

文章编号:1671-6833(2021)05-0007-06

基于特征子空间直和的跨年龄人脸识别方法

叶继华, 郭祺玥, 江爱文, 黎 欣

(江西师范大学 计算机信息工程学院, 江西 南昌 330022)

摘 要: 针对跨年龄人脸识别任务,在同时进行人脸身份识别和年龄分类这两个任务的多任务卷积神经网络的基础上加入直和模块,提出了一种基于特征子空间直和的多任务卷积神经网络(FSDS-CNN)。该网络利用 2 个并行子网分别从深度特征中提取出身份相关特征和年龄相关特征,并对这 2 个相关特征所对应的特征子空间施加直和约束,使得身份相关特征与年龄相关特征尽可能无关。通过多损失的联合监督学习,该网络可以获得随年龄变化鲁棒的年龄无关人脸身份特征。分别在 Morph Album 2、CACD-VS 和 Cross-Age LFW 数据集上进行实验,其中在 CACD-VS 数据集中,所提方法的 AUC 最优值为 99.7%;在 Cross-Age LFW 数据集中,所提方法在等错误率(EER)和错误匹配率(FMR)为 0.1 时的错误非匹配率(FNMR)上分别取得了最优值 10.1%和 10.2%。同时在 3 个数据集上的实验均进行了消融对比实验以验证直和模块的有效性。实验结果表明,身份特征与年龄特征的相关性被 FSDS-CNN 中的直和模块有效地降低,从而有效提升了模型跨年龄人脸识别的性能。

关键词: 人脸识别; 跨年龄; 多任务; 子空间直和; 特征子空间

中图分类号: TP183 **文献标志码:** A **doi:**10.13705/j.issn.1671-6833.2021.05.002

0 引言

跨年龄人脸识别问题作为人脸识别的一个分支,是对同一人在不同年龄下的人脸进行识别,受到年龄因素的影响,通常存在较大的类内差异和较高的类间相似性^[1],这极大地增加了识别难度。随着时代的发展,跨年龄人脸识别的应用场景变得越来越广,例如将多年前失踪的儿童找回等。因此,研究跨年龄人脸识别问题有着巨大的理论价值和重要的实际意义。针对这一问题,近几年学术界提出了两类方法,即生成类方法和判别类方法。生成类方法通过仿真机制对目标对象人脸骨骼的几何形状和人脸纹理进行建模,从而生成该特定年龄的人脸^[1-2],之后在此基础上进行人脸识别任务。这类方法的不足在于需要很强的参数假设与巨大的计算成本^[3],因此在实际应用中仍然具有很大的局限性。与之相比,学术界将越来越多的重心放在基于判别的方法上。对于判别类方法,尤为重要的一个步骤即人脸特征的提取^[4],其核心是将人脸特征中混合在一起的身份相关

成分和年龄相关成分最大限度地分离,以便于提取出与年龄无关的、对年龄变化具有鲁棒性的身份特征。许多基于深度卷积神经网络的判别方法被提出并取得了显著的成果^[3,5-7]。但這些方法很难将年龄成分彻底从身份相关成分中剔除,因此还有需要改进之处。

为了获得对身份更具有判别性的年龄无关特征,笔者在同时进行人脸身份识别和年龄分类这两个任务的多任务卷积神经网络的基础上引入子空间的直和,提出特征子空间直和的多任务卷积神经网络(feature subspace with direct sum CNN, FSDS-CNN)。实验结果验证了笔者提出的模型在跨年龄人脸识别任务中的有效性和优越性。

1 FSDS-CNN 模型

笔者提出的 FSDS-CNN 模型分别由 3 部分组成:提取深度人脸特征的卷积单元、学习身份相关特征和年龄相关特征并执行分类任务的 2 个子网,以及 1 个用于减少身份相关特征和年龄相关特征之间冗余成分的直和模块。FSDS-CNN 的整

体框架如图 1 所示。

FSDS-CNN 中的卷积单元由典型的 CNN 结构组成,即交替堆叠卷积层、激活层和池化层。在经过卷积单元 $F(\cdot)$ 的编码后,我们可从输入的人脸图像 p 中得到同时包含身份信息和年龄信息的深度人脸特征 x (即 $x = F(p)$),此时 x 不具备年龄无关的特性,需要再做进一步处理。

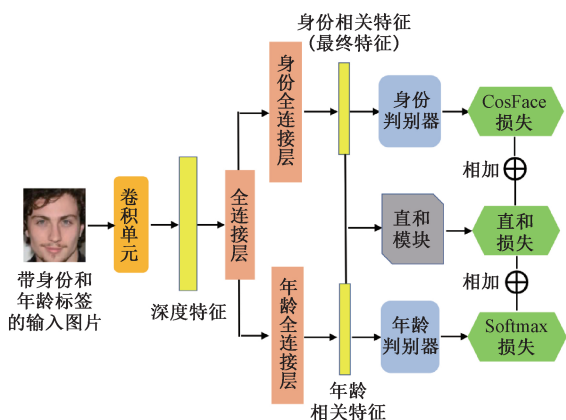


图 1 FSDS-CNN 整体框架

Figure 1 Overall framework of the FSDS-CNN

深度特征在经过第 1 层的全连接层之后,被传给 2 个并行的判别性子网,分别是身份和年龄 2 个子网,可以同时进行身份识别和年龄分类。全连接层和子网对应的判别器组成子网,2 个全连接层分别负责提取出身份相关特征 x_{id} 和年龄相关特征 x_{age} :

$$\begin{aligned} x_{id} &= W_{id}x + b_{id}; \\ x_{age} &= W_{age}x + b_{age}. \end{aligned} \quad (1)$$

式中: W_{id} 、 W_{age} 、 b_{id} 、 b_{age} 分别为 2 个全连接层的权重和偏置,通过 SGD 自动学习得到。之后将得到的 x_{id} 、 x_{age} 分别传给直和模块和相应的判别器进行多任务分类。其中 x_{id} 将作为跨年龄人脸识别任务中的最终特征。有关 2 个判别器和网络结构的具体细节将分别在 1.2 节和 2.1.1 节进行详细介绍。下面先介绍笔者提出的直和模块。

1.1 直和模块

在提取出 x_{id} 和 x_{age} 后,如何保证 x_{id} 在保持身份信息的前提下确保对年龄变化的鲁棒性至关重要。此时 x_{id} 和 x_{age} 之间的较多冗余部分导致了两者之间有着较强的潜在关联性,从而导致跨年龄人脸识别模型的性能下降。提出加入直和模块以对 x_{id} 和 x_{age} 所在空间的特征子空间施加直和约束,通过尽可能减少两者之间的冗余来尽量消除 x_{id} 中的年龄成分,以此降低两者之间的相关性。

记 V_{II} 和 V_{AA} 分别表示 x_{id} 和 x_{age} 所在特征空间,

$X_{ID} = \{x_{id_1}, x_{id_2}, \dots, x_{id_N}\}$ 和 $X_{Age} = \{x_{age_1}, x_{age_2}, \dots, x_{age_N}\}$ 分别表示所有 x_{id} 和 x_{age} 构成的矩阵。分别记 X_{ID} 和 X_{Age} 协方差矩阵的前 K 个最大特征值所对应的特征向量 $v_{id_1}, v_{id_2}, \dots, v_{id_K}$ 和 $v_{age_1}, v_{age_2}, \dots, v_{age_K}$ 构成的特征子空间为 V_{II} 和 V_{AA} , 则 $v_{id_1}, v_{id_2}, \dots, v_{id_K}$ 和 $v_{age_1}, v_{age_2}, \dots, v_{age_K}$ 分别构成 V_{II} 和 V_{AA} 的一组基^[8]。由于在基于子空间的基的形式下,两个子空间直和的充要条件是两个子空间的一组基线性无关^[9]。因此,要实现特征子空间 V_{II} 和 V_{AA} 之间为直和,即 $V_{II} \oplus V_{AA}$, 则必须要 $v_{id_1}, v_{id_2}, \dots, v_{id_K}, v_{age_1}, v_{age_2}, \dots, v_{age_K}$ 满足线性无关,故定义以下直和损失:

$$L_{DS} = \frac{1}{K^2} \sum_{i,j=1}^K |\cos(v_{id_i}, v_{age_j})|. \quad (2)$$

通过训练直和损失将最终收敛,此时有 $V_{II} \oplus V_{AA}$ 。算法 1 展示了直和计算具体过程。通过添加到最终目标函数中,以上直和损失将参与联合监督训练。

算法 1: 特征子空间 V_{II} 和 V_{AA} 的直和计算。

输入: 跨年龄人脸训练数据集

$T = \{I_i | i = 1, 2, \dots, N\}$;

需要的正交特征向量个数 K ;

步骤:

(1) $p \leftarrow 0$;

(2) 随机初始化神经网络参数 $W^{(0)}$;

(3) 通过神经网络得到 $X_{ID}^{(p)}$ 与 $X_{Age}^{(p)}$;

(4) 计算 $X_{ID}^{(p)}$ 协方差矩阵前 K 个最大特征值对应的正交特征向量 $v_{id_1}^{(p)}, v_{id_2}^{(p)}, \dots, v_{id_K}^{(p)}$, 并构成 $V_{II}^{(p)}$, 计算 $X_{Age}^{(p)}$ 协方差矩阵前 K 个最大特征值对应的正交特征向量 $v_{age_1}^{(p)}, v_{age_2}^{(p)}, \dots, v_{age_K}^{(p)}$, 并构成 $V_{AA}^{(p)}$;

(5) 通过式 (2) 计算直和损失 L_{DS} ;

(6) 更新神经网络参数 $W^{(p)}$;

(7) $p \leftarrow p + 1$;

(8) 返回第 (3) 步直到 L_{DS} 收敛;

输出: $V_{II} \oplus V_{AA}$ 。

1.2 多任务训练

1.2.1 身份判别器

为确保学到的特征对身份具有较强判别性,遵循 Wang 等^[7]的设置使用 CosFace 损失对 x_{id} 的学习进行监督。CosFace 损失可表示为下式:

$$L_{ID} = \frac{1}{N} \sum_{i=1}^N -\lg \frac{e^{s(\cos \theta_{y_i, i-m})}}{e^{s(\cos \theta_{y_i, i-m})} + \sum_{j \neq y_i} e^{s \cos \theta_{j, i}}}. \quad (3)$$

式中: y_i 为第 i 个样本对应的身份标签; N 为样本数; s 和 m 均为超参数, 分别为常数缩放因子和用来控制余弦间隔的常数间隔项; $\cos \theta_{j,i} = \frac{\mathbf{U}_j^T \mathbf{x}_{id}^i}{\|\mathbf{U}_j\| \|\mathbf{x}_{id}^i\|}$, 为第 i 个特征 \mathbf{x}_{id}^i 和身份分类器第 j 个权重向量 \mathbf{U}_j 之间夹角的余弦。

1.2.2 年龄判别器

由于年龄标签是粗略的且在实际应用中存在不确定的噪声, 对于年龄信息的学习, 在年龄分类时遵循 Li 等^[6] 的协议把年龄标签分割为多个无重复的年龄组。笔者使用 Softmax 函数作为年龄判别器的损失函数, 用下式表示:

$$L_{Age} = \frac{1}{N} \sum_{i=1}^N -\lg \frac{e^{\mathbf{V}_j^T \mathbf{x}_{age}^i + \alpha_j}}{\sum_{j=1}^n e^{\mathbf{V}_j^T \mathbf{x}_{age}^i + \alpha_j}} \quad (4)$$

式中: n 为年龄组个数; \mathbf{c}_i 为第 i 个样本所对应的年龄组标签; \mathbf{V}_j 和 α_j 分别为年龄分类器的第 j 个权重向量和其对应的偏置。

综上, FSDS-CNN 的训练由如下多任务目标函数联合监督:

$$L = L_{ID}(\mathbf{x}_{id}) + \lambda_1 L_{Age}(\mathbf{x}_{age}) + \lambda_2 L_{DS}(\mathbf{x}_{id}, \mathbf{x}_{age}) \quad (5)$$

式中: λ_1 和 λ_2 为权重超参数, 用来平衡这 3 个损失。

2 实验和分析

2.1 模型实现细节

2.1.1 网络结构

①卷积单元: 卷积单元由 5 个卷积块组成, 多个卷积层组成卷积块, 采用 3×3 、步幅为 1 的卷积核, 之后接上批归一化使模型快速收敛, PReLU 作为激活函数, 以上 5 个卷积块输出的特征图个数分别为 64、128、256、512、1 024。②判别性子网: 深度特征 \mathbf{x} 经过全连接层后的输出是 1 024 维的特征, 之后分别传给 2 个判别性子网的全连接层以输出 1 024 维的 \mathbf{x}_{id} 和 \mathbf{x}_{age} , 在 \mathbf{x}_{id} 和 \mathbf{x}_{age} 上分别进行 PReLU 和批归一化操作后再传给相应的判别器进行分类任务。

2.1.2 数据预处理

使用 MTCNN^[10] 检测人脸, 之后根据人脸关键点通过相似度变换将人脸图像尺寸裁剪为 112×96 。对裁剪后的人脸图像的每个像素进行归一化。

2.1.3 训练细节

预训练阶段使用的训练数据集为只包含身份信息信息的 CASME WebFace^[11], 大约包含 49 万张人

脸样本, 整个预训练阶段直和模块和年龄子网不参与训练, 并在数据集上进行大约 14 轮, 批尺寸大小设置为 128。式 (5) 和式 (3) 中的超参数按照文献[7] 设置为 $\lambda_1 = 0.1$, $m = 0.35$, $s = 64$; 另外, 式 (5) 和式 (2) 中的 λ_2 和 K 分别设置为 10 和 25, 具体设置细节将在 2.2 节中详细介绍。

2.1.4 模型测试细节

在著名的公开领域的跨年龄人脸数据集 Morph Album 2^[12] 和 CACD-VS^[13] 上进行了人脸识别和验证实验来评估本模型的性能, 并在数据集 Cross-Age LFW^[14] 上进一步评估了本模型的泛化能力。在测试过程中, 为了增强数据的特征, 将原始图像与翻转图像的 \mathbf{x}_{id} 进行级联作为最终特征, 再计算每对特征的余弦相似度作为配对得分。

2.2 在 Morph Album 2 数据集上的实验

Morph Album 2 是一个由 13 000 个个体的 55 000 张人脸图片组成的跨年龄人脸数据集。遵循文献[6] 的设置将整个数据集划分为训练集和测试集, 训练集和测试集的样本身份无重合。

笔者对直和模块的特征向量个数 K 和直和损失权重 λ_2 这两个重要超参数的设置进行了探索实验。当 $\lambda_2 = 1$, K 取不同的值时, 模型的识别性能如图 2 所示。

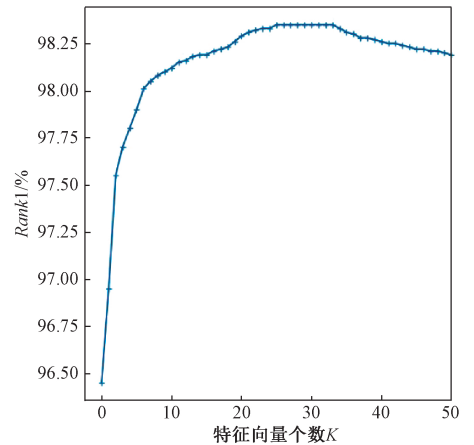


图 2 Morph Album 2 上不同 K 下的识别率

Figure 2 Rank1 identification rate changes with different K on Morph Album 2

从图 2 中看出, 当 K 取前 7 个值时模型的识别率快速增长, 因为此时直和模块对 \mathbf{x}_{id} 和 \mathbf{x}_{age} 中在跨年龄人脸识别中起主要作用的主成分特征如五官形状特征被直和约束, 两者的相关性被有效减少, 其中 \mathbf{x}_{id} 的年龄无关性被提高。识别率随着 K 的增加逐渐上升, 在 $K = 25$ 时识别率升至最高值 98.35%。之后在 K 为 25 ~ 34 时识别率不变, 在 K 升至 34 时识别率开始下降, 根据式 (2), 此时

\mathbf{x}_{id} 和 \mathbf{x}_{age} 中的一些在跨年龄人脸识别中的次要成分特征向量作为噪音被加入,这不利于主要成分特征向量之间的正交关系,降低了模型的识别性能。因此,在实验中设置 K 的值为 25。

当 $K = 25$, λ_2 取 0.001、0.01、0.1、1、10、100 时实验效果如图 3 所示。可以看出,当 λ_2 很小时,模型的识别率较低,由式(5)可知,直和约束因为直和损失的权重偏小而作用不明显。之后模型的识别率随着 λ_2 取值增大而逐渐增大,当 λ_2 为 10 时识别率达到最高值 98.41%。但当 λ_2 取值过大时识别率开始下降,原因为直和约束的作用过于明显,从而降低了身份损失和年龄损失的重要性,导致 \mathbf{x}_{id} 和 \mathbf{x}_{age} 的学习优先级降低从而不能很好地进行学习。因此,之后实验 λ_2 取 10。

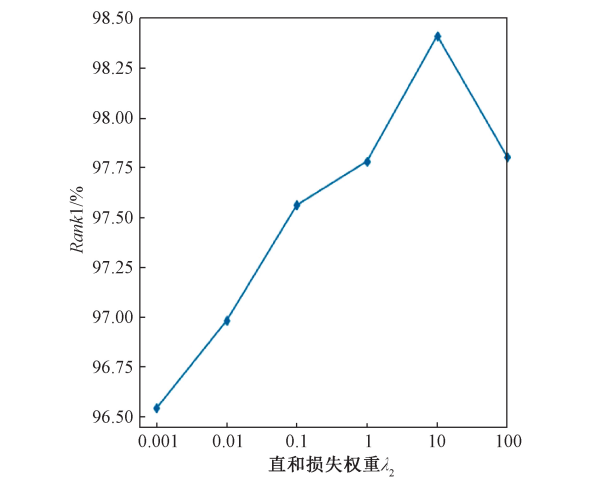


图 3 Morph Album 2 上不同 λ_2 下的识别率

Figure 3 Rank1 identification rate changes with different λ_2 on Morph Album 2

确定了两个超参数分别为 $\lambda_2 = 10$ 、 $K = 25$ 后,在 Morph Album 2 的测试集上进行了人脸识别实验并与近年来的一些主流方法进行对比,实验结果如表 1 所示。

表 1 在 Morph Album 2 数据集上的评估结果

Table 1 Evaluation results on the Morph Album 2 dataset

方法	Rank1/%
CARC ^[15]	92.80
MEFA ^[16]	93.80
MEFA+SIFT+MLBP ^[16]	94.59
LPS+HFA ^[17]	94.87
LF-CNNs ^[3]	97.51
AE-CNN ^[18]	98.13
AFJT-CNN ^[6]	97.85
OE-CNNs ^[5]	98.55
DAL ^[7]	98.93
CNN-baseline	91.83
CNN-baseline+Age	96.52
本文模型	98.41

本实验使用 Rank1 作为评价指标。从表 1 中可知,笔者提出的 FSDS-CNN 模型具有显著的优势,但是与 OE-CNNs 和 DAL 相比存在一定的差距,这是因为本模型和以上两个模型所用的训练集数量(分别为 53 万和 170 万)差距过大所致。另外,笔者实验中 batch-size 取值为 128,而 OE-CNNs 和 DAL 则为 512。为了验证直和模块的作用,进行有无直和模块情况的模型的实验。从表 1 中看出, CNN-baseline 在只用 CASIA-WebFace 一个数据集进行预训练时识别率只达到 91.83%,相较于 CARC 和 MEFA 略低;用 Morph Album 2 进行微调后, CNN-baseline + Age 的识别率增高到 96.52%,这是因为身份特征和年龄特征因年龄子网的加入而被分离,但是因为身份特征与年龄特征并没有约束的存在,模型的性能相比于 LF-CNNs 和 AE-CNN 略差。FSDS-CNN 在以上基础上加入了直和模块用来对身份特征和年龄特征所在空间的子空间施加直和约束,将模型的识别性能提升至 98.41%,证实了直和模块的显著效果。

2.3 在 CACD-VS 数据集上的实验

CACD 数据集作为公开领域的基准老化数据集常被用于年龄不变人脸识别,该数据集包含 2 000 个名人的具有年龄变化的 163 446 张人脸图片,在本实验中作为训练集对模型进行微调。遵循文献[3]和文献[6]实验设置,在 CACD 的子集 CACD-VS 上进行了测试。在本实验中,笔者将 CACD-VS 中的所有样本从训练集中剔除。

表 2 列出了本模型在 CACD-VS 数据集上与其他最新的年龄不变人脸识别方法的评估结果。本实验使用人脸识别中常用的精确度和 AUC 这两个评价指标。从表 2 可知,除了 DAL, FSDS-CNN 的精确度均不小于其他模型, AUC 则均优于其他模型,图 4 显示了多种方法的 ROC 曲线图。在反映模型性能的真实性和稳健性方面, AUC 通常优于精确度^[19],进一步证实了本模型的优越性。

2.4 在 Cross-Age LFW 数据集上的实验

笔者同样在 Cross-Age LFW^[14](CALFW)数据集上进行了验证实验以表明本模型的泛化能力。CALFW 数据集包含 4 025 个人,每个人有 2、3 或 4 张图片。因为 CACD 数据集中有噪音的存在,故使用 Morph Album 2 作为训练集对模型进行微调。由于该实验在训练和微调阶段均无 CALFW 的参与,故能有效验证模型的泛化能力。

表 3 列出了本模型在 CALFW 数据集上与其

表 2 在 CACD-VS 数据集上的评估结果

Table 2 Evaluation results on the CACD-VS dataset

方法	精确度/%	AUC/%
HD-LBP ^[20]	81.6	88.8
CARC ^[15]	87.6	94.2
文献[13]方法1	85.7	94.6
文献[13]方法2	94.2	99.0
LF-CNNs ^[3]	98.5	99.3
AFJT-CNN ^[6]	99.0	99.4
OE-CNNs ^[5]	99.2	99.5
DAL ^[7]	99.4	99.6
CNN-baseline	98.4	99.4
CNN-baseline+Age	98.8	99.5
FSDS-CNN	99.2	99.7

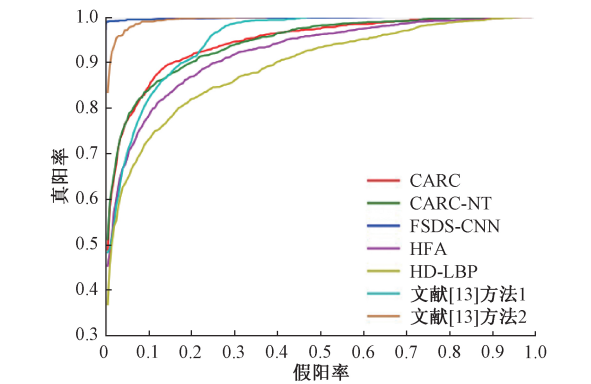


图 4 不同方法在 CACD-VS 上的 ROC 曲线

Figure 4 ROC curves of different methods on CACD-VS

表 3 在 CALFW 数据集上的评估结果

Table 3 Evaluation results on the CALFW dataset

方法	训练数据大小/M	EER/%	FNMR/%
VGG-Face ^[21]	2.6	13.5	17.6
Noisy Softmax ^[22]	0.5	17.5	29.2
AFJT-CNN ^[6]	0.7	14.8	21.8
CNN-baseline	0.5	15.1	19.4
CNN-baseline+Age	0.5	14.6	18.0
FSDS-CNN	0.5	10.1	10.2

他方法相比的验证性能。本实验采用的评价指标是 EER 和错误匹配率 (FMR) 为 0.1 时的错误非匹配率 ($FNMR$)。从实验结果可知,本模型与其他模型相比具有明显的优良性,尤其是在较小的训练集上,仍然有较好的性能,充分证实了本模型泛化能力的强大。

3 结论

提出了一种新颖网络称为特征子空间直和的多任务卷积神经网络 (FSDS-CNN) 用来处理跨年龄的人脸识别问题,并在 Morph Album 2、CACD-VS 和 CALFW 这 3 个基准老化数据集上进行了

跨年龄的人脸识别与验证实验。通过和近几年的多种具有代表性的方法进行对比,本模型在评价指标 AUC 上取得了最优值 99.7%,在 EER 和错误匹配率 (FMR) 为 0.1 时的错误非匹配率 ($FNMR$) 上分别取得了最优值 10.1% 和 10.2%。实验结果验证了直和模块的有效性,表明了本模型相比于其他年龄不变人脸识别方法具有明显的优越性。

参考文献:

[1] 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.

[2] 董锁芹. 基于生成对抗网络的跨年龄人脸识别技术研究[D]. 长春: 长春理工大学, 2019.

[3] 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). Piscataway: IEEE, 2016: 4893–4901.

[4] 苏士美, 王燕, 王明霞. 基于加权小波分解的人脸识别算法研究[J]. 郑州大学学报(工学版), 2014, 35(1): 5–9.

[5] WANG Y T, GONG D H, ZHOU Z, et al. Orthogonal deep features decomposition for age-invariant face recognition [C]//Computer Vision-ECCV 2018. Munich, Germany: ECCV, 2018: 11219.

[6] 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.

[7] 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). Piscataway: IEEE, 2019: 3522–3531.

[8] 叶继华, 万叶晶, 刘长红, 等. 基于多子空间直和特征融合的人脸识别算法[J]. 数据采集与处理, 2016, 31(1): 102–107.

[9] 孙宗明, 李振国, 梅门昌. 维数公式与子空间直和的等价条件[J]. 长沙大学学报, 1998(2): 47–50.

[10] 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.

[11] YI D, LEI Z, LIAO S C, et al. Learning face representation from scratch[EB/OL]. (2014–11–28) [2020–08–

- 31]. <https://arxiv.org/pdf/1411.7923.pdf>.
- [12] RICANEK K, TESAFAYE T. MORPH: a longitudinal image database of normal adult age-progression[C]//7th International Conference on Automatic Face and Gesture Recognition (FGR06). Piscataway: IEEE, 2006: 341-345.
- [13] 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.
- [14] 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) [2020-08-31]. <https://arxiv.org/pdf/1708.08197.pdf>.
- [15] CHEN B C, CHEN C S, HSU W H. Cross-age reference coding for age-invariant face recognition and retrieval[C]//Computer Vision-ECCV 2014. Cham: Springer, 2014: 768-783.
- [16] 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). Piscataway: IEEE, 2015: 5289-5297.
- [17] LI Z F, GONG D H, LI X L, et al. Aging face recognition: a hierarchical learning model based on local patterns selection[J]. IEEE transactions on image processing, 2016, 25(5): 2146-2154.
- [18] 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). Piscataway: IEEE, 2017: 503-511.
- [19] LING C X, HUANG J, ZHANG H. AUC: a better measure than accuracy in comparing learning algorithms[C]//Advances in Artificial Intelligence, 16th Conference of the Canadian Society for Computational Studies of Intelligence, AI 2003. Cham: Springer, 2003: 1-25.
- [20] CHEN D, CAO X D, WEN F, et al. Blessing of dimensionality: high-dimensional feature and its efficient compression for face verification[C]//2013 IEEE Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE, 2013: 3025-3032.
- [21] PARKHI O M, VEDALDI A, ZISSERMAN A. Deep face recognition[C]//British Machine Vision Conference 2015. Swansea, UK: BMVA, 2015: 1-12.
- [22] CHEN B H, DENG W H, DU J P. Noisy softmax: improving the generalization ability of DCNN via postponing the early softmax saturation[C]//IEEE Conference on Computer Vision and Pattern Recognition 2017. Piscataway: IEEE, 2017: 4021-4030.

Cross-age Face Recognition Method Based on Feature Subspace Direct Sum

YE Jihua, GUO Qiyue, JIANG Aiwen, LI Xin

(College of Computer Information Engineering, Jiangxi Normal University, Nanchang 330022, China)

Abstract: To solve cross-age face recognition tasks, this paper introduces the direct sum module on the basis of the multi-task convolutional neural network that simultaneously performs two tasks of face recognition and age classification, and proposes the feature subspace with direct sum multi-task convolutional neural network (FSDS-CNN). The network uses two parallel subnets to extract the identity-related feature and age-related feature from the deep feature, then the direct sum constraint is applied to the feature subspaces corresponding to these two related features, so that the correlation between identity-related feature and age-related feature is decreased as much as possible. Through the joint supervised learning of multiple loss functions, the network can obtain age-invariant face identity feature that is robust with age. Cross-age face recognition and verification experiments is conducted on three datasets (Morph Album 2, CACD-VS and Cross-Age LFW). In the CACD-VS dataset, the proposed method achieves the optimal result of 99.7% on the evaluation metric of *AUC*; in the Cross-Age LFW dataset, the method respectively achieves the optimal results of 10.1% and 10.2% on the evaluation metric of *EER* and *FNMR* when *FMR* is 0.1. At the same time, the ablation comparison experiments are conducted on the three datasets to verify the effectiveness of the direct sum module. The results show that the correlation between identity features and age features is effectively reduced by the direct sum module in FSDS-CNN, and then effectively improves the performance of cross-age face recognition.

Key words: face recognition; cross-age; multi-task; subspace direct sum; feature subspace