

文章编号:1008-1542(2019)06-0540-08

基于多特征融合卷积神经网络的人脸表情识别

王建霞, 陈慧萍, 李佳泽, 张晓明

(河北科技大学信息科学与工程学院, 河北石家庄 050018)

摘要:针对卷积神经网络特征提取不够充分且识别率低等问题,提出了一种多特征融合卷积神经网络的人脸表情识别方法。首先,为了增加网络的宽度和深度,在网络中引入 Inception 结构来提取特征的多样性;然后,将提取到的高层次特征与低层次特征进行融合,利用池化层的特征,将融合后的特征送入全连接层,对其特征进行融合处理来增加网络的非线性表达,使网络学习到的特征更加丰富;最后,输出层经过 Softmax 分类器对表情进行分类,在公开数据集 FER2013 和 CK+ 上进行实验,并且对实验结果进行分析。实验结果表明:改进后的网络结构在 FER2013 和 CK+ 数据集的面部表情上,识别率分别提高了 0.06% 和 2.25%。所提方法在人脸表情识别中对卷积神经网络设置和参数配置方面具有参考价值。

关键词:计算机图像处理;面部表情识别;卷积神经网络;特征融合;特征提取;表情分类

中图分类号:TP319 文献标志码:A doi:10.7535/hbkd.2019yx06012

Facial expression recognition based on multi-feature fusion convolution network

WANG Jianxia, CHEN Huiping, LI Jiaze, ZHANG Xiaoming

(School of Information Science and Engineering, Hebei University of Science and Technology, Shijiazhuang, Hebei 050018, China)

Abstract: Aiming at the problem of insufficient feature extraction and low recognition rate of convolutional neural network, a novel facial expression recognition method based on multi-feature fusion convolutional neural network is proposed. First, to increase the width and depth of the network, Inception architecture is introduced into the network to extract the diversity of features; Then, the extracted high-level features are fused with the low-level features, and the pooled features are used to send the fused features into the full connection layer, then the fused features are processed to increase the non-linear expression of the network and enrich the features learned by the network. Finally, the output layer classifies the expressions by Softmax classifier, conducts experiments on FER2013 and CK+, and analyzes the experimental results. Experimental results show that

收稿日期:2019-09-09;修回日期:2019-10-22;责任编辑:陈书欣

基金项目:河北省自然科学基金(F2018208116)

第一作者简介:王建霞(1970—),女,河北临城人,教授,硕士,主要从事网络与数据库、图像处理方面的研究。

E-mail:1084473634@qq.com

王建霞,陈慧萍,李佳泽,等.基于多特征融合卷积神经网络的人脸表情识别[J].河北科技大学学报,2019,40(6):540-547.

WANG Jianxia, CHEN Huiping, LI Jiaze, et al. Facial expression recognition based on multi-feature fusion convolution network[J]. Journal of Hebei University of Science and Technology, 2019, 40(6): 540-547.

the improved network structure improves the recognition rate of facial expressions in FER2013 and CK+ data sets by 0.06% and 2.25%, respectively. The proposed method is valuable for setting up convolution neural network and parameter configuration in facial expression recognition.

Keywords: computer image processing; facial expression recognition; convolutional neural network; feature fusion; feature extraction; expression classification

人脸面部表情在人们的日常交流中起到了举足轻重的作用,人们可以通过情绪的表达来传递一些信息。人脸表情识别现如今广泛应用于计算机视觉、公安安防、心理治疗和人际交互等领域^[1]。但是由于人脸面部表情识别具有复杂性和多样性等特点,这就使得人脸面部表情的识别有了很大的难度。

人脸表情识别大致包括图像采集、图像预处理、特征提取和表情识别分类 4 个步骤,其中特征提取是最关键的步骤,研究者们也针对这方面进行了深入研究,提出了很多方法。文献[2]提出从面部显著区域中提取 LBP 和 HOG 特征,用 Z-score 方法对这两种方法进行融合实现表情识别。文献[3]融合 Gabor 和 LBP 特征向量进行表情识别,结果表明,与单一方法相比较,融合后的效果更好。文献[4]使用 Fisher 线性判别法来改进传统的主成分分析法,通过增加类间离散矩阵以增大类别之间的距离,应用于矿工面部表情的识别,但其对光照因素比较敏感。

传统的识别方法虽然能得到较好的分类效果,但是太依赖于前期的人工提取特征,人对其干扰因素比较大。目前,由于卷积神经网络的无监督学习特性,使得特征提取和识别变得更加简单,因此在表情识别中,也常常用到卷积神经网络。文献[5]提出微笑检测的面部表情识别方法,使用卷积神经网络在 GPU 上对数据集进行训练和测试。文献[6]首次将跨连卷积网络模型应用于人脸表情识别,在 LeNet-5 网络的基础上进行改进,将网络中的池化层与全连接层融合,提取图像中的高层次特征与低层次特征,取得了不错的分类效果。文献[7]提出一种加权混合神经网络,将提取的灰度图像在经过 ImageNet 数据库训练的 VGG-16 网络上进行参数初始化,将图像提取 LBP 特征后送入由 DeepID 改进的浅层网络中,对这 2 个通道的输出以加权的方式融合,实验证明,该方法优于手工特征和单通道的神经网络。文献[8]提出融合神经网络结构和条件随机场的方法,在网络中使用了 3 个 Inception-ResNet 模块,大大提高了识别能力。文献[9]提出了一种新的神经网络岛屿损失模型,增加了类间差异,使类内距离变小。文献[10]提出了一种双通道卷积神经网络,将提取的眼睛输入第 1 通道,嘴巴输入第 2 通道,通道融合送入全连接层,提高分类的精度。文献[11]提出了 EmotionDAN 模型,将面部表情地标作为分类损失函数的一部分,同时扩展了深度对齐网络,取得了不错的效果。

本文提出了一种多特征融合卷积神经网络方法,既考虑了图像的高层次特征,又考虑了图像的低层次特征,同时又将具有降维功能的 Inception v1 模块引入卷积神经网络中,通过拓宽卷积神经网络的深度和宽度来增强提取特征的能力,并将此模型应用于人脸表情识别。本文中的神经网络借鉴了 VGG 网络,但是又不同于 VGG 网络,没有完全采用 3×3 的卷积核,而是采用 3×3 和 5×5 的卷积核, padding 采用“SAME”,池化层均采用 2×2 的池化区域,激活函数的使用更能凸显表情的边缘信息和纹理信息。为了防止过拟合,在全连接层引入了 Dropout 技术。实验结果表明,改进后的网络能有效提高人脸表情识别的准确率,达到更好的分类效果。

1 VGG-16 网络结构

VGG 是在 AlexNet 基础上进行改进的,并在 ILSVRC^[12]2014 年比赛中取得了分类项目第 2 名和定位项目第 1 名的好成绩。VGG-16 网络输入图片格式为 $224 \times 224 \times 3$,采用的卷积核尺寸为 3×3 ,卷积步长设置为 1,采用的最大池化区域为 2×2 ,步长为 2。在经过 13 个卷积层和 5 个最大池化层处理后,输出一个 4 096 维度的数据,然后经过 3 个全连接层处理,最终输入 Softmax 对图像进行分类,将分类结果输出。采用 2 个 3×3 的卷积核代替 1 个 5×5 的卷积核,3 个 3×3 的卷积核代替 1 个 7×7 的卷积核,在保证感受野相同的同时,也加深了网络的深度。增加了 ReLU^[13] 激活函数的使用次数,使特征的学习能力增强;使用最大池化层,不仅减少了网络中的参数,而且更有利于保持图像的性质。VGG-16 网络模型如图 1 所示。为了提高表情识别的识别率和鲁棒性,在文中网络架构方面做了改进。

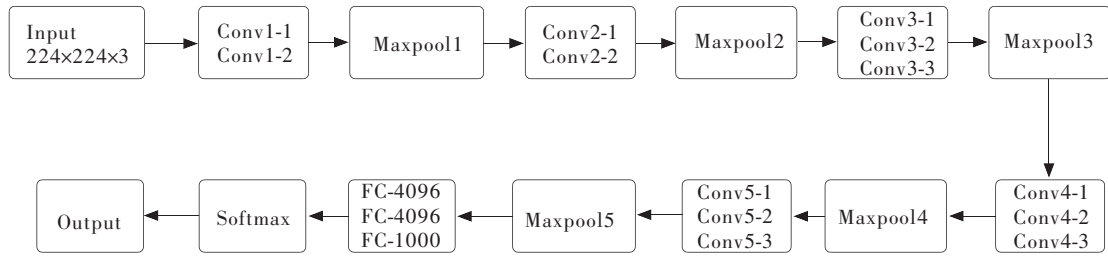


图1 VGG-16 网络结构图

Fig.1 VGG-16 network structure diagram

2 跨连卷积神经网络的改进

2.1 VGG 网络模型的改进

传统的神经网络只是把提取到的特征向后传输,没有很好地将低层次的特征进行有效利用,从而对分类结果产生影响。同时,网络层数的不断加深会产生梯度的弥散问题,进而对网络的训练也会产生一定影响。为了克服这一传统缺点,在本文中引入了跨连思想^[14],用于人脸表情的分类。该网络输入一幅图像 X ,然后分别交错经过卷积层(L1,L3,L5)、池化层(L2,L4,L6,L8)、Inception v1 模块、2 个全连接层(L9,L10)和输出层(O)。将池化层 L2,L4,L6,L8 提取到的特征直接送入全连接层进行特征融合并分类,输出层有 7 个分类。

VGG 网络通过相同的卷积核进行特征提取,在每个卷积层的后边均加入了 ReLU 激活函数以及批量正则化,可以缓解过拟合问题。改进后的网络结构是基于 VGG-16 网络结构修改得到的,输入大小为 44 像素 \times 44 像素,卷积层均采用 3×3 和 5×5 的卷积核,步长为 1,采用 padding="SAME",即卷积后保持图像空间分辨率不发生改变,其中 5×5 的卷积核可以学习到更多的空间特征。一次卷积就等价于做了一次特征变换,不同卷积核的使用,增加了网络模型对不同尺度的适应性。池化层跟在卷积层的后面,采用 2×2 的滑动窗口,步长设置为 2。全连接层中的 Dropout 设置为 0.5,即丢弃网络中一半的神经网络单元,有效防止过拟合,使网络的泛化能力更强,因而减少了网络的复杂度,同时也加快了运算速度。每个卷积层的后边都会添加 BN 层,对每层输入数据进行标准化处理。

网络中的激活函数使用缩放指数线性单元(SeLU 函数)^[15],可以让网络收敛速度更快,相较于 ReLU 来说,更不易发生过拟合。SeLU 函数如式(1)所示,其中 $\lambda=1.050\ 700\ 987\ 355\ 480\ 493\ 419\ 334\ 985\ 294\ 6$, $\alpha=1.673\ 263\ 242\ 354\ 377\ 284\ 817\ 042\ 991\ 671\ 7$ 。

$$\text{SeLU}(x) = \lambda \begin{cases} x, & x > 0, \\ \alpha e^x - \alpha, & x \leq 0. \end{cases} \quad (1)$$

2.2 Inception 模块

Inception 采用具有降维功能的 Inception v1 改进模块,使用 Inception v1 模块中的 1×1 卷积核,不仅减少了参数的数量,而且降低了特征图的通道数。

针对 VGG 网络在面部表情识别出现的问题,进行了一系列的改进。

1)将网络的深层特征与不同的浅层特征通过 Concat 连接到全连接层,将深浅特征融合到一起,能更好地提取表情特征。

2)将网络的第 4 个卷积用 Inception v1 结构取代,不仅增加了非线性的表示以及网络深度,同时提升了提取特征的丰富性,增强了网络的表达能力。

3)网络最后一个池化层使用平均池化,其余的池化使用最大池化。

本文采用的跨连卷积神经网络模型如图 2 所示。

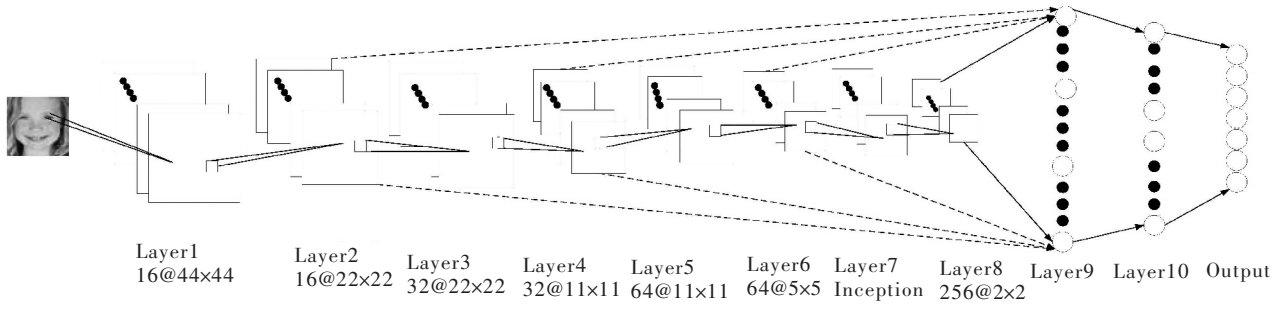


图 2 改进后的跨连卷积神经网络结构示意图

Fig.2 Improved cross-connected convolutional neural network structure diagram

2.3 改进后的跨连卷积神经网络模型描述

$$\left\{ \begin{aligned}
 &L_{1,j}^l = \sigma(x^l \times W^{1,j} + b^{1,j}), \quad 1 \leq j \leq 16, \\
 &L_{2,j}^l = \text{down}_{\lambda,\tau}(L_{1,j}^l), \quad 1 \leq j \leq 16, \\
 &L_{3,j}^l = \sigma\left(\sum_{i=1}^6 L_{2,i}^l \times W^{3,j} + b^{3,j}\right), \quad 1 \leq j \leq 32, \\
 &L_{4,j}^l = \text{down}_{\lambda,\tau}(L_{3,j}^l), \quad 1 \leq j \leq 32, \\
 &L_{5,j}^l = \sigma\left(\sum_{i=1}^{12} L_{4,i}^l \times W^{5,j} + b^{5,j}\right), \quad 1 \leq j \leq 64, \\
 &L_{6,j}^l = \text{down}_{\lambda,\tau}(L_{5,j}^l), \quad 1 \leq j \leq 64, \\
 &L_{7,j}^l = \text{FilterConcat}(L_{(7-1),j}^l, L_{(7-2),j}^l, L_{(7-3),j}^l, L_{(7-4),j}^l), \quad 1 \leq j \leq 256, \\
 &L_{8,j}^l = \text{down}_{\lambda,\tau}(L_{7,j}^l), \quad 1 \leq j \leq 256, \\
 &L_9^l = \sigma\left(\sum_{i=1}^{16} L_{2,i}^l \times W^{9-2} + \sum_{i=1}^{32} L_{4,i}^l \times W^{9-4} + \sum_{i=1}^{64} L_{6,i}^l \times W^{9-6} + \sum_{i=1}^{256} L_{8,i}^l \times W^{9-8} + b^{9,j}\right), \\
 &L_{10}^l = L_{9,(1,2,\dots,256)}^l, \\
 &O = \text{soft max}(h_9^l W^{10} + b^{10}).
 \end{aligned} \right. \tag{2}$$

式中: $\sigma(\cdot)$ 函数为 ReLU 激活函数; $\text{down}_{\lambda,\tau}(\cdot)$ 为下采样; $W^{i,j}$ 表示第 i 个隐含层的第 j 个卷积核; $b^{i,j}$ 表示第 i 个隐含层的第 j 个偏置; $L_{m,j}^l$ 表示第 l 个样本的第 m 个隐含层的第 j 个卷积面(或池化面); L_9^l 表示第 9 个隐含层的输出值; L_{10}^l 表示第 10 个隐含层的输出值; O 表示输出层。其中将 $L_{6,j}^l$ 和 $L_{8,j}^l$ 融合到全连接层,满足高低层次的特征融合。在 Inception 结构的 4 个分支中全部采用 64 的通道数,输出 $5 \times 5 \times 256$ 。

对于第 l 个样本 x^l ,改进后的跨连卷积神经网络的计算过程如下:

网络训练时分为正向传播和反向传播。正向传播主要通过卷积和池化来提取图像所含的隐层特征,反向传播会根据误差调整参数值,不断地进行迭代,直到模型达到收敛。本文采用随机梯度下降算法,对权重参数进行更新。改进后的跨连卷积神经网络结构参数描述如表 1 所示。

表1 改进后的跨连卷积神经网络结构参数描述

Tab.1 Description of improved cross-connected convolutional neural network structure parameters

输入	输入类型	区域尺寸	步长	输出尺寸
X	Input	—	—	44×44×1
L1	Convolution	5×5	1	44×44×16
L2	Maxpooling	2×2	2	22×22×16
L3	Convolution	3×3	1	22×22×32
L4	Maxpooling	2×2	2	11×11×32
L5	Convolution	5×5	1	11×11×64
L6	Maxpooling	2×2	2	5×5×64
L7	Inception	15_1 : 1×1 15_2 : 1×1;3×3 15_3 : 1×1;5×5 15_4 : 3×3;1×1		5×5×256
L8	Meanpooling	2×2	2	2×2×256
L9	Fully-connected	—	—	2 048
L10	Fully-connected	—	—	1 024
O	Output	—	—	7

3 实验

实验是基于 Python3 的深度学习框架 Keras 进行的,使用的操作系统为 Windows10,硬件配置:中央处理器(CPU)为 Intel(R) Xeon(R) E5-2630 v2,主频为 2.60 GHz,内存为 32 GB;图形处理器(GPU)为 NVIDIA Quadro K2000,显存为 3 GB。安装了 NVIDIA CUDA Framework8.0 和 cuDNN 库。

3.1 数据集

在实验中分别采用 FER2013 表情数据集和 The extended Cohn-Kanade dataset (CK+)表情数据集进行实验。

FER2013 数据集来自于 2013 年 Kaggle 比赛中的数据集,包含了 35 887 张图片,在这里选用了其中的 28 709张图片用作训练集,3 589 张图片用作测试集,3 589 张图片用作验证集。每张图片都是固定大小为 48 像素×48 像素的灰度图。其中共包含了 7 种表情:愤怒、厌恶、恐惧、高兴、悲伤、惊讶、中性。图像涵盖了不同的人种、年龄、光照、遮挡、侧脸、姿势等,同时图像的分辨率偏低,使得在该数据集上进行表情识别具有挑战性。图 3 为 FER2013 数据集中的 7 种表情样例图像。



图3 FER2013 表情库部分图像

Fig.3 Partial images of FER2013 emoji library

CK+数据集是包含 123 个人、总共 593 个表情的视频序列样本。本实验仅取某一种表情高潮状态的 3 张图片,检测到人脸并且将人脸裁剪到 48 像素×48 像素大小。共包含 7 种表情,分别为生气、厌恶、害怕、开心、悲伤、惊讶及中性表情,图 4 为 CK+数据集内的 7 种表情样例图像。



图 4 CK+表情库部分图像

Fig.4 Partial images of CK+ emoji library

由于数据集中的数据量不足,因而需对数据进行数据增强处理,如表 2 所示。表 2 中,“Rotation_range”即为图片可以随机转动的角度;“Rescale”表示对图片进行随机缩放;“Shear_range”表示逆时针方向的随机剪裁角度;“Zoom_range”表示随机缩放的幅度;“Horizontal_flip”表示随机水平翻转;“Fill_mode”表示当进行变换时超出边界的点会根据设置的参数进行处理。

3.2 实验结果及分析

将图像处理成固定大小的数据格式,方便将数据输入到神经网络中。FER2013 数据集中人脸图像的大小为 48 像素×48 像素,将图像处理成大小为 44 像素×44 像素,对图像的像素归一化到[0,1]内,对图像进行数据扩充,扩充方法的参数如表 2 所示。通过这一方法,扩大了数据集的数量并实现了空间不变性。表 3 为 FER2013 数据集中的表情识别集。

表 2 数据增强参数

Tab.2 Data enhancement parameter

参数	参数值
Rotation_range	40
Rescale	1.0/255
Shear_range	0.2
Zoom_range	0.2
Horizontal_flip	True
Fill_mode	'nearest'

表 3 FER2013 数据集中的表情识别率

Tab.3 Expression recognition rate on FER2013 database

表情	愤怒	厌恶	恐惧	高兴	中性	悲伤	惊讶
愤怒	67.03	0.81	5.18	1.28	15.92	8.56	1.22
厌恶	11.12	73.29	3.64	2.31	7.82	0	1.82
恐惧	9.23	1.19	53.30	3.22	11.21	11.82	10.03
高兴	1.52	0	0.52	90.65	3.53	1.48	2.30
中性	6.90	0.17	8.08	5.72	65.76	10.01	3.36
悲伤	2.92	1.95	7.45	4.09	9.13	71.57	2.89
惊讶	1.40	0.32	3.03	6.31	4.34	1.34	86.29

CK+数据集采用 5-折交叉验证方法,将图像分成 5 份,每份都分别含有 7 种表情,每次实验时,选择其中的 4 份作为训练集,剩余的 1 份作为测试集。训练集用于训练分类器,测试集用于测试和得到模型。进行 5 次实验,最后取 5 次实验的平均结果作为最终的实验结果。表 4 为 CK+ 数据集中的表情平均识别率。

表 4 CK+数据集中的表情平均识别率

Tab.4 Average facial recognition rate in CK+ database

表情	生气	厌恶	害怕	开心	中性	悲伤	惊讶
生气	98.07	0.39	0.79	0	0.21	0.45	0.09
厌恶	1.91	94.25	1.83	0.17	0.42	0.86	0.56
害怕	1.64	3.15	94.65	0	0.12	0.31	0.13
开心	0.41	0.23	0.42	97.68	0.53	0.11	0.72
中性	1.25	0.37	0.42	0.51	96.63	0.34	0.48
悲伤	3.19	0.18	0.26	0.13	0.91	94.98	0.35
惊讶	0	0.31	1.02	0.27	0.26	0.22	97.92

从表3可以看出,高兴和惊讶的表情识别率最高,由此可见,这些表情相较于其他表情来说比较明显,动作幅度比较大。愤怒、厌恶、恐惧、悲伤的表情识别难度比较大,其中恐惧和悲伤表情比较相似,因为它们都具有额头紧皱和嘴唇拉开的特征。生气与厌恶表情具有相同的皱眉和嘴角等特征,所以识别表情的时候容易发生混淆。

从表4可以看出,算法在CK+数据集中的整体表现较好,但在识别厌恶、害怕和悲伤表情时准确率要低一些,主要原因可能是害怕、厌恶和悲伤之间表情的关系定义不是很明确,导致检测时识别率有所下降。

由表3和表4可知,同种算法在CK+表情库的识别效果要优于在FER2013表情库,造成这种结果的原因是FER2013表情库的影响因素比较多,更接近于现实生活中的图片,更具有代表性。因CK+是标准实验室图片,影响因素比较少,所以识别率相对于FER2013来说要高很多。

为了证明本文方法的有效性,在基准数据库的基础上,对比了本文算法与表情识别主流算法的性能。表5和表6分别为在FER2013和CK+中的实验结果对比。文献[16]使用纯卷积神经网络模型,表情识别率为57.10%;文献[17]使用与GoogleLeNet^[18]相似结构的Inception,表情识别率达到71.60%;文献[19]使用深层次的VGG网络,表情识别率为72.70%;文献[20]提出基于面部表情识别其中情绪装填的结果,以表情要素系数做特征进行计算,用3-NN和MLP神经网络进行特征分类;文献[5]在LeNet-5网络的基础上提出了跨连接网络,将低层次特征与高层次特征进行融合,以弥补样本数量的不足,表情识别率达到83.74%。由表5和表6对比结果可知,在这两种数据集中,与其他方法相比,本文算法有较好的识别效果。在FER2013数据集中,本文算法比其他方法的识别率至少提高了0.06%。在CK+数据集中,本文算法较其他方法的识别率至少提高了2.55%。

表5 不同算法在FER2013的识别率对比结果

Tab.5 Comparison results of recognition rates of different algorithms in FER2013

算法	识别率/%
CNN ^[16]	57.10
Inception ^[17]	71.60
VGG ^[18]	72.70
本文算法	72.76

表6 不同算法在CK+的识别率对比结果

Tab.6 Comparison results of recognition rates of different algorithms in CK+

算法	识别率/%
并行网络 ^[20]	94.03
跨连接 LeNet-5 ^[5]	83.74
本文算法	96.58

4 结 语

针对人脸表情识别效果欠佳、鲁棒性差等问题,提出了一个改进的跨连接卷积神经网络来识别7种面部表情。在VGG网络的基础上对网络进行了改进,将网络中的低层次特征与高层次特征进行融合,并将融合后的特征输入全连接层,充分利用了每个池化层的特征;引入了Inception网络结构,增加网络深度的同时,也增加了网络的宽度,使网络学习到的特征更加丰富,同时也加快了收敛速度。实验证明,本文的特征提取方法在FER2013和CK+数据集上的识别率有明显提高。本文仅考虑了表情图片比较复杂、各个表情的类内差距比较大的情况,后续工作还需要增大表情类间差距,降低类内差距,进一步提高表情的识别率。

参考文献/References:

- [1] 徐峰,张军平.人脸微表情识别综述[J].自动化学报,2017,43(3):333-348.
XU Feng, ZHANG Junping. Facial microexpression recognition: A survey[J]. Acta Automatica Sinica, 2017, 43(3):333-348.
- [2] LIU Yanpeng, LI Yibin, MA Xin, et al. Facial expression recognition with fusion features extracted from salient facial areas[J]. Sensors, 2017, 17(4): s17040712.
- [3] SUN Yuechuan, YU Jun. Facial expression recognition by fusing Gabor and local binary pattern features[C]//International Conference on Multimedia Modeling. Cham: Springer, 2017: 209-220.
- [4] 杜云,张璐璐,潘涛.基于改进的主成分分析法的矿工表情识别[J].河北科技大学学报,2019,40(1):45-50.
DU Yun, ZHANG Lulu, PAN Tao. Miner expression recognition based on improved principal component analysis[J]. Journal of Hebei University of Science and Technology, 2019, 40(1):45-50.

- [5] IVANOVSKY L, KHRYASHCHEV V, LEBEDEV A, et al. Facial expression recognition algorithm based on deep convolution neural network[C]//2017 21st Conference of Open Innovations Association (FRUCT). Helsinki:IEEE, 2017: 141-147.
- [6] 李勇, 林小竹, 蒋梦莹. 基于跨连接 LeNet-5 网络的面部表情识别[J]. 自动化学报, 2018, 44(1): 176-182.
LI Yong, LIN Xiaozhu, JIANG Mengying. Facial expression recognition with cross-connect LeNet-5 network[J]. Acta Automatica Sinica, 2018, 44(1): 176-182.
- [7] YANG Biao, CAO Jinmeng, NI Rongrong, et al. Facial expression recognition using weighted mixture deep neural network based on double-channel facial images[J]. IEEE Access, 2017:2784096.
- [8] HASANI B, MAHOOR M H. Spatio-temporal facial expression recognition using convolutional neural networks and conditional random fields[C]//2017 12th IEEE International Conference on Automatic Face & Gesture Recognition (FG 2017). Washington DC: IEEE, 2017: 790-795.
- [9] CAI Jie, MENG Zibo, KHAN A S, et al. Island loss for learning discriminative features in facial expression recognition[C]//2018 13th IEEE International Conference on Automatic Face & Gesture Recognition (FG 2018). Xi'an:IEEE, 2018: 302-309.
- [10] NWOSU L, WANG Hui, LU Jiang, et al. Deep convolutional neural network for facial expression recognition using facial parts[C]//2017 IEEE 15th Intl Conf on Dependable, Autonomic and Secure Computing, 15th Intl Conf on Pervasive Intelligence and Computing, 3rd Intl Conf on Big Data Intelligence and Computing and Cyber Science and Technology Congress. Orlando:IEEE, 2017: 1318-1321.
- [11] TAUTKUTE I, TRZCINSKI T, BIELSKI A. I know how you feel: Emotion recognition with facial landmarks[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition(CVPR) Workshops. [S.l.]:[s.n.], 2018: 1878-1880.
- [12] RUSSAKOVSKY O, DENG Jia, SU Hao, et al. ImageNet large scale visual recognition challenge[J]. International Journal of Computer Vision, 2015, 115(3): 211-252.
- [13] 龚震霆, 陈光喜, 曹建收. 卷积神经网络在脑脊液图像分类上的应用[J]. 计算机工程与设计, 2017, 38(4):1056-1061.
GONG Zhenting, CHEN Guangxi, CAO Jianshou. Cerebrospinal fluid images classification based on convolution neural network[J]. Computer Engineering and Design, 2017, 38(4):1056-1061.
- [14] 张婷, 李玉鑑, 胡海鹤, 等. 基于跨连卷积神经网络的性别分类模型[J]. 自动化学报, 2016, 42(6): 858-865.
ZHANG Ting, LI Yujian, HU Haihe, et al. A gender classification model based on cross-connected convolutional neural networks[J]. Acta Automatica Sinica, 2016, 42(6): 858-865.
- [15] KLAMBAUER G, UNTERTHINER T, MAYR A, et al. Self-normalizing neural networks[C]//Advances in Neural Information Processing Systems. Linz:[s.n.], 2017: 971-980.
- [16] TÜMEN V, SÖYLEMEZ Ö F, ERGEN B. Facial emotion recognition on a dataset using convolutional neural network[C]//2017 International Artificial Intelligence and Data Processing Symposium(IDAP). Malatya:[s.n.], 2017: 8090281.
- [17] PRAMERDORFER C, KAMPEL M. Facial expression recognition using convolutional neural networks: State of the art[J]. CoRR, 2016:1612.02903v1.
- [18] SZEGEDY C, LIU Wei, JIA Yangqing, et al. Going deeper with convolutions[C]//2015 IEEE Conference on Computer Vision and Pattern Recognition(CVPR). Boston:IEEE, 2015: 7298594.
- [19] SIMONYAN K, ZISSERMAN A. Very deep convolutional networks for large-scale image recognition[C]//3rd International Conference on Learning Representations. San Diego:ICLR, 2014:1409.1556V6.
- [20] 徐琳琳, 张树美, 赵俊莉. 构建并行卷积神经网络的表情识别算法[J]. 中国图象图形学报, 2019, 24(2): 227-236.
XU Linlin, ZHANG Shumei, ZHAO Junli. Expression recognition algorithm for parallel convolutional neural networks[J]. Journal of Image and Graphics, 2019, 24(2):227-236.