

自然场景下人脸表情数据集的构建

叶继华 刘 燕 李汉曦 甘 莹

(江西师范大学计算机信息工程学院, 南昌, 330022)

摘 要: 目前人脸识别研究中表情数据集图像数量较少、表情信息单一, 不利于人脸表情识别的研究。本文创建了自然场景下带标签的人脸表情数据集(Facial expression dataset in the wild, FELW), 并对其进行测试。FELW表情数据集包含了多张从互联网上收集的不同的年龄、种族、性别的人脸表情图像, 采用适合的方法标注每张图像带有人脸部件的状态标签和表情标签, 并引入Kappa一致性检验, 提高人脸表情识别率。使用传统方法和深度学习的表情识别方法对数据集进行实验分析, 与其他公开的人脸表情数据集相比, FELW数据集具有更多图像和更丰富的表情类别, 并包含了两种图像标签有利于表情识别的研究。

关键词: 表情数据集; Kappa一致性检验; 表情识别; 自然场景

中图分类号: TP391 **文献标志码:** A

Construction of Facial Expression Dataset in Natural Scene

Ye Jihua, Liu Yan, Li Hanxi, Gan Ying

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

Abstract: Nowadays, there are many facial expression datasets for expression research. But, these datasets are small amounts of images and little expression information, which limits the expression research. This paper introduces the construction of the facial expression dataset in the wild (FELW) and the test case. The FELW dataset includes many facial expression images, and the images of different age, race and gender are collected from the Internet. Each image includes two labels—the state of facial part label and the expression label, that labeled by the appropriate method. And the author imported Kappa consistency check to improve the recognition rate of the FELW dataset. The author used traditional method and deep learning method to experiment and analysis on this dataset. Compared with other public facial expression datasets, the FELW dataset has much more amounts of images and more varieties of expression, and contains two labels to help to the expression research.

Key words: expression dataset; Kappa consistency check; expression recognition; deep learning; in the wild

引 言

人脸表情识别技术是涉及生物特征识别、图像处理、运动跟踪、机器视觉、模式识别、生理学、心理学等研究领域的一个富有挑战性的交叉课题, 是多年以来模式识别与人工智能领域研究的一个热点问题^[1]。

基金项目: 国家自然科学基金(61462042, 61650105)资助项目。

收稿日期: 2018-10-30; **修订日期:** 2018-11-29

人脸表情识别发展较晚,目前还不够成熟,运用较成熟的数据集有日本 Advanced Telecommunication Research Institute International (ATR)的专门用于表情识别研究的基本表情数据库 JAFFE 和 Cohn-kanade 表情数据库(CK),以及 Lucey 等人在 2010 改进的数据集 CK+,但是这几个数据集信息比较单一并且图像数量比较少,如果用于深度学习的表情识别中远远不够,而且数据集都是在实验室条件下采集的表情图像。为配合表情研究,实验室自建了一个图像数量更多且表情类别更丰富(10 种表情)的自然场景下人脸表情数据集(Facial expression datasets in the wild, FELW)。

1 人脸表情数据集的构建

1.1 自然场景下带标签的人脸表情数据集(FELW)

FELW 数据集包含 26 848 张人脸图像,每张图像带有人脸部件的状态标签和表情标签两种标签。下面简单介绍 FELW 数据集的构建过程,数据集构建分为收集人脸表情图像和标注两个阶段。

1.1.1 收集人脸表情图像

在互联网上批量下载 30 万张自然场景下的人脸图像;然后采用自动人脸检测程序,获得了将 26 848 张人脸图像;最后裁剪其灰度化成 120 像素 \times 120 像素的灰度图,得到未带标签的 FELW 的原始图像。图 1 是数据集部分原始表情图像。

1.1.2 标注

这个阶段是对灰度图标注,标注内容包括了人脸部件的状态标签(Part_Label)和表情标签(Emotion_Label)两种标签。

Part_Label,包含了眉、眼、嘴、鼻、肌肉五大类。Emotion_Label,包含 10 类表情,分别为平(中性)、喜(高兴)、怒(生气)、哀(悲伤)、惊(惊讶)、恐(恐惧)、恶(厌恶)、羞(害羞)、傲(傲慢)、鄙视。FELW 标注是用 MATLAB 中的 GUI 设计界面的,将原始表情图像分成多组,每组由一个人标注,FELW 的标签标注界面实例如图 2 所示。

因为人在各种表情下面部各个部件的状态变化大多数比较细微,有时候会达到很难辨别的程度。因此需要对所选定的 5 个面部部件的状态进行标准的设定。对眉毛来讲,设定 5 个状态,分别为正常、弯眉、皱眉、眉毛上竖和平直。对眼睛同样设置 5 个状态,分别为正常、眼眯、眼向下(可通过观察上眼皮是否下垂来辅助判断)、瞪眼和眼向上(可通过观察眼球是否朝上辅助判断);嘴巴则设置了 6 个状态,分别为正常、上翘(通过观察嘴巴是否上扬辅助判断)、嘴向下(通过观察嘴部形成的弧线是否朝下进行辅助判断)、张嘴(通过观察嘴巴是否张开或形成一个 O 形进行辅助判断)、咬牙(在能看到牙齿的情况下,观察待标注图像中的对象是否咬牙切齿)和撅嘴(嘟嘴或撅嘴);鼻子只设置了 4 个状态,分别为正常、鼻孔收缩(通过观察鼻头部分的肌肉是否紧绷进行辅助判断)、鼻孔张大(通过观察鼻头部分的肌肉是否舒张进行辅助判断)和鼻梁(通过观察鼻根部分是否紧缩进行判断);肌肉则只设定 3 个状态,分别为正



图 1 FELW 数据集部分表情图像

Fig.1 FELW partial expression image of data set



图 2 GUI 界面标注实例

Fig.2 GUI interface annotation instance

常、紧绷(通过结合观察待标注对象是否痛苦和直接观察其肌肉是否紧绷进行辅助判断)和放松(通过观察待标注对象肌肉是否舒张或传达的表情是否欢喜而进行辅助判断)^[2]。图2的左边是Part_Label部分,Part_Label是一组 1×5 的向量;右边是Emotion_Label部分,使用的是软标签形式,有10组表情,每组表情都可以进行打分,打分区间在1~5,分数越高表示该表情的概率越大。当判断不出任何表情或者脸上被遮挡时,直接对表情打6分,表示该图像不属于任何表情。最终得到的EmotionLabel是一组 1×10 的向量。FELW是一个包含了26 848张 120×120 的灰度人脸图像和其对应的Part_Label和Emotion_Label信息的数据集,这就是自然场景下带标签的人脸表情数据集。

1.2 FELW数据集的不足

建立FELW数据集后,采用传统方法和深度学习方法进行实验,FELW数据集使用传统方法的最高识别率是65.01%,使用深度学习的人脸表情识别方法最高识别率是56.53%,最好的实验结果相对于先研究的数据集实验结果来说识别率仍然不高。通过对实验结果的分析,发现原数据集存在以下不足:

(1) 图像并不全是正脸,直接用这样的图像进行表情识别的识别率不会高。

(2) Part_Label部分不合理。FELW的Part_Label部分有一些部件有主很难判断的标签,对于标注者来说,“鼻”和“肌肉”部分的细节很难在一张灰度图中区分。很难判断“鼻”的正常表情,收缩、张大,耸;同样也难以判断“肌肉”的松弛、紧绷。同时可以去掉一些重复的状态,“眉”部分的正常表情和弯(眉弯)可以合并成一个状态,因为大部分人的眉毛在正常情况下就是弯的。“眼”部分的特殊形状,向上(眼向上),向下(眼向下)部分也难以判断。“嘴”部分的“正常”可以去掉,因为大部分人的嘴巴在正常情况下就是上翘或者下垂的,与第二第三状态重合。

(3) 每个Emotion_Label都需要打分太复杂。

(4) FELW的每张图像的标签只由一个标注者进行标注,太过主观。

1.3 FELW数据集的改进

1.3.1 数据集的改进

根针对上述不足,图像并不全是正脸,可以添加各个角度方位的状态标签;将Part_Label部分冗余的状态作调整;修改Emotion_Label部分的打分情况;扩大标注规模。对数据集进行如下改进:

(1) 在人脸部件状态标签中对添加“人脸的角度”,“人脸的角度”的状态有“正脸”“左侧脸”“右侧脸”“仰视”“俯视”。

(2) 对Part_Label进行改进,去掉“鼻”和“肌肉”部分;调整“眉”“眼”“嘴”部分。

(3) Emotion_Label进行改进,直接勾选表情,去掉对各个表情的打分,最后取置信度。

(4) 所有图像的标注都由3个不同的人进行标注,提高客观性。

对改进的FELW数据集的Part_Label进行标注时需要观察图像的眉毛、眼睛、嘴巴、人脸角度方向的状态,眉毛有4种状态,眼睛有3种状态,嘴巴有4种状态,人脸角度方向有5种状态。图3是对改进后的FELW人脸的4种部件的每个状态的展示。

图4是改进后的FELW数据集的GUI标注界面,图左边是Part_Label,包含了眉、眼、嘴、人脸角度方向4大类;右边是Emotion_Label,包含10类表情,是多选的;在界面最下面是“置信度”,置信度是标注者自己对整张图像的判断打分,置信度区间为0~5,置信度越高表示判断越高,置信度为0表示看不出任何表情,用Coefficient表示置信度。

标注完后,Part_Label和Emotion_Label信息存入.mat文件,.mat文件包括EmotionLabel、Name、value、Coefficient。EmotionLabel表示Emotion_Label标签,是一个 1×10 向量,10个向量元素是10类表情,分别为平(中性)、喜(高兴)、怒(生气)、哀(悲伤)、惊(惊讶)、恐(恐惧)、恶(厌恶)、羞(害羞)、傲(傲慢)、鄙视,

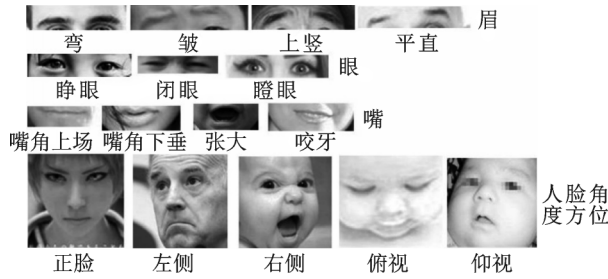


图3 改进后的FELW标签的人脸四种部分状态展示

Fig. 3 Four partial states of the face labelled by the improved FELW

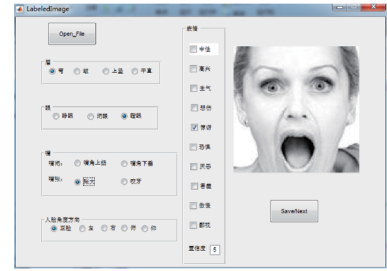


图4 改进后的FELW标注界面实例

Fig. 4 Improved FELW annotation interface instance

每个向量元素是Emotion_Label部分勾选的情况,表情被勾选是1,否则为0;Name表示表情图像的命名,实验过程中可以通过Name索引到标签对应的表情图像;value表示Part_Label,是一个1×5的向量,每个向量元素对应的是Part_Label部分中每个部件勾选的位置;Coefficient是置信度。

每张表情图像中3人分别标注,因此改进后的FELW包含了灰度表情图像和其对应的3组.mat文件,.mat文件包含了Part_Label和Emotion_Label信息。对3组数据进行分析,将3组数据结合成一组数据。

1.3.2 Kappa一致性检验

Kappa一致性检验是用Kappa系数来检验两种检验结果是否一致,可以根据Kappa系数的取值大小来衡量两种结果的一致性程度,Kappa系数的值越大说明两种结果的一致性越高^[3-4]。Kappa系数是从比较两个观测者对同一事物的观测结果或同一观测者对同一事物的两次观测结果是否一致出发,用于由机遇造成的一致性和实际观测的一致性之间的差别大小作为评价基础所提出的统计指标^[5]。Kappa公式如下

$$k = \frac{p_o - p_e}{1 - p_e} \quad (1)$$

式中: p_e 是每一类正确分类的样本数量之和除以总样本数,也就是总体分类精度;假设每一类的真实样本个数分别是 a_1, a_2, \dots, a_c ,预测出来的每一类样本个数分别是 b_1, b_2, \dots, b_c ,样本个数为 n ,则有 $p_e = \frac{a_1 \times b_1 + a_2 \times b_2 + \dots + a_c \times b_c}{n \times n}$ 。3组EmotionLabel两两进行Kappa一致性检验, $k_{ij} \geq 0.4$ (k_{ij} 表示EmotionLabel(i)和EmotionLabel(j)进行Kappa一致性检验得到的Kappa值)时,EmotionLabel(i)和EmotionLabel(j)与其Coefficient值进行下一步计算,公式如下

$$\text{EmotionLabel}' = \text{round} \left(\frac{\sum_{k=1}^n \text{EmotionLal}(k) * \text{Coefficient}(k)}{n} \right) \quad (2)$$

式中: n 表示经过Kappa一致性检验筛选下的EmotionLabel组个数, k 表示筛选出下的EmotionLabel组标号,对向量中每个元素四舍五入取整。最后得到的EmotionLabel'也是一个软标签,EmotionLabel'的向量元素的值也在0~5之间。

1.3.3 数据的合成

由于Part_Label的标注是单选的,将3组PartLabel值的相同位置的向量元素值进行比较,若3组相同位置的向量元素都不相同,则随机选择一个元素作为最终元素;若3组相同位置的向量元素有任意两个元素相同,则选取这个元素作为最终元素。

Emotion_Label是多选,则先将三组EmotionLabel用Kappa一致性检验分析的结果结合Coefficient得出最后的EmotionLabel'。

为了区别Kappa一致性检验对数据集的识别率的影响,本文将原始3组标注数据不经过Kappa一致性检验直接用式(2)进行结合生成EmotionLabel'向量,最后组成数据集FELW-1,用于FELW的对比实验。

2 人脸表情识别

2.1 传统的人脸表情识别方法

人脸表情识别系统一般包括人脸检测、图像预处理、表情特征提取和表情分类4部分^[6]。本文采用Gabor + 2DPCA + SVM和Curvelet + 2DPCA + BP两种传统的人脸表情方法进行人脸表情识别。下面介绍两种本文使用的传统人脸表情识别方法。本文数据集的表情图像已经是120像素×120像素的灰度图,可以省去人脸检测和预处理的部分,下面方法只包含了特征提取和表情识别的步骤。

(1) Gabor + 2DPCA + SVM:先将网格化的表情子图像进行Gabor小波变换^[7],然后用2DPCA^[8-10]对子图进行降维,提取特征主成分。最后用SVM分类器进行表情分类。图5(a)给出了Gabor + 2DPCA + SVM方法流程图。

(2) Curvelet + 2DPCA + BP:对输入样本图像进行Curvelet变换^[11]后再用2DPCA进行降维提取特征主成分。最后用BP神经网络^[12]进行表情分类。图5(b)给出了Curvelet + 2DPCA + BP方法流程图。

2.2 基于深度学习的人脸表情识别

本文表情数据集是大小为120像素×120像素的图像,表情分类是10类。所以模型输入是120×120的图像矩阵,输出是大小1×10的向量。本文使用了用于人脸表情识别的3个卷积神经网络模型FERNet (Facial Expression Recognition Net),图6是本文采用的3个CNN模型。

如图6(a)所示,模型1输入120像素×120像素的灰度表情图像,经过连续两个卷积层,一个池化层,再经过一个卷积层,一个池化层,然后经过三个连续的卷积层,再进行一个池化层,最后经过3个全连接层,输出一个1×10的向量。有12层网络,前3个卷积层进行ReLU^[13]调整,再进行LRN^[14]归一化后输出;后3个卷积层只经过ReLU调整。

如图6(b)所示,模型2是参照Zeiler & Fergus二人提出的ZF模型^[15]设计的一个CNN模型。输入120像素×120像素的灰度表情图像,经过一个卷积层,一个池化层,再经过一个卷积层,一个池化层,然后经过3个连续的卷积层,再进行一个池化层,最后经过3个全连接层,输出一个1×10的向量。有11层网络,虽然网络层数没有模型1多,但是其卷积层的卷积核个数远多于模型1。前3个卷积层进行ReLU调整,再进行LRN归一化后输出;后3个卷积层只经过ReLU调整。

如图6(c)所示,模型3与模型2的网络类似,但是没有添加LRN,目的是区别未使用LRN对卷积和ReLU的影响,体现出LRN是否会对实验结果产生影响。。输入120像素×120像素的灰度表情图像,经过连续两个卷积层,一个池化层,再经过一个卷积层,一个池化层,然后经过3个连续的卷积层,再进行一个池化层,最后经过3个全连接层,输出一个1×10的向量。整个网络的卷积层只经过ReLU调整。

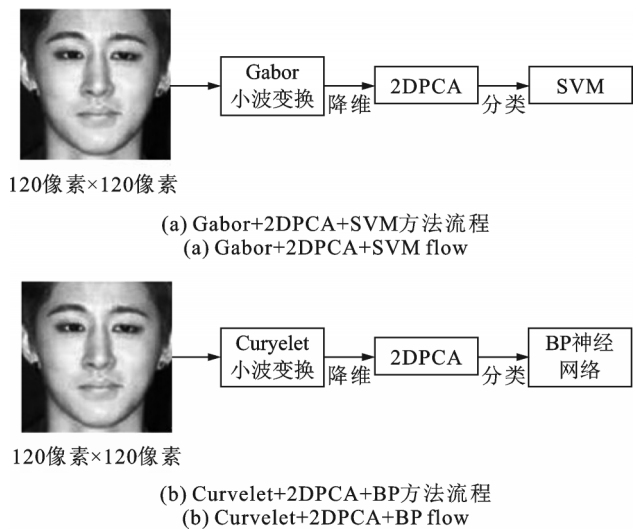


图5 两种传统人脸表情识别方法流程图
Fig.5 Flowchart of two traditional facial expression recognition methods

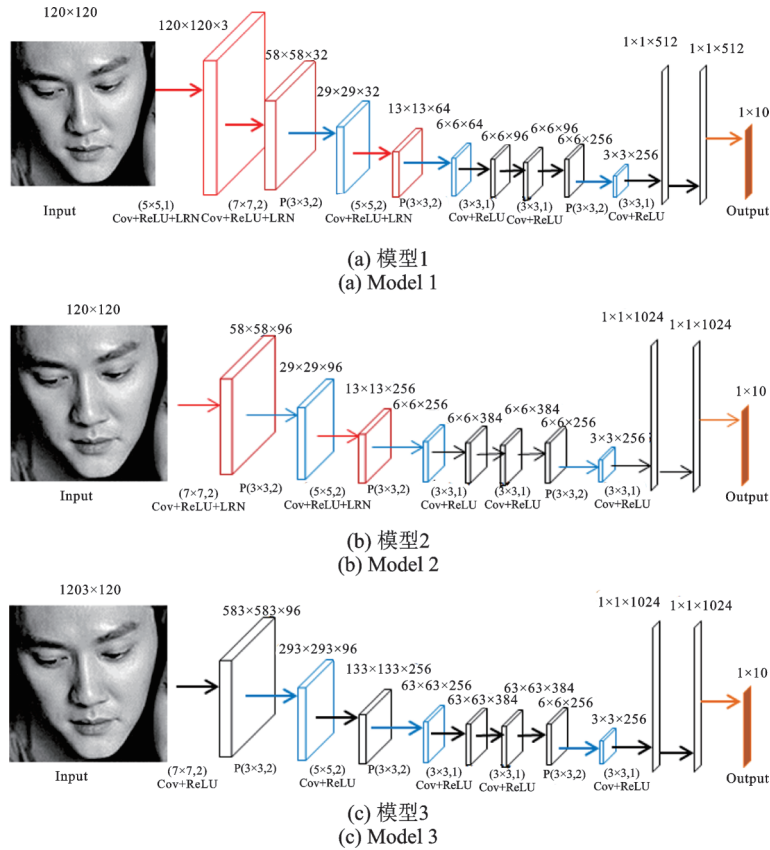


图6 本文采用FERNet模型

Fig.6 FERNet model

3 实验结果与分析

将FELW 分别用传统的人脸表情识别方法和深度学习的人脸表情识别方法进行实验,并用 JAFFE数据集和FELW 系列数据集进行实验对比。本文实验中用到的术语:

软标签:由于FELW 系列数据集图像的 Emotion_Label 不止一种表情,将采用所有表情实验的表情标签称为软标签。

硬标签:在 Emotion_Label 向量元素中选择最大值作为该图像的表情标签。

FELW-原:未改进的数据集。

FELW:改进后的数据集。

FELW-1:未经过Kappa 一致性检验数据集。

3.1 传统的人脸表情识别实验

(1) 实验 1

在FELW-原、FELW 和FELW-1 数据集和 JAFFE 数据集上分别用上述两种传统方法进行实验。由于 JAFFE 数据集只有7种表情且样本太少。首先,将FELW 系列数据集转化为硬标签形式进行实验;然后,选取跟 JAFFE 数据集一样的表情类别,每种表情30张图像作为实验样本,Train:Test=1:1 或2:1 进行实验,取200次实验平均值。实验结果如表1所示。

(2) 实验 2

由于 JAFFE 数据集没有软标签,并且两种传统方法中只有 Gabor+2DPCA+SVM 方法可以进行

软标签实验,所以接下来的实验仅在FELW系列数据集上采用软标签形式进行Curvelet+2DPCA+BP实验。本文做了10组实验,用了5种不同的实验样本数量,分别是180,210,240,300,360,每种数量的Train:Test=1:1或2:1进行实验,实验结果如表2所示。

表1 JAFFE数据集和使用硬标签的三组数据集传统实验结果

Tab. 1 Traditional experimental results for JAFFE datasets and three sets of datasets using hard tags

数据集	Train/ Test	传统实验方法准确率/%	
		Curvelet+2DP- CA+BP	Gabor+2DP- CA+SVM
		FELW-原	15/15
FELW-原	20/10	14.29	14.285 7
FELW	15/15	24.38	22.61
FELW	20/10	28.67	26.78
FELW-1	15/15	22.12	20.31
FELW-1	20/10	23.33	20.89
JAFFE	15/15	61.25	55.48
JAFFE	20/10	68.86	59.60

(3) 实验3

在FELW系列数据集中采用Part_Label+软标签形式进行Curvelet+2DPCA+BP方法实验。实验样本参数与实验2一样,实验结果如表3所示。

(4) 实验4

由于FELW2.0的图像没有经过图像配准,并不全是正脸,可能会影响人脸表情识别。将FELW数据集和FELW-1数据集挑出正脸图像重复实验3,实验结果如表4所示。

表3 Part_Label+软标签形式对3组数据集进行Curvelet+2DPCA+BP实验准确率

Tab. 3 Curvelet+2DPCA+BP experimental accuracy for three sets of data sets in the form of Part_Label+ soft label

实验样本 总数	Train/ Test	不同数据集实验准确率/%		
		FELW-原	FELW	FELW-1
180	90/90	52.47	57.48	52.56
180	120/60	53.77	59.67	53.34
210	105/105	60.31	64.34	61.13
210	140/70	61.07	64.67	61.56
240	120/120	61.37	65.43	62.87
240	160/80	62.80	67.66	62.45
300	15/150	56.55	63.42	59.67
300	200/100	57.91	62.66	58.89
360	180/180	63.76	67.90	64.43
360	240/120	65.01	69.09	63.12

表2 软标签形式对3组数据集进行Curvelet+2DPCA+BP实验准确率

Tab. 2 Curvelet+2DPCA+BP experimental accuracy for the three data sets in the form of soft tag

实验样本 总数	Train/ Test	不同数据集实验准确率/%		
		FELW-原	FELW	FELW-1
180	90/90	41.67	46.88	40.89
180	120/60	43.73	48.23	43.97
210	105/105	44.01	50.45	46.11
210	140/70	44.98	48.11	43.01
240	120/120	40.38	44.92	40.19
240	160/80	41.49	46.44	40.34
300	150/150	40.12	45.89	40.99
300	200/100	41.50	46.34	41.89
360	180/180	46.91	51.69	46.13
360	240/120	47.24	51.33	45.91

表4 正脸+Part_Label+软标签形式对3组数据集进行Curvelet+2DPCA+BP实验准确率

Tab. 4 Curvelet+2DPCA+BP experimental accuracy for three sets of data sets in the form of positive face +Part_Label+ soft label

实验样本 总数	Train/ Test	不同数据集实验准确率/%	
		FELW	FELW-1
180	90/90	62.43	60.90
180	120/60	64.67	60.06
210	105/105	69.43	63.05
210	140/70	69.23	64.23
240	120/120	69.59	64.69
240	160/80	72.93	67.27
300	15/150	68.26	63.17
300	200/100	66.49	61.18
360	180/180	72.05	67.11
360	240/120	74.91	69.48

(5) 实验分析

实验1可知,JAFPE数据集的识别率在两种传统人脸表情识别方法中识别率比LFEW-原,LFEW和LFEW-1高出很多。原因是:JAFPE数据集图像都是正脸,而LFEW系列数据集是自然状态下的人脸表情图像,图像复杂;训练样本太少使得训练模型获取的特征表达不够全面;强制转换成硬标签也可能导致效果不好。

将实验2,3,4结果统计得到图7的折线图,从图中可以看出:①由折线图趋势可以看出,样本越多,识别率越高;Train:Test=2:1的识别率普遍比Train:Test为1:1的识别率高。②使用相同数据集,采用Part_Label加软标签的形式实验效果比仅采用软标签形式的效果更好。③使用相同数据集,筛选正脸的实验效果比未筛选正脸的效果更好。④使用相同形式实验时,FELW数据集实验效果最好。

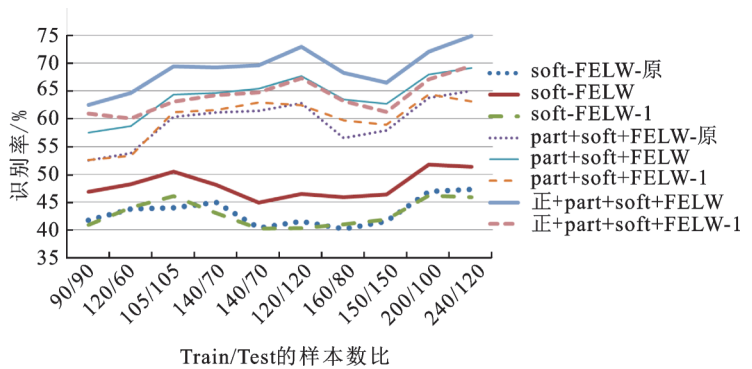


图7 实验2—4的结果统计

Fig.7 Statistically results of Experiments 2—4

3.2 基于深度学习的人脸表情识别实验

数据集包含26 848张图像,从中随机选择20 000个样本进行训练,其余的作为测试样本,训练时将20 000个样本平均分成10份,分次取9份作为训练样本,剩下1份则作为验证样本。每个数据集进行10次实验,最终的识别率是10次实验的识别率的平均值。

(1) 实验1

依照本章所建立的模型,采用硬分类规则进行表情识别,所得结果如表5所示。

(2) 实验2

为了验证软标签的作用,使用软标签形式进行识别,将软标签用在网络最后一层,其他实验参数与实验1一致,实验结果如表6所示。

表5 硬标签形式在3种深度模型上实验准确率

数据集	识别方法			%
	6卷积+3全连接(模型1)		5卷积+3全连接	
	模型2	模型3		
FELW-原	46.72	56.53	46.22	
FELW	50.31	59.44	50.12	
FELW-1	48.19	57.87	47.09	

表6 软标签形式在3种深度模型上实验准确率

数据集	识别方法			%
	6卷积+3全连接(模型1)		5卷积+3全连接	
	模型2	模型3		
FELW	55.01	65.73	55.45	
FELW-1	52.29	61.71	52.94	

(3) 实验3

将Part_Label标签作为特征的表达,与原图像一起作为网络的输入。即Part_Label+软标签形式识别,其他实验参数与实验1一致,实验结果如表7。

(4) 实验4

同时为了提高识别率采用筛选正脸的方式进行实验。使用软标签+正脸进行识别(软标签使用在

模型最后一层),实验结果如表8所示。

表7 Part_Label+软标签形式在3种深度模型上实验准确率

Tab. 7 Part_Label+ soft label form experiment in three depth models %

数据集	识别方法		
	6卷积+3全连接(模型1)	5卷积+3全连接	
	模型2	模型3	
FELW	61.45	72.22	60.09
FELW-1	59.15	66.71	58.56

表8 正脸+Part_Label+软标签形式在3种深度模型上实验准确率

Tab. 8 Positive face +Part_Label+ soft label form experiment in three depth models %

数据集	识别方法		
	6卷积+3全连接(模型1)	5卷积+3全连接	
	模型2	模型3	
FELW	66.45	78.22	64.09
FELW-1	64.15	71.71	63.56

(5) 实验分析

将实验1—4结果统计得到柱形图8,由图8可以看出:①在相同数据集实验时,模型1识别率比模型3略高,因为模型1的网络更深;模型2实验识别率比模型1高,因为模型2使用的卷积核个数远比模型1多;模型2比模型3高,因为模型2使用了LRN。②采用相同形式使用相同模型实验时,FELW数据集识别率最高。③相同数据集在相同模型实验时,正脸+Part_Label+软标签的形式识别率最高。

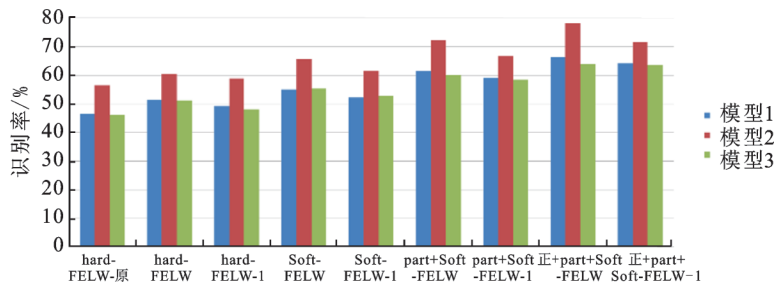


图8 实验1—4准确率统计柱形图

Fig.8 Accuracy histogram of Experiments 1—4

4 结束语

本文介绍了人脸表情数据集FELW的建立与改进,并另建了一个FELW-1数据集与其对比。对这3组数据集分别用传统人脸表情识别方法和深度学习的人脸表情识别方法实验。两种方法实验结果表明:FELW数据集的识别率比FELW-原、FELW-1数据集识别率高,同时体现了Kappa一致性检验对数据集的作用;采用正脸+Part_Label+软标签的形式的实验识别率效果最好;采用深度学习的人脸表情识别方法时,使用LRN的模型2比没有使用LRN的模型3识别率高,同样体现了LRN的作用。由于FELW数据集图像是未经过图像配准的归一化操作,而且图像不是在实验室采集而是在网络上采集的自然场景下的各种类型图片,图像的质量远远比不上现在成熟的人脸表情数据集图像质量,但是在传统方法和深度学习方法的识别率达到了70%以上的效果,后续还可以继续改进数据集以达到更好的识别率。

参考文献:

- [1] 周飞燕,金林鹏,董军.卷积神经网络研究综述[J].计算机学报,2017,40(6):1229-1251.
Zhou Feiyan, Jin Linpeng, Dong Jun. Review of convolutional neural network[J]. Chinese Journal of Computers, 2017, 40(6):

- 1229-1251.
- [2] 甘荧. 自然场景下人脸表情数据集的构建与应用研究[D]. 南昌:江西师范大学, 2016.
Gan Ying. Construction and application research of facial expression datasets in wild[D]. Nanchang: Jiangxi Normal University, 2016.
- [3] 华琳, 阎岩, 张建. 关于对诊断一致性Kappa系统的探讨[J]. 数理医药学杂志, 2006(5):518-520.
Hua Lin, Yan Yan, Zhang Jian. Discussion on diagnostic consistency Kappa system[J]. Journal of Mathematical Medicine, 2006 (5):518-520.
- [4] Cyr L, Francis K. Measures of clinical agreement for nominal and categorical data: The Kappa coefficient[J]. Computer Sin Biology and Medicine, 1992, 22(4):239-246.
- [5] 文万青. 关于Kappa系数的进一步评价和校正[J]. 中国卫生统计, 1994(6):4-7.
Wen Wanqing. Further evaluation and correction of Kappa coefficient [J]. Chinese Journal of Health Statistics, 1994(6):4-7.
- [6] 李思泉, 张轩雄. 基于卷积神经网络的人脸表情识别研究[J]. 软件导刊, 2018, 17(1):28-31.
Li Siqian, Zhang Xuanxiong. Study on face expression recognition based on convolutional neural network[J]. Software Guide, 2018, 17(1):28-31.
- [7] 陈海永, 徐森, 刘坤, 等. 基于Gabor小波和加权马氏距离的带钢表面缺陷检测[J]. 电子测量与仪器学报, 2016, 30(5):786-793.
Chen Haiyong, Xu Sen, Liu Kun, et al. Strip defect detection based on Gabor wavelet and weighted Mahalanobis distance[J]. Journal of Electronic Measurement and Instrumentation, 2016, 30(5):786-793.
- [8] 陈伏兵, 陈秀宏, 张生亮, 等. 基于模块2DPCA的人脸识别方法[J]. 中国图象图形学报, 2006(4):580-585.
Chen Fubing, Chen Xiuhong, Zhang Shengliang, et al. A human face recognition method based on modular 2DPCA[J]. Journal of Image and Graphics, 2006(4):580-585.
- [9] 颜雪军, 赵春霞, 袁夏. 2DPCA-SIFT:一种有效的局部特征描述方法[J]. 自动化学报, 2014, 40(4):675-682.
Yan Xuejun, Zhao Chunxia, Yuan Xia. 2DPCA-SIFT: An efficient local feature descriptor[J]. Acta Automatica Sinica, 2014, 40(4):675-682.
- [10] 赵丽美, 贾维敏, 王标标, 等. 基于最大相关熵准则的2维主成分分析[J]. 中国图象图形学报, 2015, 20(12):1684-1688.
Zhao Limei, Jia Weimin, Wang Biaobiao, et al. Two-dimensional principal component analysis based on maximum correntropy criterion[J]. Journal of Image and Graphics, 2015, 20(12):1684-1688.
- [11] 杨勇, 童松, 黄淑英. 快速离散Curvelet变换域的图像融合[J]. 中国图象图形学报, 2015, 20(2):219-228.
Yang Yong, Tong Song, Huang Shuying. Image fusion based on fast discrete Curvelet transform [J]. Journal of Image and Graphics, 2015, 20(2):219-228.
- [12] Bai Yun, Li Yong, Wang Xiaoxue, et al. Air pollutants concentrations forecasting using back propagation neural network based on wavelet decomposition with meteorological conditions[J]. Atmospheric Pollution Research, 2016, 7(3):1-5.
- [13] Krizhevsky A, Sutskever I, Hinton G E. ImageNet classification with deep convolutional neural networks [J]. Communications of the ACM, 2012, 60 (6): 84-90.
- [14] LeCun Y, Boser B, Denker J S, et al. Backpropagation applied to handwritten zip code recognition[J]. Neural Computation, 1989, 1(4): 541-551.
- [15] Zeiler M D, Fergus R. Visualizing and understanding convolutional networks[C]//Computer Vision-ECCV 2014. [S.l.]: Springer International Publishing, 2014: 818.

作者简介:



叶继华(1966-),男,教授,博士生导师,研究方向:数据融合、物联网技术、嵌入式系统、图像处理, E-mail: yjhwcl@163.com。



刘燕(1993-),女,硕士研究生,研究方向:数据融合、图像处理。



李汉曦(1982-),男,博士,特聘教授,研究方向:计算机视觉。



甘荧(1991-),女,硕士,研究方向:图像处理。