了验证改进的 ECBAM 和相关性约束模块 BKCCA 的有效性, CNN 特征提取模块加上 FC1 组成了 Baseline 模型。预训练后, 设置了 CA-CNN 模型、Baseline+基于 ECBAM 的特征分解模块+BCCA 及 Baseline+基于 CBAM 的特征分解模块+BKCCA 的对比实验, 结果如图 6 所示。从图 6 中可以看出, 3 种模型的 Rank-1 准确率最高分别为 99.03%, 98.80%, 98.67%, 充分说明了所提基于 ECBAM 的特征分解模块和 BKCCA 的有效性。

在 Morph Album 2 数据集上对比了近年来主要的跨年龄人脸识别算法的 Rank-1 准确率, 实验结果如表 2 所示。从表中可以看出, 所提 CA-CNN 模型的 Rank-1 准确率高其他模型, 虽然相较于 DAL 模型只有 0.06 个百分点的提升, 但 DAL 模型的训

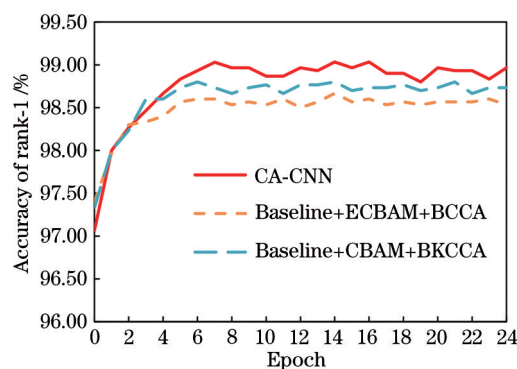


图 6 训练过程中不同模块作用下的 Rank-1 曲线

Fig. 6 Rank-1 curves of different models during training

表 2 不同算法在 Morph Album 2 上的 Rank-1 结果

Table 2 Rank-1 results of different algorithms on Morph Album 2

Algorithm	Accuracy of Rank-1 / %
HFA (2013) ^[10]	91.14
MEFA (2015) ^[11]	93.80
LF-CNN (2016) ^[12]	97.51
AE-CNN (2017) ^[14]	98.13
AFJT-CNN (2018) ^[34]	97.85
JMCNN (2018) ^[26]	98.36
OE-CNN (2018) ^[15]	98.67
DAL (2019) ^[16]	98.97
FSDS-CNN (2020) ^[32]	98.41
Baseline	95.67
Baseline+Age	97.73
Proposed algorithm	99.03

练集是未开源的大规模跨年龄人脸数据集及其他一些通用人脸数据集, 训练集人脸图片数量达到了 170 万张, 而所提 CA-CNN 模型仅在 CASIA-WebFace 数据集上预训练和 Morph Album 2 数据集上微调就可达到 99.03% 的准确率, 训练集照片数量仅有 51 万张, 说明所提 CA-CNN 模型在小训练集下就可达到较好效果。同时设置了对比试验, Baseline 利用 Softmax 损失函数在 CASIA-WebFace 数据集上预训练后, 未在 Morph Album 2 上微调, 此时 Rank-1 准确率仅为 95.67%, 说明年龄对身份识别确实有影响。Baseline+Age 表示加入年龄子网后的模型, 但年龄子网仅有 FC3, 采用多任务学习的方式在 Morph Album 2 进行模型微调, 此时模型的准确率达到 97.73%, 说明多任务学习对身份识别任务有利, 但模型仍有改进空间。图 7 为错误识别的部分人脸图像, 第 1 行为测试的人脸图像, 第 2 行为模型得到的 top-1 图像, 第 3 行为对应的真



图7 检索失败的一些示例

Fig. 7 Some examples of failed retrievals

实结果,每张图像底部数据为年龄。从图7可以看出,识别失败大多是由于姿态、光照等问题。

3.3 在 CALFW 数据集上的实验

CALFW 数据集^[33]是用来评估模型在大年龄间隔下的人脸验证性能的,该数据集挑选了3000对有年龄间隔的相同个体为正样本对,还挑选了3000对具有相同性别和种族的不同个体为负样本对。由于 CALFW 没有年龄标签,在 CASIA-WebFace 数据集上预训练和 Morph Album 2 数据集上微调后,只在 CALFW 数据集上进行人脸验证实验,没有进行任何训练和微调。首先进行人脸图像的对齐

正,测试时,采用与 Morph Album 2 相同的测试方法,将原始图像与水平翻转后的图像输入网络,得到1024维身份特征向量,计算每对身份特征向量的余弦距离作为配对得分,根据错误匹配率(FMR)和错误非匹配率(FNMR)得到等错率(EER)及错误匹配率为0.1时的错误非匹配率[FNMR@FMR(0.1)],EER越低,说明模型的性能越好。

在 CALFW 数据集上对比了近几年主要的跨年龄人脸识别模型,实验结果如表3所示,所提 CA-CNN 模型的 EER 和 FNMR@FMR(0.1) 都优于其他模型,充分说明了所提模型具有较强的泛化能力。

表3 不同算法在 CALFW 数据集上的验证结果

Table 3 Evaluation results of different algorithms on CALFW dataset

Algorithm	Training image size /Million	EER /%	FNMR@FMR(0.1) /%
VGG-Face(2015) ^[35]	2.6	13.5	17.6
Nosiy softmax(2017) ^[36]	0.5	17.5	29.2
AFJT-CNN(2018) ^[34]	0.7	14.8	21.8
FSDS-CNN(2020) ^[32]	0.5	10.1	10.2
Proposed algorithm	0.5	9.8	10.0

4 结 论

针对跨年龄人脸识别,提出了一种新的跨年龄人脸识别模型(CA-CNN),该模型仅需在开源数据集上进行网络训练,就可以达到非常不错的效果。同时,所提 ECBAM 后续也可以考虑应用于 CNN 特征提取模块中,BKCCA 方法也可为多通道网络提供一种新的研究思路;而年龄因素的干扰也可以替

换为其他干扰因素(如光照、表情、角度等),从而解决类似的干扰性问题。

参 考 文 献

- [1] Chen B C, Chen C S, Hsu W H. Face recognition and retrieval using cross-age reference coding with cross-age celebrity dataset[J]. IEEE Transactions on Multimedia, 2015, 17(6): 804-815.

- [2] Best-Rowden L, Jain A K. Longitudinal study of automatic face recognition[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2018, 40(1): 148-162.
- [3] Sawant M M, Bhurchandi K M. Age invariant face recognition: a survey on facial aging databases, techniques and effect of aging[J]. *Artificial Intelligence Review*, 2019, 52(2): 981-1008.
- [4] Agrawal A K, Singh Y N. Evaluation of face recognition methods in unconstrained environments [J]. *Procedia Computer Science*, 2015, 48: 644-651.
- [5] Khiyari H, Wechsler H. Age invariant face recognition using convolutional neural networks and set distances[J]. *Journal of Information Security*, 2017, 8(3): 174-185.
- [6] Mortezaie Z, Hassanpour H. A survey on age invariant face recognition methods[J]. *Jordanian Journal of Computers and Information Technology*, 2019, 5(2): 87-96.
- [7] Wang M, Deng W H. Deep face recognition: a survey [EB/OL]. (2018-04-18) [2020-12-25]. <https://arxiv.org/abs/1804.06655>.
- [8] Lei Z, Zhu P F, Hu Q H, et al. A linear subspace learning approach via sparse coding[C]//2011 International Conference on Computer Vision, November 6-13, 2011, Barcelona, Spain. New York: IEEE Press, 2011: 755-761.
- [9] Li Z F, Park U, Jain A K. A discriminative model for age invariant face recognition[J]. *IEEE Transactions on Information Forensics and Security*, 2011, 6(3): 1028-1037.
- [10] Gong D H, Li Z F, Lin D H, et al. Hidden factor analysis for age invariant face recognition[C]//2013 IEEE International Conference on Computer Vision, December 1-8, 2013, Sydney, NSW, Australia. New York: IEEE Press, 2013: 2872-2879.
- [11] Gong D H, Li Z F, Tao D C, et al. A maximum entropy feature descriptor for age invariant face recognition[C]//2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), June 7-12, 2015, Boston, MA, USA. New York: IEEE Press, 2015: 5289-5297.
- [12] Wen Y D, Li Z F, Qiao Y. Latent factor guided convolutional neural networks for age-invariant face recognition[C]//2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), June 27-30, 2016, Las Vegas, NV, USA. New York: IEEE Press, 2016: 4893-4901.
- [13] Xu C F, Liu Q H, Ye M. Age invariant face recognition and retrieval by coupled auto-encoder networks[J]. *Neurocomputing*, 2017, 222: 62-71.
- [14] Zheng T Y, Deng W H, Hu J N. Age estimation guided convolutional neural network for age-invariant face recognition[C]//2017 IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), July 21-26, 2017, Honolulu, HI, USA. New York: IEEE Press, 2017: 503-511.
- [15] Wang Y T, Gong D H, Zhou Z, et al. Orthogonal deep features decomposition for age-invariant face recognition[M]//Ferrari V, Hebert M, Sminchisescu C, et al. *Computer vision-ECCV 2018. Lecture notes in computer science*. Cham: Springer, 2018, 11219: 764-779.
- [16] Wang H, Gong D H, Li Z F, et al. Decorrelated adversarial learning for age-invariant face recognition [C]//2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), June 15-20, 2019, Long Beach, CA, USA. New York: IEEE Press, 2019: 3522-3531.
- [17] Hotelling H. Relations between two sets of variates [J]. *Biometrika*, 1936, 28(3/4): 321-377.
- [18] Yu J H, Lin Z, Yang J M, et al. Free-form image inpainting with gated convolution[C]//2019 IEEE/CVF International Conference on Computer Vision (ICCV), October 27-November 2, 2019, Seoul, Korea (South). New York: IEEE Press, 2019: 4470-4479.
- [19] Krizhevsky A, Sutskever I, Hinton G E. ImageNet classification with deep convolutional neural networks [J]. *Communications of the ACM*, 2017, 60(6): 84-90.
- [20] Simonyan K, Zisserman A. Very deep convolutional networks for large-scale image recognition[EB/OL]. (2014-09-04) [2020-12-25]. <https://arxiv.org/abs/1409.1556>.
- [21] Szegedy C, Liu W, Jia Y Q, et al. Going deeper with convolutions[C]//2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), June 7-12, 2015, Boston, MA, USA. New York: IEEE Press, 2015: 1-9.
- [22] He K M, Zhang X Y, Ren S Q, et al. Deep residual learning for image recognition[C]//2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), June 27-30, 2016, Las Vegas, NV, USA. New York: IEEE Press, 2016: 770-778.
- [23] Deng J K, Guo J, Yang J, et al. ArcFace: additive angular margin loss for deep face recognition[C]//

- 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), June 15-20, 2016, Long Beach, USA. New York: IEEE Press, 2019: 4685-4694.
- [24] Ioffe S, Szegedy C. Batch normalization: accelerating deep network training by reducing internal covariate shift[EB/OL]. (2015-02-11) [2020-12-25]. <https://arxiv.org/abs/1502.03167>.
- [25] He K M, Zhang X Y, Ren S Q, et al. Delving deep into rectifiers: surpassing human-level performance on ImageNet classification[C]//2015 IEEE International Conference on Computer Vision (ICCV), December 7-13, 2015, Santiago, Chile. New York: IEEE Press, 2015: 1026-1034.
- [26] Yu J B, Jing L P. A joint multi-task CNN for cross-age face recognition[C]//2018 25th IEEE International Conference on Image Processing (ICIP), October 7-10, 2018, Athens, Greece. New York: IEEE Press, 2018: 2411-2415.
- [27] Woo S, Park J, Lee J Y, et al. CBAM: convolutional block attention module[M]//Ferrari V, Hebert M, Sminchisescu C, et al. Computer vision-ECCV 2018. Lecture notes in computer science. Cham: Springer, 2018, 11211: 3-19.
- [28] Wang Q L, Wu B G, Zhu P F, et al. ECA-net: Efficient channel attention for deep convolutional neural networks[C]//2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), June 13-19, 2020, Seattle, WA, USA. New York: IEEE Press, 2020: 11531-11539.
- [29] Akaho S. A kernel method for canonical correlation analysis[EB/OL]. (2006-09-13)[2020-12-25]. <https://arxiv.org/abs/cs/0609071>.
- [30] 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, June 18-23, 2018, Salt Lake City, UT, USA. New York: IEEE Press, 2018: 5265-5274.
- [31] Zhang K P, Zhang Z P, Li Z F, et al. Joint face detection and alignment using multitask cascaded convolutional networks[J]. IEEE Signal Processing Letters, 2016, 23(10): 1499-1503.
- [32] Guo Q Y. Research on cross-age face recognition based on age-invariant identity feature[D]. Nanchang: Jiangxi Normal University, 2020: 37-38.
郭祺玥. 基于年龄无关身份特征的跨年龄人脸识别方法研究[D]. 南昌: 江西师范大学, 2020: 37-38.
- [33] Zheng T Y, Deng W H, Hu J N. Cross-age LFW: a database for studying cross-age face recognition in unconstrained environments[EB/OL]. (2017-08-28) [2020-12-25]. <https://arxiv.org/abs/1708.08197>.
- [34] Li H X, Hu H F, Yip C. Age-related factor guided joint task modeling convolutional neural network for cross-age face recognition[J]. IEEE Transactions on Information Forensics and Security, 2018, 13(9): 2383-2392.
- [35] Parkhi O M, Vedaldi A, Zisserman A. Deep face recognition[C]//Proceedings of the British Machine Vision Conference 2015, September 7-10, 2015, Swansea, UK. Durham: British Machine Vision Association, 2015: 41.1-41.12.
- [36] Chen B H, Deng W H, Du J P. Noisy softmax: improving the generalization ability of DCNN via postponing the early softmax saturation[C] //2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), July 21-26, 2017, Honolulu, HI, USA. New York: IEEE Press, 2017: 4021-4030.