

控制与决策

Control and Decision

人脸性别约束下的深度随机森林表情识别

钟良骥, 廖海斌

引用本文:

钟良骥, 廖海斌. 人脸性别约束下的深度随机森林表情识别[J]. *控制与决策*, 2021, 36(7): 1693–1698.

在线阅读 View online: <https://doi.org/10.13195/j.kzyjc.2019.1703>

您可能感兴趣的其他文章

Articles you may be interested in

行人重识别中度量学习方法研究进展

A survey on metric learning in person re-identification

控制与决策. 2021, 36(7): 1547–1557 <https://doi.org/10.13195/j.kzyjc.2020.0801>

基于地标特征和元学习方法推荐最适用优化算法

Recommending best suitable metaheuristic based on landmarking feature and meta-learning approach

控制与决策. 2021, 36(5): 1223–1231 <https://doi.org/10.13195/j.kzyjc.2019.0993>

融合稀疏编码与深度学习的草图特征表示

A feature representation of sketch based on fusion of sparse coding and deep learning

控制与决策. 2021, 36(3): 699–704 <https://doi.org/10.13195/j.kzyjc.2019.0941>

基于卷积神经网络的云雾遮挡舰船目标识别

Obscured ship target recognition based on convolutional neural network

控制与决策. 2021, 36(3): 661–668 <https://doi.org/10.13195/j.kzyjc.2019.0781>

基于无标签、不均衡、初值不确定数据的设备健康评估方法

Equipment health risk assessment based on unlabeled, unbalanced data under uncertain initial condition

控制与决策. 2020, 35(11): 2687–2695 <https://doi.org/10.13195/j.kzyjc.2018.1493>

人脸性别约束下的深度随机森林表情识别

钟良骥¹, 廖海斌^{1,2†}

(1. 湖北科技学院 计算机科学与技术学院, 湖北 咸宁 437100; 2. 江西省智慧城市产业技术研究院, 南昌 330096)

摘要: 由于人脸表情类内变化和类间干扰因素的存在, 人脸表情识别仍面临着巨大挑战. 提出一种基于性别条件约束随机森林的深度人脸表情识别方法, 解决人脸表情识别中噪声、性别等变化和干扰问题. 首先, 采用深度多示例学习方法提取鲁棒性人脸特征, 解决人脸光照、遮挡和低分辨率等图像变化问题; 其次, 采用性别条件随机森林分类方法进行人脸表情分类器设计, 解决人脸性别因素干扰问题. 在公开的 CK+, BU-3DEF、LFW 人脸表情数据库上进行广泛实验结果表明: 所提出方法在 3 大人脸数据库上分别达到了 98.83%、90%、60.58% 的识别率, 与先进方法相比具有更好的性能和鲁棒性. 另外, 与其他先进的深度学习方法 (需要大量训练数据库) 相比, 所提出方法只需要少量训练样本就能达到较好效果.

关键词: 特征提取; 人脸表情识别; 性别估计; 深度学习; 随机森林

中图分类号: TP391

文献标志码: A

DOI: 10.13195/j.kzyjc.2019.1703

开放科学(资源服务)标识码(OSID):



引用格式: 钟良骥, 廖海斌. 人脸性别约束下的深度随机森林表情识别[J]. 控制与决策, 2021, 36(7): 1693-1698.

Facial expression recognition using deep random forest under gender constraints

ZHONG Liang-ji¹, LIAO Hai-bin^{1,2†}

(1. School of Computer Science and Technology, Hubei University of Science and Technology, Xianning 437100, China; 2. Jiangxi Smart City Industrial Technology Research Institute, Nanchang 330096, China)

Abstract: Facial expression recognition is still faced with great challenges due to the intra class variation and interclass interference. In this paper, a method of deep facial expression recognition based on random forest with gender constraints is proposed to solve the problems of noise, gender and other variation and interference. Firstly, robust facial features are extracted by deep multi-instance learning to solve the problems of illumination, occlusion and low resolution. Secondly, the face expression classifier is designed by using gender conditional random forest classification to solve the problem of gender interference. Extensive experiments on the public CK+, BU-3DEF and LFW databases show that the recognition rate of the proposed method is 98.83%, 90% and 60.58% respectively, which has better performance and robustness compared with the state-of-the-art methods. In addition, compared with other advanced deep learning methods (requiring a large number of training databases), the proposed method only needs a small number of training samples to achieve better results.

Keywords: feature extraction; facial expression recognition; gender estimation; deep learning; random forest

0 引言

人脸表情是人类一种最直接、自然和有效的情感表达方式, 它是人工智能和计算机领域的一个重要研究方向. 人脸表情识别可以广泛应用于安全驾驶、智能教室、视频监控、虚拟现实和人工机交互等场景^[1-3]. 由于人脸表情识别是一种特殊的模式识别问题, 一般的模式分类方法可以自然而然地应用

于人脸表情识别问题. 受到人脸识别和 ImageNet 图像识别巨大突破的启示, 深度学习技术被广泛应用于人脸表情识别^[4-7]. 例如, Kim 等^[7]通过多样化 CNNs 设计出一种层次化的网络结构应用于人脸表情识别, 在国际人脸表情竞赛 (EmotiW2015) 上获得第 1 名成绩; Bargal 等^[8]通过结合 VGG 和残差网络 (ResNet) 设计了一种混合神经网络进行人脸特征提取, 并使用

收稿日期: 2019-12-05; 修回日期: 2020-02-02.

基金项目: 国家自然科学基金项目 (61701174); 咸宁市自然科学基金项目 (2019kj130); 湖北科技学院培育基金项目 (2020-22GP03).

责任编辑: 李登峰.

†通讯作者. E-mail: liao_haibing@163.com.

SVM进行人脸表情分类,最终获得了较好的效果.但是,他们的方法严重依赖于大量的训练样本和高性能计算.最近,Bulo等^[9]提出深度学习和随机决策树相结合的图像语义标注方法,获得了很大成功,他们利用深度神经网络提取图像特征,利用决策树分类方法进行图像分类标注.与SVM相比,随机森林(RF)分类方法具有很好的大数据处理、容易实现、高泛化和快速性等特征,在计算机视觉领域被广泛应用^[10-12].因此,RF被成功应用于人脸表情识别、人脸姿态识别、人脸关键特征点定位和行为识别方面.Sun等^[11]提出基于深度信息的RF身体姿态识别方法;Fanelli等^[13]利用霍夫森林方法进行视频人脸表情识别,在MMI人脸表情库获得76%的识别率.由于RF优良的特性,多类别RF方法在人脸分类领域变得越来越重要.目前,在可控环境下,已经出现了大量的人脸表情识别方法并获得很大的成功.但是,只有少数工作去研究非可控环境下的人脸表情识别问题^[1,14-16].文献[15]为了解决人脸表情识别中人脸姿态变化问题,采用成对条件随机森林的方法进行人脸姿态变化处理,在多姿态BU3D-EF人脸表情库获得了76.1%的识别率.另外,本文发现在实际的应用场景中,男性和女性有着不同的表情变化模式,即人脸性别对人脸表情识别存在较大影响.因此,非可控环境和小样本下的人脸表情识别问题仍然是一个具有挑战的问题.本文提出基于深度学习和条件随机森林的鲁棒性人脸表情识别方法(G_DRF),力求缓解深度学习需要大量训练样本支持的问题,追求非可控环境下的人脸表情识别准确性和有效性.

1 基于G_DRF的人脸表情识别模型

G_DRF人脸表情识别技术流程如图1所示.

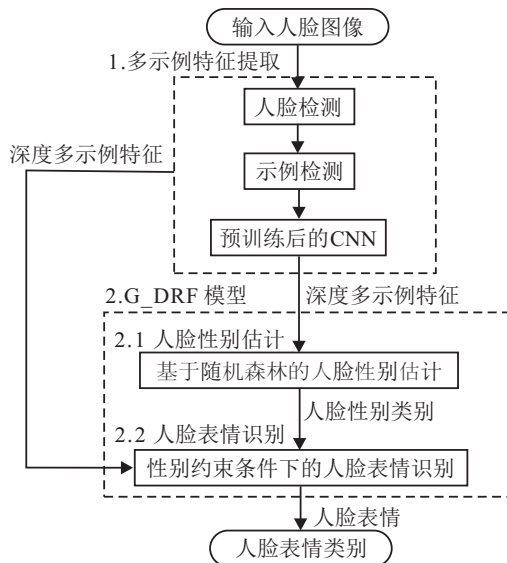


图1 G_DRF人脸表情识别技术流程

首先,采用深度多示例学习提取鲁棒性人脸特征以降低人脸图像遮挡、光照和低分辨率等变化问题;然后,基于提取的人脸特征,利用随机森林分类方法进行人脸性别估计;最后,基于提取的人脸特征及其性别属性,利用条件随机森林进行人脸表情识别.

1.1 基于深度多示例学习的鲁棒性特征提取

人脸特征提取是人脸表情识别的关键,本文提出深度多示例学习的人脸特征提取方法:首先,根据人脸结构特性选取人脸表情示例;然后,将选取的人脸表情示例送入经过预训练的GoogLeNet网络模型进行示例表示和多示例融合,提取人脸表情特征.

1) 人脸表情示例选择.

参考图像块随机/密集采样^[13]和显著检测图像块选取方法^[16],同时根据人脸的“三眼五庭”结构特性,提取7个人脸图像块作为表情示例,如图2所示.其中包括整个人脸图像、左眼区域块、右眼区域块、眼部区域块、嘴部区域块1、嘴部区域块2和嘴部区域块3等7个示例.之所以选取以上7个人脸图像块作为表情示例的原因是由于人脸表情变化主要集中在眉毛、眼睛、嘴巴等关键区域.

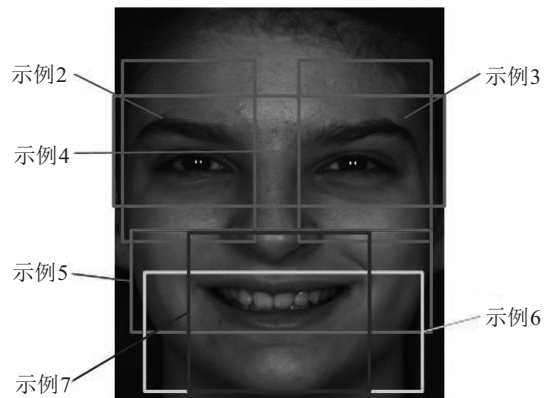


图2 人脸表情示例选取

2) 多示例特征提取与融合.

将上步得到的人脸表情示例送入GoogleNet网络模型进行表示学习.GoogleNet通过LFW和YTF人脸库进行预训练,使其具有高层语义信息表示能力.基于GoogleNet的人脸表情示例特征提取可以表示为

$$y^i = x_c^i w + b. \tag{1}$$

其中: y^i 为最后的高层特征输出, x_c^i 为第*i*个示例在“Depthconcat”层中的特征图, w 为权值项, b 为偏置项.因此,多示例融合特征可以通过将各示例特征串连而得到,即

$$y = [y^1 \ y^2 \ y^3 \ y^4 \ y^5 \ y^6 \ y^7]. \tag{2}$$

最后,在深度学习网络模型中增加一个全连接层对式

(2)得到的多示例特征 y 进行降维处理.

1.2 G_DRF人脸表情识别

由于人脸性别等因素的存在,导致了人脸表情特征空间的多样性,使得很难找到一个合适的分类器进行高精度的人脸表情分类识别.因此,本文使用人脸性别作为隐含条件进行人脸表情空间划分,提出一种基于性别约束的条件随机森林人脸表情识别方法.

模型训练 首先,使用整个训练集 S 去训练生成一个基于人脸性别分类的随机森林 T^G ;然后,训练集 S 根据人脸性别类别被分成 $N = 2$ 个子集,每一个子集被用于训练生成一组条件随机森林 $\{T^S(\Omega)\}_{n=1}^N$,用于人脸表情分类.本文将人脸性别划分成两类 $\{\Omega_1 : \text{female}, \Omega_2 : \text{male}\}$.

表情识别 首先,利用人脸性别随机森林 T^G 对测试人脸图像进行性别估计;然后,根据性别估计结果,随机地从条件随机森林 $\{T^S(\Omega)\}_{n=1}^2$ 中选择相应数量的决策树动态构建人脸表情识别随机森林 T^E ;最后,根据构建的人脸表情识别随机森林进行人脸表情分类识别.

1) 人脸性别估计.

采用深度多示例学习提取人脸特征 y ,采用随机森林 T^G 进行人脸性别分类.在训练生成性别随机森林 T^G 模型时,采用如下不确定性测度:

$$H(y) = -\sum_g P(g|y) \log_2 P(g|y), \quad (3)$$

其中 g 表示人脸性别类别.不确定性测度引导各节点从二值测试候选库中选择最优的二值测试,以保证能不断地将当前节点分裂为不确定性降低的两个子节点.人脸性别以高斯模型的方式储于随机森林 T^G 的每个叶子节点 l 上,即

$$P(g|l(y)) = N(g; \bar{g}_l; \sigma_l), \quad (4)$$

其中 \bar{g}_l 和 σ_l 表示叶子节点 l 上性别均值和方差.生成叶子节点后,人脸特征 y 在随机森林 T^G 上的性别类别概率可以通过投票方式获得,即

$$P(g|y) = \frac{1}{M} \sum_m P(g|l_m(y)). \quad (5)$$

其中: l_m 为决策树对应的叶子节点, M 为决策树的数量.

2) 条件随机森林训练.

条件随机森林 $\{T^S(\Omega)\}_{n=1}^2$ 中的每棵决策树 $T^S(\Omega_n)$ 采用相同的方法独立训练.为构建每棵决策树,从相应的数据子集 S_{Ω_n} 中随机选取图像构成训练数据集;然后,从每张训练样本特征 y 中随机提取一组子特征 $\{y_i = (e_i, I_i)\}$,其中 e_i 为表情类别, $I_i = \{I_i^1, I_i^2, \dots, I_i^F\}$ 为提取的子特征集合;最后,提取的子特征集合被用于决策树节点的分裂,以便生

成最终的决策树.

为了更好地学习和生长分裂节点,本文采用 NCSF 分裂模型^[16],结合决策树的度量函数信息增益 (IG) 和深度学习模型中的损失函数进行深度网络增强树的节点生长.基于 CNN 隐藏层的联接函数 f_n 强化人脸样本的条件特征表示 y ,用增强后特征表示作为网络增强森林的节点特征选择,即

$$d_n(y, K|\Omega_g) = \sigma(f_n(y, K|\Omega_g)). \quad (6)$$

其中: $\sigma(x) = (1 + e^{-x})^{-1}$ 为 sigmoid 函数; Ω_g 为不同性别下的表情子森林,为增强森林中的一个分裂节点; K 为 CNN 网络模型参数,可通过 Adam 算法更新.当信息增益 IG 最大时,分裂生成树的左、右子节点

$$\hat{\psi} = \operatorname{argmax}_{\psi} \left(H(d_n) - \sum_{S \in \{N_r, N_l\}} \frac{|d_n^S|}{|d_n|} H(d_n) \right). \quad (7)$$

其中: $d_n^S/d_n (S \in \{R, L\})$ 为左、右子节点特征样本数量概率, $H(d_n)$ 为节点 d_n 的熵.当树的深度达到最大或者损失函数迭代收敛后,生成叶子节点;否则,继续迭代节点学习.

3) 基于性别条件随机森林的人脸表情识别.

在人脸性别 $g \in \Omega_n$ 已知的条件下,人脸样本 y 属于表情 e 的概率 $P(e|\Omega_n, y)$ 由随机森林 T^E 中所有树投票得到,即

$$P(e|\Omega_n, y) = \frac{1}{M} \sum_m P(e|\Omega_n, l_m(y)). \quad (8)$$

其中: M 为树的数量, l_m 为第 m 棵树上样本特征 y 达到的叶子节点.

由式(8)可以看出,在进行人脸表情识别时:首先,根据性别估计的结果从条件随机森林 $T^S(\Omega_n)$ 中随机选取 k_n 棵决策树动态构建随机森林 T^E ;然后,由 T^E 中决策树投票得出测试样本 y 的表情类别概率 $P(e|y)$.

2 实验分析

2.1 实验数据和参数设置

为了验证所提出方法的有效性,在公开的 Cohn-Kanade (CK+)^[17]、BU-3DFE^[18] 和 LFW^[19] 数据集上进行训练和测试,部分识别效果如图3所示. CK+ 是目前使用最多的公开人脸表情识别测试库,它采集了123个人,每个人包括593幅不同的表情图片;本文选取其中最常用的6种表情共1500幅图像作为训练和测试之用. BU-3DFE 是一个常用的多姿态表情人脸库,它采集了100个人(男性占56%),每个人包括了7种不同表情的2D和3D人脸样本. LFW 收集的是自然场景下的人脸数据库,包含了人脸姿态、表情、光照和遮挡等各种变化,本文对该数据集进行常见

6大类表情标注,收集5000幅复杂场景的不同表情人脸图像作为训练和测试之用.在模型训练阶段:从CK+、BU-3DFE和LFW中分别随机挑选1100、1800和4500张,共7400张不同性别、表情人脸图像作为G_DRF模型的训练样本集.在测试阶段:从CK+、BU-3DFE和LFW中分别挑选400、600和500张,共1500张不同于训练样本的人脸图像作为测试样本集.模型训练关键参数设置为:学习率 $\lambda = 0.001$, epochs为4500,分裂迭代次数为1000,树的深度为16.

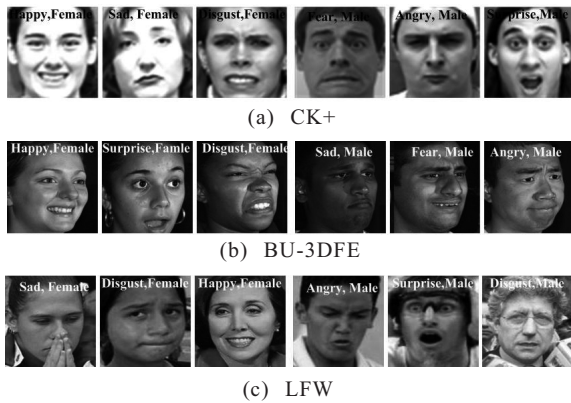


图3 CK+、BU-3DFE和LFW数据库识别结果示例

2.2 人脸性别估计实验

本节收集3万张人脸性别图像样本对性别随机森林进行训练,在1500张测试样本集上进行性别估计测试,人脸性别估计结果如图3所示.表1给出了本文方法与CNN^[20]、ROR^[21]的比较结果.CNN方法采用AlexNet网络结构进行人脸特征提取和性别分类,取得了95.17%的识别率.ROR采用基础块和瓶颈块构建ROR网络结构,获得了97.26%的识别率.本文方法得到了最好的识别率98.03%.由表1可以看出本文方法的优越性,另外0.5的方差表明了其具有很好的鲁棒性.

表1 不同方法的性别估计结果

methods	features	accuracy	STD
CNN	image	95.17	0.7
ROR	image	97.26	0.65
本文	deep transfer feature	98.03	0.5

2.3 人脸特征提取方法分析实验

为了验证所提出的特征提取方法的有效性,采用人脸表情识别领域常用的特征提取方法进行比较分析,其中包括深度学习特征、SIFT、HOG和几何结构特征.表2给出了不同特征提取方法在SVM分类器下的平均表情识别率.实验结果表明,本文提出的深度多示例特征提取方法优于其他方法,特别是在较难的LFW数据集上优势更加明显.

表3给出了不同特征提取方法在本文提出的性别条件随机森林分类器下的表情识别比较结果.可

以清楚地看到本文方法同样取得了最好效果.在较难的LFW数据集上提出的深度多示例特征提取方法取得了55.72%的识别率,高于第2名5.4个百分点.另外,由表2、表3对比可以看出,提出的性别条件森林分类方法明显优于SVM分类器,在BU-3DFE和LFW数据集上分别提高了6.64和8.48个百分点.

表2 基于SVM的不同特征提取方法比较结果

features	CK+	BU-3DFE	LFW
deep multi-instance feature	97.12	82.78	47.24
CONV.13 from VGG	93.28	74.85	37.65
SIFT	78.66	56.02	29.83
HOG	75.83	67.53	30.57
SIFT+HOG	71.35	53.91	32.33
distances between 21points	86.25	65.37	44.84

表3 基于G_DRF的不同特征提取方法比较结果

features	CK+	BU-3DFE	LFW
deep multi-instance feature	98.83	89.42	55.72
CONV.13 from VGG	94.67	74.24	43.20
SIFT	80.23	72.57	31.62
HOG	80.73	67.58	32.67
SIFT+HOG	83.58	73.93	36.61
distances between 21points	90.87	74.68	50.32

2.4 人脸表情识别实验

为了验证所提出方法的有效性,本节将分别在CK+、UB-3DFE和LFW表情库进行人脸表情识别实验,实验结果如表4~表6所示.在经典的CK+数据集上平均识别率达到98.83%,最低识别率也达到了95.5%;在具有姿态变化的UB-3DFE数据集上平均识别率达到89.42%;在最难的LFW数据集上平均识别率达到55.72%,高兴表情达到了最高识别率74.6%,害怕和悲伤表情都高于53%,效果最差的厌恶表情也达到了44.4%.

为了进一步验证本文方法的先进性,将其与目前先进的方法进行对比实验,实验结果如表7所示.其

表4 基于CK+的表情识别混淆矩阵

	anger	disgust	fear	happy	sadness	surprise
anger	1	0	0	0	0	0
disgust	0.030	0.955	0	0	0	0.015
fear	0.0015	0	0.9985	0	0	0
happy	0	0	0	1	0	0
sadness	0	0.01	0.004	0	0.986	0
surprise	0	0	0	0.010	0	0.990

表5 基于BU-3DFE的表情识别混淆矩阵

	anger	disgust	fear	happy	sadness	surprise
anger	0.915	0	0.04	0	0.045	0
disgust	0.08	0.863	0	0.017	0.04	0
fear	0	0.08	0.861	0.059	0	0
happy	0	0.022	0.021	0.916	0	0.041
sadness	0.055	0.045	0	0	0.90	0
surprise	0	0	0.04	0.05	0	0.91

表6 基于LFW的表情识别混淆矩阵

	anger	disgust	fear	happy	sadness	surprise
anger	0.487	0.265	0.176	0	0.059	0.013
disgust	0.157	0.444	0.098	0.02	0.142	0.139
fear	0.054	0.136	0.545	0.016	0.195	0.014
happy	0.021	0.17	0.04	0.746	0.032	0.021
sadness	0.155	0.163	0.126	0	0.531	0.025
surprise	0.021	0.151	0.166	0.033	0.036	0.593

表7 基于CK+、BU-3DFE和LFW数据集的不同方法比较结果

methods	accuracy		
	CK+	BU-3DFE	LFW
T-DCN ^[4]	80.49	—	—
HF ^[13]	87.1	—	—
PCRF ^[23]	96.4	76.1	51.42
AU-DNN ^[6]	92.22	—	—
M-SVM ^[24]	93.6	71.1	—
JFDNN ^[25]	95.21	72.5	—
CNN ^[5]	93.2	72.56	—
C-CNN ^[22]	96.76	86.24	—
SIFT-CNN ^[11]	—	80.5	—
GSRRR ^[14]	—	79.12	—
SVM based on Gabor feature	—	—	43.67
RF based on Gabor feature	—	—	46.56
G_DRF	98.83	89.42	55.72

中,有常规卷积神经网络的人脸表情识别方法,如CNN^[5],在CK+和BU-3DFE上获得了93.2%和72.56%的识别率;也有综合不同CNN模型加上特殊图像预处理的C-CNN^[22]方法,在CK+和BU-3DFE数据集上分别获得了96.76%和86.24%的识别率;另外,还有采用传统LBP+LGBP特征和多类别SVM的人脸表情识别方法在CK+和BU-3DFE数据集上分别获得了93.6%和71.1%的识别率.可以看出,本文方法具有最好的效果,究其原因是因为其综合了两种优秀的特征提取和分类识别方法,在特征提取端采用了鲁棒的深度多示例学习方法,在分类器设计端采用了性能较好的条件随机森林分类方法.

2.5 小样本实验

为了验证本文方法对小量级训练样本的鲁棒性,利用5种不同的深度学习方法在同样的小样本集进行比较实验.5种不同的深度学习方法都采用预训练迁移学习的政策进行网络初始化.小样本集从CK+抽样组建,包括480张训练集,60张测试集.实验采用

表8 基于小样本集的不同方法比较结果

convolutional models	number of images		accuracy / %	STD
	#training	# testing		
ResNet-50	480	60	80.6	0.8
AlexNet	480	60	76.96	0.9
T-DCN ^[4]	480	60	76.56	0.7
JFDNN ^[25]	480	60	93.3	1.0
G_DRF	480	60	95.65	0.5

交叉验证方法得出最终识别结果,表8给出了不同方法在小样本集的比较结果.可以看出,本文方法获得了95.65%的最好结果,充分展示了其对小样本的鲁棒性.

2.6 性别条件约束实验

为了验证所提出的性别条件随机森林人脸表情识别方法的有效性,本节特意比较了没有性别属性对齐和有性别属性对齐的人脸表情识别结果.实验结果如图4所示,在LFW上采用性别属性对齐的条件随机森林方法比没有使用性别对齐的方法高出了5%左右,在BU-3DFE上高出了2%左右.实验显示了性别属性对人脸表情识别存在干扰影响.

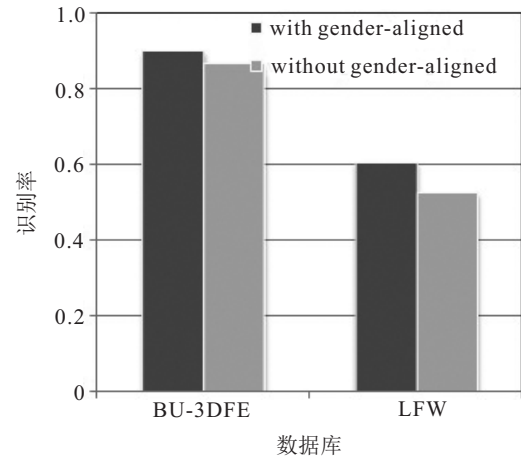


图4 带性别约束和不带性别约束对比识别结果

3 结论

本文提出了一种基于性别对齐的条件随机森林方法对人脸表情进行鲁棒性识别,并提出了一种深度多示例特征提取方法,在利用迁移学习策略的基础上充分利用多示例学习方法提出鲁棒性人脸表情特征.在分类器设计方面,为了充分利用人脸性别属性特征,提出了一种性别属性对齐的条件随机森林分类方法进行人脸鲁棒性人脸表情识别.实验结果表明,本文方法先进,对人脸遮挡、噪声和分辨率变化具有较好的鲁棒性.下一步工作,考虑将人脸年龄属性加入分析.因为,在现实中,人脸表情不仅受性别因素影响,还受年龄因素影响.即不同性别和不同年龄段的人表现出不同的情感表现方式.如果能同时将人脸性别和年龄因素加入到人脸表情识别中,一定会对识别率的提高有所帮助.

参考文献(References)

[1] Zhang T, Zheng W M, Cui Z, et al. A deep neural network-driven feature learning method for multi-view facial expression recognition[J]. IEEE Transactions on Multimedia, 2016, 18 (12): 2528-2536.
 [2] Yuan X H, Xie L J, Abouelenien M. A regularized

- ensemble framework of deep learning for cancer detection from multi-class, imbalanced training data[J]. *Pattern Recognition*, 2018, 77: 160-172.
- [3] 刘汝涵, 徐丹. 视频放大和深度学习在微表情识别任务上的应用[J]. *计算机辅助设计与图形学学报*, 2019, 31(9): 1535-1541.
(Liu R H, Xu D. Video amplification and deep learning in micro-expression recognition[J]. *Journal of Computer-Aided Design & Computer Graphics*, 2019, 31(9): 1535-1541.)
- [4] Xu M, Cheng W, Zhao Q, et al. Facial expression recognition based on transfer learning from deep convolutional networks[C]. *Proceedings of the International Conference on Natural Computation. Zhangjiajie: IEEE*, 2015: 702-708.
- [5] Mollahosseini A, Chan D, Mahoor M H. Going deeper in facial expression recognition using deep neural networks[C]. *Proceedings of the IEEE Winter Conference on Applications of Computer Vision. New York: IEEE*, 2016: 1-10.
- [6] Liu M, Li S, Shan S, et al. Au-aware deep networks for facial expression recognition[C]. *Proceedings of the IEEE International Conference and Workshops on Automatic Face and Gesture Recognition. Shanghai: IEEE*, 2013: 1-6.
- [7] Kim B K, Lee H, Roh J, et al. Hierarchical committee of deep cnns with exponentially-weighted decision fusion for static facial expression recognition[C]. *Proceedings of the ACM International Conference on Multimodal Interaction. Seattle: ACM*, 2015: 427-434.
- [8] Bargal S A, Barsoum E, Ferrer C C, et al. Emotion recognition in the wild from videos using images[C]. *Proceedings of the ACM International Conference on Multimodal Interaction. Tokyo: ACM*, 2016: 433-436.
- [9] Buló S R, Kotschieder P. Neural decision forests for semantic image labeling[C]. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Columbus: IEEE*, 2014: 81-88.
- [10] Zhou Z H, Feng J. Deep forest: Towards an alternative to deep neural networks[C]. *Proceedings of the 26th International Joint Conference on Artificial Intelligence. Melbourne: IJCAI*, 2017: 3553-3559.
- [11] Sun M, Kohli P, Shotton J. Conditional regression forests for human pose estimation[C]. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Rhode Island: IEEE*, 2012: 3394-3401.
- [12] 邵琦, 陈云浩, 杨淑婷, 等. 基于随机森林算法的玉米品种高光谱图像鉴别[J]. *地理与地理信息科学*, 2019, 35(5): 34-39.
(Shao Q, Chen Y H, Yang S T, et al. Identification of maize seed varieties based on random forest and hyperspectral technique[J]. *Geography and Geo-Information Science*, 2019, 35(5): 34-39.)
- [13] Fanelli G, Yao A, Noel P L, et al. Hough forest-based facial expression recognition from video sequences[C]. *Proceedings of the European Conference on Computer Vision. Crete: IEEE*, 2010: 195-206.
- [14] Zheng W. Multi-view facial expression recognition based on group sparse reduced-rank regression[J]. *IEEE Transactions on Affective Computing*, 2014, 5(1): 71-85.
- [15] Dapogny A, Bailly K, Dubuisson S. Dynamic pose-robust facial expression recognition by multi-view pairwise conditional random forests[J]. *IEEE Transactions on Affective Computing*, 2019, 10(2): 167-181.
- [16] Liu Y Y, Yuan X H, Gong X, et al. Conditional convolution neural network enhanced random forest for facial expression recognition[J]. *Pattern Recognition*, 2018, 84: 251-261.
- [17] Lucey P, Cohn J F, Kanade T, et al. The extended cohn-kanade dataset (ck+): A complete dataset for action unit and emotion-specified expression[C]. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. San Francisco: IEEE*, 2010: 94-101.
- [18] Yin L, Wei X, Sun Y, et al. A 3d facial expression database for facial behavior research[C]. *Proceedings of the International Conference on Automatic Face and Gesture Recognition. Southampton: IEEE*, 2006: 211-216.
- [19] Huang G B, Ramesh M, Berg T, et al. Labeled faces in the wild: A database for studying face recognition in unconstrained environments[R]. Amherst: University of Massachusetts, 2007: 7-49.
- [20] Levi G, Hassner T. Age and gender classification using convolutional neural networks[C]. *Proc of the IEEE Conference on Computer Vision and Pattern Recognition. Boston: IEEE*, 2015: 34-42.
- [21] Zhang K, Gao C, Guo L, et al. Age group and gender estimation in the wild with deep RoR architecture[J]. *IEEE Access*, 2017, 5: 22492-22503.
- [22] Lopes A T, Aguiar E, De Souza A F, et al. Facial expression recognition with convolutional neural networks: Coping with few data and the training sample order[J]. *Pattern Recognition*, 2017, 61: 610-628.
- [23] Dapogny A, Bailly K, Dubuisson S. Pairwise conditional random forests for facial expression recognition[C]. *Proceedings of the IEEE International Conference on Computer Vision. Santiago: IEEE*, 2015: 3783-3791.
- [24] Zhang X, Mahoor M H, Mavadati S M. Facial expression recognition using lp-norm MKL multiclass-SVM[J]. *Machine Vision and Applications*, 2015, 26(4): 467-483.
- [25] Jung H, Lee S, Yim J, et al. Joint fine-tuning in deep neural networks for facial expression recognition[C]. *Proceedings of the IEEE International Conference on Computer Vision. Santiago: IEEE*, 2015: 2983-2991.

作者简介

钟良骥(1980—), 男, 副教授, 从事大数据、图像处理等研究, E-mail: 41676637@qq.com;

廖海斌(1982—), 男, 副教授, 博士, 从事图像处理、模式识别等研究, E-mail: liao_haibing@163.com.

(责任编辑: 孙艺红)