

江先辉, 徐名海, 王子轩. 基于 AMF-VGG16 的面部表情识别[J]. 智能计算机与应用, 2024, 14(7): 1-9. DOI: 10.20169/j.issn.2095-2163.240701

基于 AMF-VGG16 的面部表情识别

江先辉, 徐名海, 王子轩

(南京邮电大学 通信与信息工程学院, 南京 210003)

摘要: 随着深度学习的再度兴起, 情感识别领域得到了飞速的发展, 表情作为人类情感表现最直接的形式, 引起许多学者的关注。本文对面部表情识别 (Facial Expression Recognition, FER) 进行深入探讨, 以经典 VGG16 模型为切入点, 针对其存在的问题, 提出基于 AMF-VGG16 算法的面部表情识别模型。首先, 通过局部二值模式 (Local Binary Pattern, LBP) 及 Sobel 算子处理图像以强化特征作为模型输入; 其次, 针对 VGG16 模型存在的缺陷做出优化, 包括精简训练参数, 即将第四个卷积组的特征通道数量减半, 前两层全连接层中的神经元数量减少至 256, 以及深浅层特征融合完成信息互补; 最后, 引入空间及通道注意力机制进一步提升模型识别性能。在 CK+ 及 Fer2013 数据集上实验的准确率分别达到了 99.742 1% 和 66.78%, 较 VGG16 模型提升约 6% 和 4%, 较其它模型的性能也有一定的提升, 说明了本文提出模型的有效性。

关键词: 表情; VGG16; 特征融合; 注意力机制

中图分类号: TP391.41

文献标志码: A

文章编号: 2095-2163(2024)07-0001-09

Facial expression recognition based on AMF-VGG16

JIANG Xianhui, XU Minghai, WANG Zixuan

(College of Communication and Information Engineering, Nanjing University of Posts and Telecommunications, Nanjing 210003, China)

Abstract: With the resurgence of deep learning, the field of emotion recognition has experienced rapid development. As one of the most direct forms of human emotional expression, facial expressions have attracted the attention of many scholars. This article delves into Facial Expression Recognition (FER) and proposes a facial expression recognition model based on AMF-VGG16, taking the classic VGG16 model as the starting point to address its existing problems. Firstly, the image is processed using Local Binary Pattern (LBP) and Sobel operator to enhance features as model input. Secondly, optimization is made to address the shortcomings of the VGG16 model, including simplifying the training parameters. Specifically, the number of feature channels in the fourth convolutional group is halved, the number of neurons in the first two fully connected layers is reduced to 256, and deep and shallow feature fusion is completed. Information complementarity, and finally, the introduction of CBAM attention mechanism further enhances the recognition performance of the model. The accuracy of the experiment on the CK+ and Fer2013 datasets reached 99.742 1% and 66.78%, respectively, which is about 6% and 4% higher than the VGG16 model, and there is also a certain improvement in performance compared to other models, indicating the effectiveness of the proposed model in this paper.

Key words: expression; VGG16; feature fusion; attention mechanism

0 引言

识别人的情绪状态, 有广泛的应用场景, 如: 安全驾驶的自动检测, 通过车机设备读取驾驶员的情绪状态相关信息, 分析驾驶员的情绪状态, 若为消极状态则发出警报; 辅助治疗, 通过医疗设备监控患者

的情绪状态, 以便给病患更好的医疗体验; 多媒体推荐, 通过识别用户的情绪状态, 推荐播放能调节用户情绪的多媒体^[1]。早在二十世纪七十年代, 科学家赫伯特·西蒙就曾在一次认知心理学的讲座中强调“解决问题要结合情感”, 说明分析人的情绪状态对解决问题有线索性意义。有心理学研究表明, 在人

基金项目: 国家自然科学基金青年科学项目 (60902015)。

作者简介: 江先辉 (1999-), 男, 硕士研究生, 主要研究方向: 机器学习, 表情识别, 情感识别; 王子轩 (1999-), 女, 硕士研究生, 主要研究方向: 复杂网络建模, 机器学习。

通讯作者: 徐名海 (1976-), 男, 博士, 副教授, 主要研究方向: 无线异构网络融合, 环境感知, 智能交通。Email: d0207@njupt.edu.cn

收稿日期: 2023-05-18

们的日常交际中,理解他人情感的各方面因素,面部表情的占比达到了55%,而语言信息传达的只有7%,剩余的38%则是通过音色、音调等声音信息传达^[2]。可见面部表情对安全检测、疾病治疗、娱乐措施等有较大的意义,因而表情识别也成为了研究和应用领域的重大课题。

传统的面部表情识别方法实现简单,但识别效果受环境和个体差异影响、存在不适应面部表情变化的快速性和瞬时性等缺点。随着深度学习的再度兴起,各种卷积神经网络(Convolutional Neural Network, CNN)在面部表情识别领域比传统方法有更高的识别准确率和更强的适应性,但也存在一些问题,如算法复杂、需要大量运算资源、对数据集质量要求高,出现过拟合问题等。

视觉几何组(Visual Geometry Group, VGG)网络是最经典的卷积神经网络之一,其用多个小卷积核替代CNN的大卷积核,保证相同的感受野同时大大加深了网络深度,另外整个模型采用固定3×3大小的卷积核,因此较其它模型简洁。本文以VGG16为基础对面部表情识别进行研究,构建基于CBAM注意力机制的多层特征融合改进VGG16模型,简称AMF-VGG16。

1 国内外研究进展

面部表情识别主要包括3个部分:图像获取与预处理、特征提取、特征分类。图像的预处理包括人脸的检测与定位、几何归一化、直方图均衡化数据增强等。

传统的特征提取主要包括基于几何和基于外观纹理的手工提取方式^[3]。如P. Chenniappan等^[4]基于几何特征生成特征图;王涛等^[5]将局部二值模式(Local Binary Pattern, LBP)特征与几何融合;Islam B等^[6]切割人脸区域并融合方向梯度直方图(Histogram Of Gradient, HOG)特征和LBP特征。然而这些提取特征的方法受手工特征设计的影响,而手工特征又具有不确定性,易丢失关键的面部表情特征,此外易受复杂的背景、阳光照射、遮挡等情况的影响,导致识别率降低,因此泛化能力不强。

2000年以来,以CNN为核心的深度学习被广泛应用于面部表情特征提取。CNN模拟人类的大脑神经元,显著提升了提取图片特征的能力,在面部表情识别任务中表现出色,为面部表情识别的研究起到了决定性的作用^[7-8]。研究人员提出了许多经典的CNN模型,例如Krizhevsky A^[9]2012年提出的

深层卷积神经网络 AlexNet; Simonyan K^[10]2014年提出的VGGNet; He K等^[11]提出的ResNet。这些模型仍存在一些问题,因此许多学者在这些模型的基础上提出了改进的模型,如石翠萍等^[12]提出改进AlexNet模型,采用单图形处理单元进行训练,减少了两层卷积层和一层全连接层,在每层卷积层后加上批标准化代替原来的局部归一化,并在全连接层后加上Dropout正则化,进一步防止过拟合,比原模型更简单,参数量少;兰江海等^[13]以ResNet18作为基本网络,改进得到一个新的表情识别模型,增强了网络的鲁棒性,提升了模型的识别率。

除此之外,也有学者将手工特征和CNN相结合,完美融合了两者的优点。例如Zhang Q等^[14]将LBP特征与VGG网络相结合,提高了识别的准确率;Liu S等^[15]对数据集进行Sobel算子的边缘检测,送入CNN中进行特征提取,取得了不错的效果;Zhang L等^[16]提出改进的LBP算子并集成多个CNN网络完成特征的融合。

对提取到的特征进行分类是很重要的环节,算法的优劣直接影响到表情识别的效果,常见的一些分类方法及其优缺点和适用场景见表1。

2 应用场景

面部表情识别技术有很多应用场景,如识别用户的面部表情,推荐能调节用户情绪的多媒体;通过商场的摄像头获取用户在商品前的情绪状态,以便更好的推荐商品;通过车机设备获取驾驶员的面部表情,分析得到驾驶员的情绪状态,若为消极状态则发出警报等等。个人的音乐推荐系统示例如图1所示,通过电子产品或者车机上的摄像头获取用户的面部图片,进行面部表情识别得到用户的情绪状态,当用户使用音乐指令后根据识别出的用户情绪状态播放对应情感属性的音乐,音乐反过来调节用户的情绪。

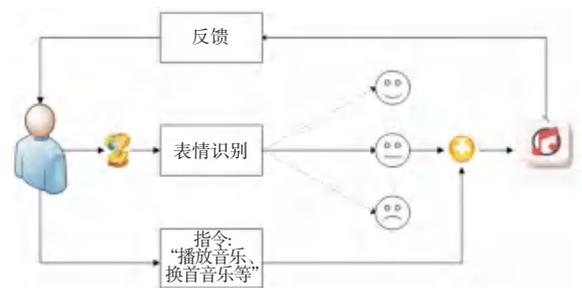


图1 个人音乐推荐

Fig.1 Personal music recommendation

表 1 常见的分类算法

Table 1 Common classification algorithms

分类算法	优点	缺点	适用场景
朴素贝叶斯	简单有效、易实现和理解等	对先验概率敏感、类别特征值较多则分类效果差等	文本分类、天气预报、商品推荐等
决策树	简单易理解、结构可视化等	易过拟合、类别过多时错误增加的会比较快等	情感分析、产品选择等
K 最近邻算法	简单易懂、易实现、训练时间短等	特征数多时计算量大、预测速度相对较慢等	垃圾邮件识别、图像内容识别、检测异常值等
支持向量机	高维空间下有良好的泛化能力、可调整核函数适应特定问题等	运行和调参麻烦、对大规模数据处理效率极低等	图像分类、文本分类、表情识别等
Adaboost 算法	很好的综合多个弱分类器, 获得更好的性能、不易过拟合等	对弱分类器质量和数量要求高、训练时间长等	图像识别、人脸识别、情感识别等
随机森林	引入随机性, 不易过拟合、适合大数据集、容易实现并行化	实施复杂、黑盒模型不易解释等	股票市场预测、信用卡违约等
深度学习算法	很强分类能力、可自动提取特征、能处理大规模高纬度数据	对硬件资源、算力等要求高、优化周期和训练时间长等	计算机视觉、语音识别、自然语言处理等

3 基于 AMF-VGG16 的面部表情识别

3.1 特征分析

直接将表情数据集送入 CNN 中训练就能得到不错的分类效果。为了进一步提高识别效果, 可以先对图像做传统的特征提取。面部的纹理、轮廓等特征对表情识别有不可忽略的作用, 因此, 本文考虑对图片做 LBP 和 Sobel 预处理。

LBP 最大的作用就是提取图像的纹理特征, 有计算复杂度低、光照不变形、易实现等优点。最初的 LBP 算子提取特征有着严重的缺陷, 如提取的特征过于局部、无法捕获大尺度的纹理特征、不宜旋转不变特征的设计等, 后来出现了大量 LBP 改进方法, 如圆形 LBP、旋转不变 LBP、均匀 LBP 等。本文使用 Zhang L 等^[16]提出的改进 LBP 算法对原图像处理, 原理图如图 2 所示。



图 2 改进的 LBP 原理图

Fig.2 Improved LBP schematic diagram

改进的 LBP 由原先的单个像素扩展为 $m \times m$ 大小的像素块, 原先的中心像素点扩大为包含中心像素点及其领域的 $m \times m$ 大小的中心像素块, 用于扩大对比度范围, 强化面部整体的纹理特征。当 m

为奇数时取中间点作为中心像素, m 为偶数时取像素块中间最小的 2×2 正方形的右上角像素作为中心像素。

g_p 是边缘像素块的中心像素值, 等于对应像素块的像素平均值, 公式(1):

$$g_p = Ave(p) \tag{1}$$

g_c 是中间像素块的中心像素值, 公式(2):

$$g_c = Ave(c) \tag{2}$$

LBP_c 是经改进的 LBP 处理后的中间像素块的中心像素值, 公式(3):

$$LBP_c = \sum_0^{p-1} s(g_p - g_c) \times 2^p, p \in (0, 1, \dots, p - 1) \tag{3}$$

$$s(x) = \begin{cases} 1, & x \geq 0 \\ 0, & x < 0 \end{cases} \tag{4}$$

Sobel 算子是一种一阶微分算子, 用于边缘检测。Sobel 根据图像的灰度梯度大小来确定边缘位置, 检测到的边缘比较清晰, 不容易产生噪声。对于 FER, 脸部的肌肉变化是区分表情直接有效的特征, 因而对表情图片做 Sobel 算子处理, 强化不同表情的轮廓边界, 有助于特征的提取, 包括水平和垂直两个方向的检测算子, 公式(5)和公式(6):

$$G_x = \begin{matrix} \hat{e}-1 & 0 & +1\hat{u} \\ \hat{e}-2 & 0 & +2\hat{u} \times A \\ \hat{e}-1 & 0 & +1\hat{u} \end{matrix} \tag{5}$$

$$G_y = \begin{matrix} \hat{e}+1 & +2 & +1\hat{u} \\ \hat{e} & 0 & 0\hat{u} \times A \\ \hat{e}-1 & -2 & -1\hat{u} \end{matrix} \tag{6}$$

其中, A 为待处理的图像。

G 为最终的灰度值,公式(7):

$$G = \sqrt{G_x^2 + G_y^2} \quad (7)$$

经 LBP 及 Sobel 算子处理后图像如图 3 所示。



图 3 原始图像以及 LBP 和 Sobel 特征图像

Fig. 3 Original image as well as LBP and Sobel feature images

3.2 VGG16 模型优化分析

传统的情感分类模型易理解和实现,但较目前主流的深度学习模型识别准确率低,适合处理大量数据的深度学习模型目前更受欢迎,其中的 CNN 在 FER 任务中表现优异,如常见的 VGG16、GoogLeNet 以及 ResNet,对比见表 2。

VGG16 提出时间相对早,易实现、易调整、易进行特征可视化和解释,且在图像识别任务上有出彩的表现,本文基于 VGG16 模型对 FER 进行研究, VGG16 网络结构如图 4 所示。

表 2 常见 CNN 模型对比

Table 2 Comparison of common CNN models

CNN 模型	优点	缺点
VGG16	结构简单、易实现和调整、易于进行特征可视化和解释、不易出现梯度消失、图像识别中具有较高的准确性	参数较多,训练时间较长、内存消耗大、不适合处理大尺寸图片
GoogLeNet	模型轻量化、用 Inception 模块进行卷积操作,计算效率高、灵活易调整和扩展	网络参数较难调整、特征可解释性较差
ResNet	残差块缓解梯度消失问题,加快了模型学习速度、以训练更深的神经网络,提高精度	结构复杂、模型训练需要更多数据、不易解释

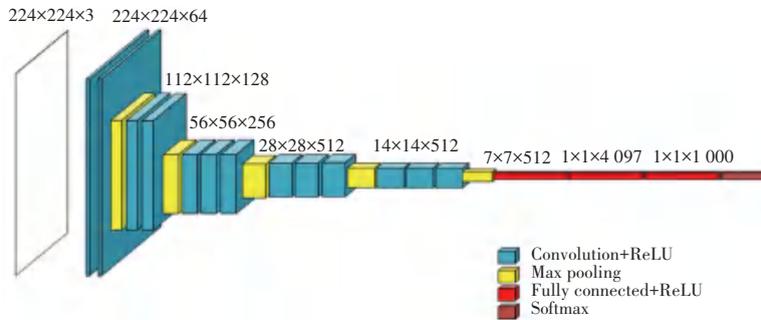


图 4 VGG16 网络结构

Fig. 4 VGG16 network structure

3.2.1 参数优化分析

VGG16 网络结构非常简单,每个卷积层采用固定的 3×3 的卷积核,并在每个卷积层后都有 ReLU 激活函数进行非线性变换,特征图最大池化采用 2×2 池化核,并且每个卷积组之间都会有最大池化层以减小特征图大小,在计算机视觉领域有着很好的表现。VGG16 最大的缺陷之一就是训练参数特别多,导致训练速度也非常慢,而训练的主要参数在卷积神经网络的卷积层及全连接层。卷积层训练参数计算见公式(8):

$$P_x = f_x \times (f_{x-1} \times 3 \times 3) \quad (8)$$

其中, P_x 是本层所需训练的参数数量; f_x 是本层设置的特征图通道大小; f_{x-1} 是上一层的特征图通道大小。

卷积神经网络的全连接层训练参数计算见公式(9):

$$P_x = c_x \times o_{x-1} \quad (9)$$

其中, P_x 是本层所需训练的参数数量, c_x 为本连接层神经元个数, o_{x-1} 为上一层输出的总尺寸大小,如上一层为卷积层或者池化层,则 o_{x-1} 为特征图通道乘输出特征图尺寸;如上一层为全连接层,则 o_{x-1} 为神经元个数。

VGG16 模型中的多个卷积层将特征图通道数翻倍,并且全连接层的神经元数量非常庞大,导致需要训练的参数数量非常巨大,但实际上这些参数中很多都是冗余的,例如卷积层设置的特征图通道数即输出的特征图通道数,在每个特征通道中提取输入图像的某些特定特征,也就是说,一些设置的通

道无法提取有效的特征。

卷积层的特征图通道数量以及全连接层的神经元数量的增加会导致训练参数急剧增加。因此, 本文针对 VGG16 的训练参数问题结合 VGG16 本身的网络结构, 对模型进行优化, 将第四个卷积组的特征图通道数量减半, 前两层全连接层中的神经元数量减少至 256。

3.2.2 深浅层特征优化分析

对于 VGG16 来说, 网络层越深, 提取的特征就越高级、越抽象, 可能会丢失很多信息。深层神经网络所提取的特征离输出更接近, 包含更抽象的高级语义信息, 然而其分辨率很低, 对细节的感知能力较差。相比之下, 浅层神经网络提取的特征与输入数据集相似, 包含更多的像素和细粒度信息, 如图像的纹理、边缘和棱角等一些细节信息。浅层神经网络的感受野相对较小且感受野重叠区域也较小, 因此可以保证网络提取的特征具有更高的分辨率和更多的位置和细节信息。但浅层神经网络的卷积层较少, 其所提取的特征的语义性较低, 噪声也比较多。通过特征图的可视化可以得到不同网络层提取到的特征, VGG16 模型含有多层网络层且每层提取的特征图数量巨大, 其中第二层和第八层卷积层提取到的原图特征如图 5 所示。

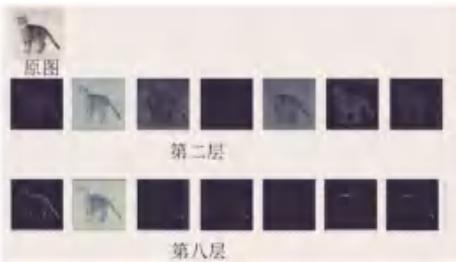


图 5 VGG16 第二层及第八层卷积层提取到的原图特征

Fig. 5 Original image features extracted from the second and eighth convolutional layers of VGG16

VGG16 模型含有多层网络层且每层提取的特征图数量巨大, 因而挑选部分进行展示说明。从图 5 能看出深层次网络提取的特征是更抽象的信息, 对细节的感知能力较差。而浅层网络分辨率更高, 包含更多位置、纹理、边缘、棱角、轮廓、细节信息。因此, 本文针对深层特征的缺陷, 将深浅层提取的特征融合, 弥补深层网络特征的不足。

特征融合方法有两种实现方式: Concat 和 Add。Concat 将特征图进行连接, 即增加特征图的数量, 扩展了输入特征数; 而 Add 则是对特征图中对应位置的像素进行相加, 增加单个特征图上的信息量, 但特

征图的数量不变。

VGG16 包含 5 个卷积组, 每个卷积组后有一个步长为 2 的最大池化层, 每个卷积组中的卷积层都有完全相同的参数。基于 VGG16 网络结构, 文本将 5 个卷积组的最后 5 个池化层的输出特征图进行融合。由于各层维度都不符合这两种方法的输入要求, 为此, 借鉴 Inception 模块中的思想, 即将池化层的输出送入 1×1 卷积层, 进一步进行特征抽取以改变特征图数; 送入池化层中, 改变特征图的尺寸。通过这种方式, 满足了特征融合的输入要求, 特征融合的具体结构如图 6 所示。

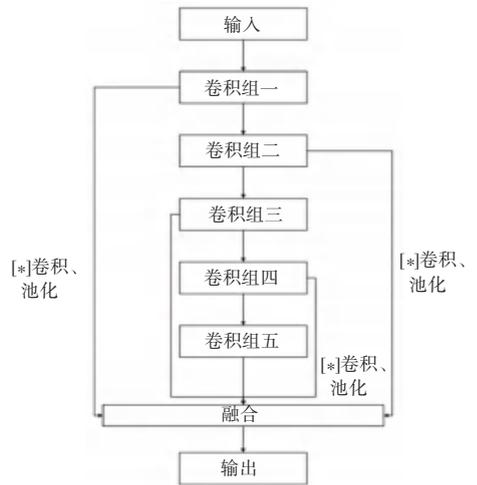


图 6 多层特征融合示意图

Fig. 6 Schematic diagram of multi-layer feature fusion

3.3 AMF-VGG16 算法模型

注意力机制 (Attention Mechanism) 类似于大脑的注意力机制, 能够在一张图片中发现焦点。神经网络中的 Attention 机制是在计算能力有限的情况下, 将计算资源分配给更重要的任务, 同时解决信息超载问题的一种资源分配方案。通过这种机制可以利用有限的资源从大量信息中快速筛选出高价值的信息, 常用于提升模型的整体性能。

卷积块注意力模块 (Convolutional Block Attention Module, CBAM) 适用于计算机视觉任务, 能够根据输入数据自适应调整权重并且利用激活函数引入了非线性关系, 可以更好地捕捉表情的细微变化, 同时综合考虑了通道和空间特征, 因而, 可以很好地与 VGG16 的多通道输出契合, 提高识别的准确性和鲁棒性, 因此本文引入 CBAM 模块, 其结构如图 7 所示。

本文构建基于 CBAM 注意力机制的多层特征融合改进 VGG16 模型, 简称 AMF-VGG16 模型, 模型结构如图 8 所示。

常用的评价指标包括准确率 (*Accuracy*)、精确率 (*Precision*)、召回率 (*Recall*) 和 *F1* 分数等。准确率是正确分类的样本数占总样本数的比例,是最简单直观的评价指标,但是当样本分布非常不均匀时,效果就非常差,由于本数据集采用了数据增强防止数据倾斜,各类数据集分布比较均衡,因此,采用最为直观的准确率作为评价指标。

4.2 数据集及处理

本文采用的数据集为 CK+ 及 Fer2013。CK+ 是目前常用的表情数据集,对 CK 数据进行了纠正和扩展,包含丰富的标注信息,本文选取其图片序列中的除中性外的 7 种峰值表情,作为实验的数据集。图片包含许多和面部表情无关的背景成分,直接输入进模型会造成资源的浪费,增加训练时间,对表情识别的效果造成干扰,降低识别准确率,所以对数据集人脸感兴趣区域进行识别并完成裁剪,去除掉不相关的部分。此外,使用 Keras 框架中封装的 *ImageDataGenerator* 函数,对样本图片做诸如翻转、

平移、旋转之类的数据增强操作,处理后的数据集如图 9 所示。Fer2013 仅做数据增强处理。实验将数据集按 8 : 1 : 1 的比例分为训练集、验证集、测试集。



图 9 处理后的 CK+数据集
Fig. 9 Processed CK+dataset

4.3 实验结果及分析

本文对提出的 AMF-VGG16 模型,在 CK+数据集上完成了两种特征融合方式即 Add 和 Concat 的对比实验,并对优化模型的性能进行验证。原 VGG16 模型及本文提出的 AMF-VGG16 模型训练时的训练集和验证集的准确率及损失曲线如图 10 所示。

此外,VGG16 及 AMF-VGG16 模型在 CK+测试集上得到准确率混淆矩阵如图 11 所示。

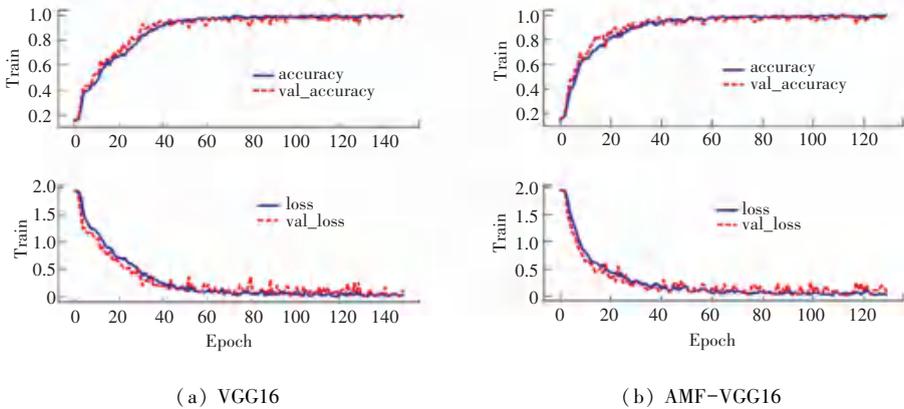


图 10 VGG16 和 AMF-VGG16 模型在 CK+数据集上的训练准确率及损失曲线

Fig. 10 Training accuracy and loss curves of VGG16 and AMF-VGG16 models on the CK+dataset

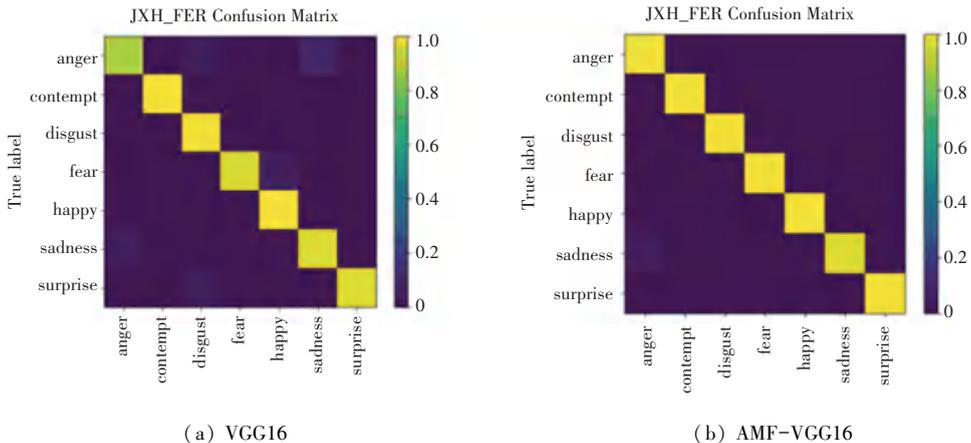


图 11 VGG16 和 AMF-VGG16 模型在 CK+数据集上的测试集混淆矩阵

Fig. 11 Test set confusion matrix of VGG16 and AMF-VGG16 models on CK+dataset

图 11 的混淆矩阵中,在右边数值相同的情况下,颜色越深代表识别准确率越高,可见 AMF-VGG16 模型弥补了原 VGG16 模型对生气、害怕、惊讶表情识别的缺陷,提高了识别的准确率。

陈加敏等^[17]提出改进的残差网络模型 APRNET50,利用金字塔卷积提取图像的细节信息,并通过引入通道和空间注意力提升模型识别性能。对比 VGG16 模型、AMF-VGG16 模型及 APRNET50 模型,在 CK+数据集上实验结果见表 4。

表 4 CK+数据集上不同模型实验结果对比

Table 4 Comparison of experimental results of different models on the CK+dataset

模型	准确率/%	参数量/m
VGG16	93.710 0	33
优化模型(Add)	97.790 0	11
优化模型(Concat)	97.146 4	12
AMF-VGG16(Add)	95.283 0	11
AMF-VGG16(Concat)	99.742 1	12
APRNET50	94.949 0	

从表 4 中可以看出,AMF-VGG16 模型较 VGG16 模型识别准确率提升了约 6 个百分点,较 APRNET 提升了近 5 个百分点。针对模型参数的优化使得参数量减少了近三分之二,AMF-VGG16 模型的识别准确率较 VGG16 模型有不小的提升,说明对输入图像的强化特征处理以及网络深浅层特征的融合使得模型抽取的特征更加有效且全面,CBAM 对特征图的聚焦使得分类准确度进一步提高。

另外,郭昕刚等^[18]提出的改进的 VGG 模型,在 VGG 网络基础上加入改进型高阶残差模块和参数共享反馈子网络降低网络退化度和减少网络参数量,并利用通道注意力机制给通道分配不同权重,来提高表情识别度,整体训练参数约 9 M,略低于本文的 AMF-VGG16 模型。AMF-VGG16 模型、VGG16 模型及改进的 VGG 模型^[18]在 Fer2013 数据集上对比实验结果如图 12 所示。

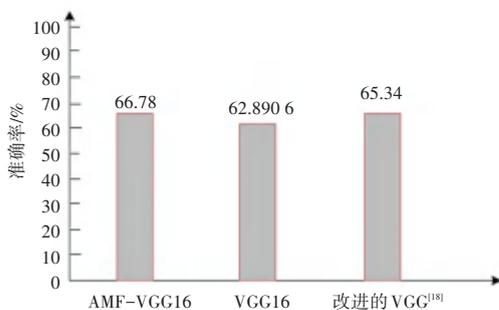


图 12 Fer2013 数据集上与其它模型对比

Fig. 12 Comparison with other models on the Fer2013 dataset

由上图 12 可知 AMF-VGG16 模型较 VGG16 模型及改进的 VGG 模型在 Fer2013 数据集上的识别率更高,证明了本模型的有效性。

5 结束语

本文对面部表情识别进行研究,针对 VGG16 模型的缺点,提出 AMF-VGG16 模型。使用了多种传统的特征提取方式强化表情纹理、轮廓特征,在维持识别效果几乎不变的情况下优化减少了模型参数,并且分析深浅层特征的优缺点,完成深浅层特征之间的融合互补,最后使用通道和空间注意力机制进一步提升模型的识别性能和泛化能力。在 CK+和 Fer2013 数据集上的实验结果表明此模型在面部表情识别上的有效性。本文提出的模型也存在一些不足,如参数的优化缺乏解释性,减少参数后没有降低多少模型的性能,但参数的设置何时为最优却无法解释;验证了多层特征融合能提高识别的准确率,但没有确定具体哪层是最有效的,缺乏对比实验,这些问题都有待进一步的研究。

参考文献

- [1] DAMIA A H, ESNAASHARI M M. Automated test data generation using a combination of firefly algorithm and asexual reproduction optimization algorithm[J]. International Journal of Web Research, 2020, 3(1): 19-28.
- [2] CHEN N, KIM S. Puzzle-based automatic testing: bringing humans into the loop by solving puzzles [C]//Proceedings of IEEE/ACM International Conference on Automated Software Engineering. IEEE, 2012:140.
- [3] 王浩,栗永泽,方宝富.基于局部特征聚类损失和多类特征融合的面部表情识别[J].模式识别与人工智能,2020,33(3):9.
- [4] CHENNIAPPAN P, RAJAN S, MADIAN N, et al. Geometric based feature extraction techniques for facial expression recognition using SVM [C]// Proceedings of 2021 Smart Technologies, Communication and Robotics (STCR). IEEE, 2021: 1-6.
- [5] 王涛,彭欣荣,朱经睿,等.基于几何特征和LBP特征融合的笑脸识别算法的研究[J].电子测试,2021(23):52-54.
- [6] ISLAM B, MAHMUD F, HOSSAIN A. High performance facial expression recognition system using facial region segmentation, fusion of HOG & LBP features and multiclass SVM [C]// Proceedings of 2018 10th International Conference on Electrical and Computer Engineering (ICECE). IEEE, 2018: 42-45.
- [7] HINTON G E, OSINDERO S, TEH Y W. A fast learning algorithm for deep belief nets[J]. Neural Computation, 2006, 18(7): 1527-1554.
- [8] 景晨凯,宋涛,庄雷,等.基于深度卷积神经网络的人脸识别技术综述[J].计算机应用与软件,2018,35(1):9.
- [9] KRIZHEVSKY A, SUTSKEVER I, HINTON G E. ImageNet classification with deep convolutional neural networks [J]. Communications of the ACM, 2017, 60(6): 84-90.
- [10] SIMONYAN K, ZISSERMAN A. Very deep convolutional

- networks for large-scale image recognition [J]. arXiv preprint arXiv:1409.1556, 2014.
- [11] HE K, ZHANG X, REN S, et al. Deep residual learning for image recognition [C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. IEEE, 2016; 770-778.
- [12] 石翠萍, 谭聪, 左江, 等. 基于改进 AlexNet 卷积神经网络的人脸表情识别[J]. 电讯技术, 2020, 60(9): 1005-1012.
- [13] 兰江海, 林国军, 游松, 等. 基于改进的残差网络面部表情识别[J]. 国外电子测量技术, 2023, 42(3): 123-130.
- [14] ZHANG Q. Facial expression recognition in VGG network based on LBP feature extraction [C]// Proceedings of 2020 5th International Conference on Mechanical, Control and Computer Engineering (ICMCCE). IEEE, 2020; 2089-2092.
- [15] LIU S, TANG X, WANG D. Facial expression recognition based on sobel operator and improved CNN-SVM [C]// Proceedings of 2020 IEEE 3rd International Conference on Information Communication and Signal Processing (ICICSP). IEEE, 2020; 236-240.
- [16] ZHANG L, XU C, LI S. Facial expression recognition of infants based on multi-stream CNN fusion network [C]// Proceedings of 2020 IEEE 5th International Conference on Signal and Image Processing (ICSIP). IEEE, 2020; 37-41.
- [17] 陈加敏, 徐杨. 注意力金字塔卷积残差网络的表情识别[J]. 计算机工程与应用, 2022, 58(22): 9.
- [18] 郭昕刚, 沈紫琪, 王晓林. 改进 VGG 网络的人脸表情识别[J]. 长春工业大学学报, 2023, 44(1): 52-57.