

DOI:10.19651/j.cnki.emt.2107612

基于深度对抗丢弃正则化的年龄估计*

朱昱¹ 樊航¹ 王鹏² 马莞悦² 周媛³(1.南京信息工程大学长望学院 南京 210044; 2.南京信息工程大学电子与信息工程学院 南京 210044;
3.南京信息工程大学人工智能学院(未来技术学院) 南京 210044)

摘要: 成年人面部变化非常缓慢,因此相邻年龄段的成人年龄估计仍是一个挑战。针对该问题,将对抗学习思想引入年龄估计任务,提出了基于深度对抗丢弃正则化的年龄估计模型。通过年龄特征学习器与判别器的对抗训练,提升年龄特征学习器对年龄段特征(特别是对相邻年龄段人脸年龄特征)的学习能力。在3个经典数据集(UTKFace、MORPH和Adicnce)上的实验显示,所提出的模型将UTKFace数据集的预测正确率由42.8%提升至81.6%,MORPH数据集的准确率由39.8%提升至69.8%,对Adicnce数据集的预测正确率为63.3%;和其他4个经典模型相比,该模型仅用5层神经网络就达到了比深层神经网络更好的效果,特别中青年年龄段(15~53岁)年龄估计准确率比其他模型平均高出17.5%,说明本文模型对年龄估计任务性能有显著提升,有很好的实用价值。

关键词: 年龄估计;对抗学习;对抗丢弃正则化;卷积神经网络

中图分类号: TP391 **文献标识码:** A **国家标准学科分类代码:** 520.2

Age estimation based on deep adversarial dropout regularization

Zhu Yu¹ Fan Hang¹ Wang Peng² Ma Wanyue² Zhou Yuan³(1. Changwang School of Honors, Nanjing University of Information Science and Technology, Nanjing 210044, China;
2. School of Electronics & Information Engineering, Nanjing University of Information Science and Technology,
Nanjing 210044, China; 3. School of Artificial Intelligence/School of Future Technology, Nanjing University of
Information Science and Technology, Nanjing 210044, China)

Abstract: For adults' facial appearances changing slowly, the age estimation of adults in adjacent age groups is still a challenge. Aiming at this problem, this paper introduced the adversarial training method into the age estimation and proposed an age estimation method based on adversarial dropout regularization(ADR). The age feature learner and the discriminator are trained via the adversarial training method, then the ability of age feature learning(especially the adjacent age groups features) gets improved. Experimental results on three classic datasets (UTKFace, MORPH and Adience) show that the proposed model improves the accuracy of UTKFace from 42.8% to 81.6%, and improves the accuracy of MORPH from 39.8% to 69.8%. Moreover, the accuracy of Adience is 63.3%. Being compared with other 4 classic models, the model using the neural networks of 5 layers achieves better results than other deep neural networks, and outperforms other methods with averagely 17.5% higher accuracy for the young and middle-aged(15~53 years old), which shows that our model improves the performance significantly on age estimation task, and has the practical value.

Keywords: age estimation; deep learning; adversarial dropout regularization; convolutional neural network

0 引 言

年龄信息作为人类的一种重要生物特征,除了对身份认证提供有效的信息外,也影响着年龄人脸识别系统的性能^[1-2],因此对年龄进行估计有着重要作用。年龄估计(age estimation)即根据人脸图片的面部年龄特征对该人脸所属

年龄或年龄段进行估计,可被用于网络安全(如未成年人访问网络内容控制)或门禁控制系统^[3-5],在电子客户关系管理(ECRM)中通过摄像头采集人脸图像并自动进行年龄估计,可以实现对各年龄段顾客的消费特点和产品喜好进行自动、无干扰地收集与分析,从而便于对不同年龄段的顾客进行产品开发并提供有针对性的营销服务^[6-7]。因其有很

收稿日期:2021-08-17

* 基金项目:江苏省大学生创新训练重点项目(201910300044z)、2019年第一批产学研合作协同育人项目(201901134029)资助

大的应用前景,近年来人脸年龄估计受到了国内外学者的广泛关注,成为计算机视觉领域的一个研究热点。

典型的年龄估计模型是基于学习到的人脸年龄特征对年龄回归或对年龄段进行分类,可见人脸年龄特征学习是年龄估计性能的决定因素之一。经典的人脸年龄估计模型有人体测量学模型^[8]、主动外观模型(active appearance models, AAM)^[9]、年龄成长模式子空间(aging pattern subspace, AGES)^[10]、年龄流形(age manifold)^[11],这些方法主要通过对面脸骨骼、皮肤等年龄因素建立复杂的生物模型来表示人脸年龄特征或者手工提取人脸年龄特征,建模过程复杂、计算复杂度高。

深度学习被用于图像相似度检测^[12]、人脸检测^[13]、亲属关系验证^[14]等,其发展使得端到端的年龄估计成为可能。2015年,Levi等^[15]用深度卷积神经网络(deep convolutional neural network, DCNN)首次在非受限条件的数据集上进行年龄分类,在 Adience 上预测准确率为 50.7%,误差在一个年龄段内的准确率为 84.7%;2016年,Rothe等^[16]提出了深度期望模型(deep expectation, DEX),利用 VGG 网络对年龄进行估计,该方法将 Adience 上的估计准确率提升至 55.6%,误差在一个年龄段以内的准确率为 89.7%;2017年,Dehghan等^[17]利用 Sighthound 新型深度神经网络架构加上大量数据训练,最终取得了比较好的年龄分类效果,在 Audience 上的准确率为 61.3%。同年 Hou等^[18]利用 SAAF 微调了 VGG 网络在 Adience 上取得了 61.3%准确度,误差一个年龄段以内的准确率为 95.1%。2019年,Mallouh等^[19]通过减少 VGG 网络输出层的维度来预训练网络,最终在 Adience 数据集上的预测准确率为 60.6%,误差一个年龄段以内的准确率为 92.3%,该文章同时提出了一种组合模型,通过将不同 CNN 网络组合(VGG-Face + GoogleNet + ResNet-50 + VGG-VD-16 + VGG-VD-19 + FNC-8s),结合减少维度的方法,最终在 Adience 数据集上的预测准确率为 62.3%,误差一个年龄段以内的准确率为 92.6%。

值得注意的是,成年后的人面部变化缓慢,表现为相邻年龄段的人脸年龄特征相似,相邻年龄段人脸年龄估计成为一个挑战,目前年龄估计算法都没有考虑这个问题^[20]。为解决相邻年龄段人脸年龄估计问题,本文将对抗训练思想引入年龄估计任务,提出了一种基于对抗丢弃正则化的年龄估计模型并给出了算法步骤。通过在 UTKFace、MORPH 数据集和 Adience 数据集上的实验表明,本文模型和单纯的深度学习模型相比,在 UTKFace 数据集上的预测正确率由 42.8%提升至 81.6%,在 MORPH 数据集上的准确率由 39.8%提升至 69.8%;和其他 4 种常用方法相比,本文模型对于 Adience 数据集预测的准确率为 63.3%,达到了常用方法的效果,且在青中年年龄段(15~53岁)估计上本文模型的准确率比较文献[19]中的方法年龄估计准确率平均提高 17.5%。实验结果表明通过对抗

学习过程强化了年龄特征学习器对决策边界点的特征学习能力,增强了年龄估计性能(特别是对相邻年龄段估计的准确度)。

1 对抗丢弃正则化

在传统的对抗学习方法中(以对抗生成网络为例),分类器只是判断生成器生成的特征在不在各个类中^[21],而并没有考虑到分类边界上样本特征,这可能导致生成器生成模棱两可的特征,从而降低分类的准确率。对抗丢弃正则化(adversarial dropout regularization, ADR)^[22]是 Kuniaki 等在 2018 年提出的一种新的对抗学习方法,因其对于分类边界点的特征有较好的学习能力而被用于域迁移领域,该方法通过对抗训练的方式训练生成网络 G 与判别网络 C 。

图 1 所示为丢弃正则化^[23]的说明。 x_s 是来自源域的样本, $G(x_s)$ 表示通过特征学习器 G 得到的 x_s 的特征。首先通过监督学习同时训练 G 和 C 。然后,对于判别器 C ,通过设置 Dropout 系数获得两个不同的判别网络 C_1 和 C_2 ,将目标域样本 x_t 的特征 $G(x_t)$ 分别输入到 C_1 和 C_2 得到两个输出 $C_1(G(x_t))$ 和 $C_2(G(x_t))$,相应的后验概率表示 $P_1(y|x_t)$ 和 $P_2(y|x_t)$,缩写为 P_1 和 P_2 。

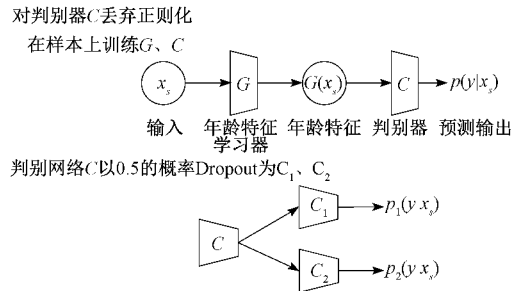


图 1 丢弃正则化说明

图 2 为丢弃正则化对抗训练流程,点状阴影模块表示固定不训练。训练分为两步:第 1 步,固定特征学习器 G ,通过最大化 C_1 和 C_2 的差异 $d(p_1, p_2)$ 训练 C_1, C_2 以提高分类器效果;第 2 步,固定判别器 C_1, C_2 ,通过最小化它们之间的差异来增强特征学习器 G 的性能。

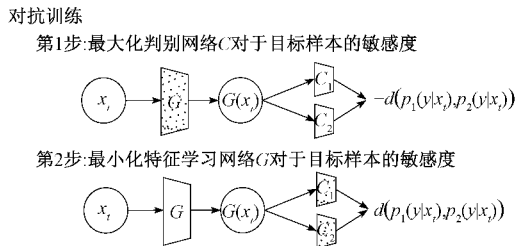


图 2 对抗训练流程

其中,两个分类器的差异 $d(p_1, p_2)$ 用 KL 散度衡量,如式(1)所示。

$$d(p_1, p_2) = \frac{1}{2}(D_{KL}(p_1 | p_2) + D_{KL}(p_2 | p_1)) \quad (1)$$

2 本文方法

2.1 模型描述

针对成年人相邻年龄段之间人脸年龄面部特征变化不明显的问题,本文提出一种基于对抗丢弃正则化的年龄估计模型,图 3 为模型结构。

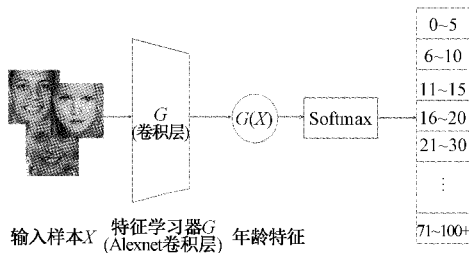


图 3 基于对抗丢弃正则化的年龄估计模型

模型由两部分组成:年龄特征学习器 G 和 Softmax 分类器。将 Alexnet^[24] 作为年龄特征学习器 G , 人脸图片 X 输入到年龄特征学习器 G 后得到年龄特征 $G(X)$, 通过 Softmax 分类器对该人脸年龄进行估计。模型输出根据实际年龄估计任务设置(如此处将年龄分成了 10 个年龄段, 分别对应 0~5 岁, 6~10 岁, 11~15 岁等)。

2.2 模型训练步骤

本模型训练分为 3 步进行。

第 1 步:训练分类器 C 和年龄特征学习器 G 。输入样本 x , 其年龄标签为 y , 训练样本集记为 X , 训练样本标签集记为 Y , 为了正确获取年龄特征信息对样本进行分类, 目标函数如式(2)所示。

$$\min_{C,G} L(X, Y) = -E_{x,y \sim X, Y} \sum_{k=1}^K 1_{k=y} \log C(G(x))_k \quad (2)$$

其中, K 为年龄段数目, $C(G(x))_k$ 为样本 x 被分类到第 k 个年龄段的概率, $1_{k=y}$ 表示样本年龄估计正确输出 1。

第 2 步:固定年龄特征学习器 G , 训练分类器 C 。通过 Dropout 将分类器 C 分为 C_1 和 C_2 , 将同一特征 $G(x)$ 分别输入到 C_1 和 C_2 , 得到两个输出分别为 $C_1(G(x))$ 、 $C_2(G(x))$, 它们的差异 L_{adv} 定义如式(3)所示。

$$L_{adv}(X) = E_{x \sim X} [d(C_1(G(x)), C_2(G(x)))] \quad (3)$$

为增强分类器 C 性能, 目标最大化式(3), 因此本步骤目标函数表示为式(4):

$$\min_C L(X, Y) - L_{adv}(X) \quad (4)$$

第 3 步:固定分类器 C , 训练年龄特征学习器 G 。使得 G 对同一输入的判别差异 L_{adv} 最小化, 因此目标函数如式(5)所示。

$$\min_G L_{adv}(X) \quad (5)$$

3 实 验

3.1 数据集及模型参数说明

本文选用经典人脸数据集 UTKFace^[25]、MORPH^[26]、

Adience^[27] 进行实验, 其中 UTKFace 数据集样本年龄为 1~100 岁, MORPH 数据集中样本年龄为 18~70 岁, Adience 数据集中样本年龄为 1~100 岁。按照留出法, 实验过程中从每个年龄段中随机取 20%(100 张)的图片作为测试集, 剩余 80% 图片作为训练集。

为保证各年龄段数据分布均匀, 通过随机旋转、亮度调整、对比度调整等方式对数据进行了增强, 最终使训练集中每个年龄段都有 10 000 张训练样本。实验在华为云 ModelArts 上进行, 采用 Python3.6.4+Tensorflow1.14.0 框架。训练时 batchsize 设置为 64, Dropout 设置为 0.5。学习率的设置根据数据集不同有所调整, 以 Adience 为例, 前 5 000 次迭代学习率均设置为 0.005, 之后以 0.001 的学习率迭代 15 000 次, 接着以 0.0005 的学习率迭代 30 000 次, 最后以 0.0001 的学习率迭代 60 000 次。

3.2 实验结果及分析

本文模型年龄特征学习器采用 AlexNet 为主干网络, 为考察对抗训练对年龄估计的作用, 先在 UTKFace、MORPH 数据集上分别用本文模型和未使用对抗训练思想的 AlexNet 进行了实验, 然后在 UTKFace、MORPH 和 Adience 数据集上将本文模型与其他 4 个经典年龄估计方法进行性能比较。本文选择准确率作为模型性能评价指标, 基于混淆矩阵计算出各年龄段的预测准确率与总准确率。

1) UTKFace 数据集年龄估计性能

将 UTKFace 样本根据年龄段分为 10 个子集, 分别对应 0~5、6~10、11~15、16~20、21~30、31~40、41~50、51~60、61~70 以及 70+ 年龄段。

表 1、2 分别为在 UTKFace 数据集上本文模型和 AlexNet 模型的年龄估计混淆矩阵。针对表 1 中本文模型将 70~100 年龄段样本误判为 30~40 岁的情况, 选取了部分被判断错误的老年人图片(如图 4 所示)以及部分 31~40、61~70 年龄段的图片(如图 5、6 所示)进行分析。对比表 1 和 2, 可看出引入对抗训练思想的本文模型相对于 AlexNet 模型, 对各个年龄段进行年龄估计的准确率均有显著提升, 本文模型的“1-off”准确率均达到 90% 以上, 这说明对抗学习确实能够加强年龄特征学习器的性能。

通过观察, 如图 4(a) 和 (b) 所示被误判的老年人因其妆容等因素导致其面部年龄特征与低年龄段较为相近, 或如图 4(c) 所示图片作为训练数据对应的年龄标签错误, 从而导致了误判。值得一提的是, 图 4(a) 实际年龄就是 30~40 岁之间, 这也从另一个角度证明了本文模型对年龄特征学习有效性。

表 3 为本文模型和 AlexNet 在 UTKFace 数据集上准确率, 其中“Exact”表示年龄预测完全正确的比例, “1-off”为预测年龄与实际年龄相差一个年龄段的比例, 都是越高越好。

表 1 本文模型在 UTKFace 数据集上年龄估计混淆矩阵(粗体表示预测正确的数目)

年龄段	预测结果									
	0~5	6~10	11~15	16~20	21~30	31~40	41~50	51~60	61~70	70~100+
0~5	94	4	0	0	1	0	0	1	0	0
6~10	4	90	2	2	1	0	0	1	0	0
11~15	2	7	82	5	3	1	0	0	0	0
16~20	1	0	0	88	8	1	1	1	0	0
21~30	1	0	0	6	69	15	3	5	0	1
31~40	0	0	0	1	19	69	5	5	0	1
41~50	0	0	0	2	5	13	65	14	0	1
51~60	0	0	0	0	1	0	6	87	3	3
61~70	0	0	0	0	0	0	3	13	79	4
70~100+	0	0	0	0	0	1	2	3	1	93

表 2 AlexNet 对 UTKFace 数据集预测结果的混淆矩阵(粗体表示预测正确的数目)

年龄段	预测结果									
	0~5	6~10	11~15	16~20	21~30	31~40	41~50	51~60	61~70	70~100+
0~5	80	14	3	1	1	1	0	1	0	0
6~10	16	45	17	13	6	2	0	0	0	1
11~15	4	28	29	18	14	2	1	2	1	1
16~20	2	5	9	28	40	12	0	3	0	1
21~30	1	0	4	14	53	16	6	3	2	1
31~40	0	1	3	6	24	36	13	15	1	1
41~50	0	0	0	3	10	28	24	27	5	2
51~60	0	0	1	3	6	10	11	48	10	11
61~70	0	0	1	0	2	6	4	30	24	33
70~100+	0	0	0	3	2	0	5	15	14	61



(a) 样本年龄:70岁 (b) 样本年龄:70岁 (c) 样本年龄:72岁

图 4 误判图像



(a) 样本年龄:73岁 (b) 样本年龄:74岁 (c) 样本年龄:75岁



(d) 样本年龄:80岁 (e) 样本年龄:83岁 (f) 样本年龄:90岁

图 6 60~100 年龄段部分训练样本图像



(a) 样本年龄:31岁 (b) 样本年龄:32岁 (c) 样本年龄:35岁



(d) 样本年龄:36岁 (e) 样本年龄:38岁 (f) 样本年龄:40岁

图 5 30~40 年龄段部分训练样本图像

进一步的,为了验证模型对于中青年年龄估计的效果,针对 21~50 岁样本按照 5 岁一个间隔划分,最终 UTKFace 数据集分为 0~5、6~10、11~15、16~20、21~25、26~30、31~35、36~40、41~45、46~50、50~60、60~70 及 70+,共 13 个年龄段,本文模型和文献[15-17]及文献[19]中的方法对 UTKFace 数据 13 个年龄段估计结果如表 4 所示。

表 3 UTKface 数据集两种方法性能比较
(粗体表示某年龄段最高指标值) %

年龄段	本文模型		AlexNet	
	Exact	1-off	Exact	1-off
0~5	94	98	80	94
6~10	90	96	45	78
11~15	82	94	29	75
16~20	88	96	28	77
21~30	69	90	53	83
31~40	69	93	36	73
41~50	65	92	24	79
51~60	87	96	48	69
61~70	79	96	24	87
71+	93	94	61	75
平均准确率	81.6	94.5	42.8	79

表 4 UTKFace 数据集上各方法的整体性能
(粗体表示最高准确率) %

方法	Exact	1-off
DCNN ^[15]	9.0	23.2
DEX ^[16]	29.9	62.8
Sighthoun ^[17]	33.4	73.1
文献[19](VGGFace)	27.1	61.9
本文模型	36.2	68.5

表 4 结果显示本文模型对年龄段的准确估计优于其它未加入对抗学习思想的模型;本文模型的 1-off 低于 Sighthound 是因为准确预测率最高,因此预测结果落在相邻年龄段的减少。进一步对比表 4 和 3 的结果,UTKFace 数据集年龄段细化分后的年龄估计准确率整体降低,这是因为人类成年后(21~50 岁)人脸年龄特征变化缓慢,该部分的样本原本各年龄段间人脸年龄特征变化就不明显,若将年龄段进一步拆分(如 31~40 岁拆分为 31~35 岁和 36~40 岁两个年龄段),成年年龄段间样本的人脸年龄特征变化非常小,从而影响了特征学习的效果和年龄估计准确。

2) MORPH 数据集年龄估计性能

表 5、6 分别为 MORPH 数据集上本文模型和 AlexNet 模型的年龄估计结果。通过计算得到 MORPH 数据集上年龄估计的准确率如表 7 所示,可见特别在年龄分类正确率上本文模型性能明显优于 AlexNet。MORPH 数据集“61~70”年龄段估计的准确率为 31%,是因为 MORPH 数据集中主要为黑人样本且图片多为高光拍摄,因此不同年龄段样本之间的年龄特征变化更加不明显,从而影响了 MORPH 数据集中该年龄段年龄估计的准确率。

对照表 3 和 7 可知对于 UTKface 数据集,年龄估计的正确率由 42.8% 提升至 81.6%,而“1-off”准确率也由

表 5 本文模型在 MORPH 数据集上年龄估计混淆矩阵
(粗体表示预测正确的数目)

年龄段	预测结果				
	21~30	31~40	41~50	51~60	61~70
21~30	77	20	3	0	0
31~40	8	83	7	2	0
41~50	2	19	72	7	0
51~60	1	6	7	86	0
61~70	0	9	22	38	31

表 6 AlexNet 在 MORPH 数据集上年龄估计混淆矩阵
(粗体表示预测正确的数目)

年龄段	预测结果				
	21~30	31~40	41~50	51~60	61~70
21~30	64	27	5	4	0
31~40	26	46	21	7	0
41~50	10	32	46	12	0
51~60	6	15	43	36	0
61~70	4	6	23	60	7

79.0% 提升至 94.5%; 同样,对于 MORPH 数据集,该两项准确率分别由 39.8% 和 84% 提升至 69.8% 和 91%,说明加入对抗训练思想的本文模型对各个年龄段的估计性能相对于未加入对抗训练思想的 AlexNet 有显著提升。又因为 UTKface 数据集的样本均拍摄于自然环境下、光照柔和且面部清晰,样本质量明显高于 MORPH 数据集,因此 UTKface 数据集年龄估计的准确率整体高于 MORPH 数据集。

表 7 MORPH 数据集两种方法性能比较
(粗体表示某年龄段最高指标值) %

年龄段	本文模型		AlexNet	
	Exact	1-off	Exact	1-off
21~30	77	97	64	91
31~40	83	98	46	93
41~50	72	98	46	90
51~60	86	93	36	79
61~70	31	69	7	67
平均准确率	69.8	91	39.8	84

同样为了验证模型对于中青年年龄估计的效果,将 MORPH 数据集细分为 20~25、26~30、31~35、36~40、41~45、46~50、51~60 和 61~70 共 8 个年龄段,本文模型和文献[15-17]及文献[19]中方法在 MORPH 数据集上对 8 个年龄段的预测结果如表 8 所示,可知在 MORPH 数据集上,引入对抗思想的本文模型对准确估计年龄性能也优于其他 4 种方法,进一步说明了在模型中加入对抗训练的有效性。

表 8 MORPH 数据集上各方法的整体性能
(粗体表示最高准确率) %

方法	Exact	1-off
DCNN ^[15]	12.4	36.5
DEX ^[16]	33.8	80.8
Sighthoun ^[17]	31.8	72.5
文献[19](VGGFace)	23.9	52.3
本文模型	34.6	70.9

表 9 本文模型在 Adience 数据集上年龄估计混淆矩阵(粗体表示预测正确的数目)

年龄段	预测结果							
	0~2	4~6	8~13	15~20	25~32	38~43	48~53	60+
0~2	85	8	1	2	1	2	1	0
4~6	16	67	10	0	4	0	3	0
8~13	6	6	73	4	5	4	1	1
15~20	1	1	9	58	19	7	3	2
25~32	0	3	3	7	66	11	8	2
38~43	0	2	5	4	20	54	11	0
48~53	2	2	11	4	13	16	49	3
60+	2	4	5	3	13	13	6	54

的“1-off”指标却又都在 90% 以上,这说明“15~20”、“38~43”这两个年龄段的测试样本大多数都误判到了相邻的年龄段中;并且“15~20”、“25~32”、“38~43”、“48~53”这 4 个年龄段预测的准确率呈现一好一差的特征,该结果说明 Adience 数据集上青中年相邻年龄段之间的年龄特征变化不明显,对其进行年龄估计时的确容易产生误判。本文模型在“15~20”、“25~32”、“38~43”、“48~53”这 4 个年龄段预测的准确率则较为均匀,尤其是“15~20”、“38~43”这两个年龄段的预测准确率相较于文献[19]方法有较大的提升,这表明本文模型在相邻年龄段年龄特征变化不明显的中青年年龄段(15~53 岁)有更好的年龄估计效果,再一次表明对抗训练思想可以提高模型对于边界点特征的学习能力,从而进一步提高年龄估计的准确率。

表 10 两种方法在 Adience 数据集上各年龄段性能
(粗体表示某年龄段预测最高准确率) %

年龄段	本文模型		文献[19]方法	
	Exact	1-off	Exact	1-off
0~2	85	93	93.1	99.6
4~6	67	93	62.1	96.3
8~13	73	83	42.0	61.2
15~20	58	86	24.2	93.4
25~32	66	84	86.1	92.7
38~43	54	85	8.8	91.5
48~53	49	68	38.1	80.5
60+	54	60	60.7	95.7

3) Adience 数据集年龄估计性能

将 Adience 数据集按照 0~2、4~6、8~13、15~20、25~32、38~43、48~53、60+ 分为 8 个年龄段,在 Adience 数据集上本文模型的混淆矩阵如表 9 所示。

参照文献[19]中结果,表 10 为本文模型和文献[19]方法在 Adience 上各年龄段性能表现。表 10 显示,在中青年年龄段(15~53 岁),文献[19]方法在“15~20”、“38~43”这两个年龄区间表现较差、准确率偏低,而这两个年龄段

3.3 模型整体性能对比分析

表 11 为本文模型与现有其他方法的年龄估计整体准确率。相对于文献[15]中使用的 DCNN(Inception 网络)、文献[16, 19]中使用的 VGG 网络、文献[17]中使用的 Sighthound 网络架构,本文模型使用浅层(5 层 Alexnet)网络结构就能达到 63.3% 的准确率,比其他 4 种方法准确率都高,在 1-off 指标上也能达到 DCNN 方法性能,这充分说明了本文模型的可行性和实用性。

表 11 Adience 数据集上各方法的整体性能
(粗体表示最高准确率) %

方法	Exact	1-off
DCNN ^[15]	50.7	84.7
DEX ^[16]	55.6	89.7
Sighthoun ^[17]	61.3	—
文献[19](VGGFace)	60.6	92.3
本文模型	63.3	81.5

总体来说,在训练中加入对抗丢弃正则化后,年龄估计的准确度显著提高,并且对于预测错误的年龄段,其误差也显著缩小。同时,由不同数据集之间的对比,以及相同数据集不同年龄段之间的对比还可以看出,估计的准确度也与数据集的选取和具体的年龄段相关,数据集的各年龄段的图片数据越均匀,以及该年龄段的人脸与相邻年龄段的人脸特征差距越大,越容易获得更好的效果。

4 结 论

本文在人脸年龄估计任务中引入对抗训练思想,提出基于对抗正则化的人脸年龄估计模型。在 UTKface、MORPH 和 Adience 3 个人脸数据集上的实验结果表明,该模型显著提高了年龄估计的准确率,能够提高年龄特征学习器性能,特别对相邻年龄段特征学习有效。另外年龄估计准确率与数据集的选取有关,通过对 UTKface 和 MORPH 两个数据集的性能对比分析可知,预测结果受样本性别、种族、光照、年龄段分布、妆容等因素影响,这将是我们的下一步的研究内容。

参考文献

- [1] CHEN S, ZHANG C, DONG M, et al. Using ranking-cnn for age estimation[C]. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2017: 5183-5192.
- [2] ANGULU R, TAPAMO J R, ADEWUMI A O. Age estimation via face images: A survey[J]. EURASIP Journal on Image and Video Processing, 2018(1): 1-35.
- [3] 潘思宇,赵雯婷,唐鲲,等.人脸年龄估计和年龄面貌合成技术研究进展[J]. 刑事技术,2017,42(4):270-276.
- [4] PORTO L F, LIMA L N C, FRANCO A, et al. Estimating sex and age from a face: A forensic approach using machine learning based on photoanthropometric indexes of the Brazilian population[J]. International Journal of Legal Medicine, 2020, 134(6): 2239-2259.
- [5] SAGAR S, STAMATIADIS N, WRIGHT S, et al. Use of codes data to improve estimates of at-fault risk for elderly drivers [J]. Accident Analysis & Prevention, 2020, 144: 105637.
- [6] ISLAM M, BAEK J H. Deep learning based real age and gender estimation from unconstrained face image towards smart store customer relationship management[J]. Applied Sciences, 2021, 11(10): 4549.
- [7] CHATTERJEE S, RANA N P, TAMILMANI K, et al. The effect of AI-based CRM on organization performance and competitive advantage: An empirical analysis in the B2B context[J]. Industrial Marketing Management, 2021, 97: 205-219.
- [8] KWON Y H, LOBO N D V. Age classification from facial images [J]. Computer Vision and Image Understanding, 1999, 74(1): 621-628.
- [9] COOTES T F, EDWARDS G J, TAYLOR C J. Active appearance models[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2001, 23(6): 681-685.
- [10] XIN G, ZHOU Z H, SMITH M K. Automatic age estimation based on facial aging patterns[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2007, 29(12): 2234-2240.
- [11] FU Y, HUANG T S. Human age estimation with regression on discriminative aging manifold[J]. IEEE Transactions on Multimedia, 2008, 10(4): 578-854.
- [12] 周嫣然,罗毛欣,赵博.基于多分辨率特征提取网络的图像相似度学习[J]. 国外电子测量技术,2021,40(5): 27-32.
- [13] 何其霖,穆平安. VGG 网络与多特征融合的遮挡人脸检测[J]. 电子测量技术,2021,44(18):150-154.
- [14] 王仕伟,闫海滨.利用判别采样的视频人脸亲属关系验证[J]. 电子测量与仪器学报,2021,35(8):12-19.
- [15] LEVI F, HASSNER T. Age and gender classification using convolutional neural networks[C]. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2015:34-42.
- [16] ROTHE R, TIMOFTE R, GOOL L V. Deep expectation of real and apparent age from a single image without facial landmarks [J]. International Journal of Computer Vision, 2016:1-14.
- [17] DEHGHAN A, ORTIZ E G, SHU G, et al. Dager: Deep age, gender and emotion recognition using convolutional neural network [J]. ArXiv Preprint, 2017, ArXiv:1702.04280.
- [18] HOU L, SAMARAS D, KURC T, et al. ConvNets with smooth adaptive activation functions for regression[C]. Proceedings of the 20th International Conference on Artificial Intelligence and Statistics, PMLR, 2017:430-439.
- [19] MALLOUH A A, QAWAQNEH Z, BARKANA B D. Utilizing CNNs and transfer learning of pre-trained models for age range classification from unconstrained face images[J]. Image and Vision Computing, 2019, 88: 41-51.
- [20] EIDINGER E, ENBAR R, HASSNER T. Age and gender estimation of unfiltered faces [J]. IEEE Transactions on Information Forensics & Security, 2014, 9(12):2170-2179.
- [21] 罗佳,黄晋英.生成式对抗网络研究综述[J]. 仪器仪表学报,2019,40(3):74-84.
- [22] SAITO K, USHIKU Y, HARADA T, et al. Adversarial dropout regularization[C]. ICIR, 2018.
- [23] NITISH S, GEOFFREY E H, KRIZHEVSKY A, et al. Dropout: A simple way to prevent neural networks from overfitting[J]. JMLR, 2014, 15(1):1929-1958.
- [24] KRIZHEVSKY A, SUTSKEVER I, HINTON G E. Image classification with deep convolutional neural

- networks[C]. Proceedings of the 25th International Conference on Neural Information Processing Systems, 2012:1097-1105.
- [25] ZHIFEI Z, YANG S, HAIRONG Q. Age progression/regression by conditional adversarial autoencoder [C]. IEEE Conference on Computer Vision and Pattern Recognition, 2017:4352-4360.
- [26] RICANEK K, TESAFAVE T. MORPH: A longitudinal image database of normal adult age-progression [C]. International Conference on Automatic Face and Gesture Recognition, IEEE, 2006:341-345.
- [27] EIDINGER E, ENBAR R, HASSNER T. Age and gender estimation of unfiltered faces [J]. IEEE Transactions on Information Forensics & Security, 2014, 9(12):2170-2179.

作者简介

朱昱, 本科生, 主要研究方向为机器视觉。

E-mail: 18851967721@163.com

樊航, 硕士研究生, 主要研究方向为大气科学、资料同化。

E-mail: fhang1999@163.com

周媛(通信作者), 讲师, 主要研究方向为机器视觉。

E-mail: zhouyuan@nuist.edu.cn