

基于 CBAM-DSC 网络的表情识别方法

宋文博¹, 高璐², 苗壮¹, 林克正¹

(1. 哈尔滨理工大学 计算机科学与技术学院, 哈尔滨 150080; 2. 哈尔滨信息工程学院, 哈尔滨 150431)

摘要:针对人脸表情网络模型参数复杂和计算性能偏低等问题,提出一种基于卷积注意力机制和深度可分离卷积(convolutional block attention module-depthwise separable convolution, CBAM-DSC)网络的表情识别方法。网络使用深度可分离卷积与传统卷积相结合,提出的改进的 Inception 模块通过不同分支提取不同特征信息的同时减少了网络参数量,提高网络运行的效率。最后添加了卷积注意力机制模块,能够使网络提取特征时重点关注关键信息,从而使得提取到的特征表达信息更准确,更利于分类。在 RAF-DB 数据集和 CK+ 数据集上的仿真实验表明,网络模型 CBAM-DSC 具有较高的识别率,消融实验中相比传统 CNN(convolutional neural network)网络的参数量减少了 6.4%,提升了计算性能。

关键词:Inception; 深度可分离卷积; 注意力机制; 表情识别

DOI:10.15938/j.jhust.2023.06.012

中图分类号: TP391.4 **文献标志码:** A **文章编号:** 1007-2683(2023)06-0103-08

Expression Recognition Method Based on CBAM-DSC Network

SONG Wenbo¹, GAO Lu², MIAO Zhuang¹, LIN Kezheng¹

(1. School of Computer Science and Technology, Harbin University of Science and Technology, Harbin 150080, China;

2. Harbin Institute of Information Technology, Harbin 150431, China)

Abstract: Aiming at the problems of complex parameters and low computational performance of facial expression network model, an expression recognition method based on Convolutional Block Attention Module-Depthwise Separable Convolution network is proposed. The network usage depth separable convolution is combined with the traditional convolution. The improved Inception module extracts different feature information through different branches while reducing the network parameters and improving the network operation efficiency. Finally, the convolutional block attention module is added, which can make the network focus on the key information when extracting features, so that the extracted feature expression information is more accurate and more conducive to classification. The simulation experiments on RAF-DB dataset and CK+ dataset show that the network model Convolutional Block Attention Module-Depthwise Separable Convolution has a high recognition rate, and the number of parameters in the ablation experiment is reduced by 6.4% compared with the traditional Convolutional Neural Network, which improves the computing performance.

Keywords: Inception; deep separable convolution; attention mechanism; expression recognition

0 引言

随着卷积神经网络在计算机视觉研究中的优异

表现,越来越多人开始使用深度学习的技术完成表情分类任务,也使用了一些经典的卷积神经网络提取特征,如 VGGNet(visual geometry group network)、GoogLeNet 等。但由于人脸表情数据受环境特殊、

收稿日期: 2022-05-18

基金项目: 国家自然科学基金(62071157); 黑龙江自然科学基金(F2015040)。

作者简介: 宋文博(1998—),男,硕士研究生;

苗壮(1996—),男,硕士研究生。

通信作者: 高璐(1982—),女,硕士,副教授, E-mail: hxeigaolu@126.com。

人脸姿态不同和背景光照条件等因素的影响,导致有些模型在表情识别任务中效果并不理想。

近年来随着人脸表情识别领域的深入研究,提出了大量模型。文元美等^[1]提出利用 ROI (regions of interest) 区域二级投票机制识别人脸表情,提高模型在复杂环境下的应对表情识别问题的能力,但是识别率提升幅度较小,整体识别率并不突出。Jung 等^[2]提出的深度网络基于两种不同的模型,但只适用于部分小样本数据集。Kim 等^[3]使用不同学习策略训练多个深度卷积神经网络 (convlutional neural network, CNN) CNN 完成识别任务,但是该网络结构和连接方式过于复杂导致训练过程繁琐耗时。张鹏等^[4]提出的网络模型引入注意力机制和空洞卷积来提取人脸表情的多尺度特征信息,但空洞卷积对相似表情识别效果不好。程学军等^[5]提出的改进型 VGG-16 网络解决了像素特征分布不均的问题,在黑暗条件下也有 90% 的准确率。李春虹等^[6]利用人脸分割网络分割出人脸图像中与表情识别相关的区域,利用深度可分离卷积构建两个基分类器,减少了非重要区域的影响,取得了很好的识别率。魏赞等^[7]提出一种注意力机制的轻量级 CNN 通道和卷积自编码器预训练通道的双通道模型,卷积自编码器对输入人脸表情进行无监督预处理,有效克服了单一神经网络模型提取表情特征困难,以及堆叠网络结构造成训练复杂的问题。亢洁等^[8]提出利用迁移学习在目标函数中构造一个基于 log-Euclidean 距离的损失项来减少迁移学习中源域与目标域之间的相关性差异,并结合注意力机制进行人脸表情识别。韦赛远等^[9]对戴口罩的人群进行人脸表情识别,利用 Dlib 的 Face-Mask 技术对 FER2013 数据集进行遮挡处理生成 FERM 数据集进行实验,取得不错的实验结果。

随着网络规模变得越来越大,导致参数数量的大幅度增加,网络结构不够轻量,出现计算性能不高、识别效率低等问题。因此,提出一种基于 CBAM-DSC 网络的表情识别方法,以 Inception 模块为基础,使用了深度可分离卷积 (depthwise separable convolution, DSC),将其和传统卷积相结合使用,并添加了卷积注意力机制 (convolutional block attention module, CBAM) 模块构造出一种新型的网络结构 CBAM-DSC 网络,其中深度可分离卷积分步提取通道信息和空间信息。卷积注意力机制模块添加在卷

积层后使网络关注有效信息^[10],最后使用 RAF-DB 和 CK+ 中的数据样本进行大量实验,实验结果表明, CBAM-DSC 网络具有较好的识别效果。

1 相关基本理论

1.1 注意力机制

卷积注意力机制模块^[11]由通道注意力^[12]和空间注意力两部分组成,卷积层的输出结果,会先通过一个通道注意力模块,得到加权结果之后,会再经过一个空间注意力模块,最终进行加权得到结果。将两种方式融合在一起能够进行多方位卷积^[13-14],提取到更重要的信息,增强了网络提取有效信息的能力。图 1 为 CBAM 原理图。

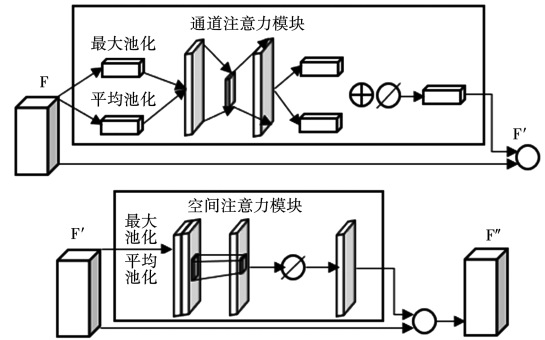


图 1 CBAM 的原理图

Fig. 1 Schematic diagram of CBAM

CBAM 主要是对权重重新赋值将特征图中的关键特征信息进行标识,使后续的特征提取重点关注关键区域,增强提取特征的有效性。CBAM 模块的输入是从卷积层得到的矩阵 $F_1 \in \mathbf{R}^{C \times H \times W}$ 。先对矩阵在通道维度上进行操作,分别是最大池化和平

均池化操作,将提取到的两种特征压缩合并在一起,合并后的特征图经过单个卷积核的隐藏层生成通道权值矩阵 $F_1 \in \mathbf{R}^{C \times H \times W}$,如式(1)所示。

$$F_1 = M_s(F) \odot F \quad (1)$$

空间注意力机制是对通道注意力机制的结果进行后续操作,是在空间维度上进行操作,再一次进行最大池化和平均池化操作,与通道注意力不同的是,空间注意力将特征图叠加压缩为一个信道描述符,经过空间压缩操作的计算得到空间权值矩阵 F_2 ,如式(2)所示。经过操作后可通过矩阵建立像素点之间的重要关系,突出重要的信息区域。

$$F_2 = M_s(F_1) \odot F_1 \quad (2)$$

其中: $M_s(F)$ 表示三维的通道压缩权值矩阵; $M_s(F_1)$ 表示空间压缩权值矩阵; “ $\odot$ ” 表示矩阵间乘操作。

1.2 深度可分离卷积

与传统神经网络相比,卷积神经网络优势是网络中的权值可以共享,在图像分类、图像处理等方面非常实用,可以得到较好的效果。卷积核可以完成在特征图上的卷积运算,卷积核具有3个维度,分别是通道维度和空间维度上的宽和高。卷积操作映射出的结果是对深度和空间相关性的联合表达,卷积神经网络可以自动完成对图像数据的特征提取,经过研究者们不断地研究,对其进行优化和改进,演生出许多类型的卷积方式,它们都有不同于其他卷积的特点和优势,解决遇到的各类问题。深度可分离卷积的优势和特点就在于对网络参数数量的减少。

深度可分离卷积^[15]的思想是分开提取通道维和空间维上的信息,这样相较于原本的共同卷积使用的参数量更少,有效减少了计算量,同时这样多次操作得到的结果更好,具体改变。将传统卷积分成两部分,分别是 1×1 卷积和深度卷积,如图2和图3所示。

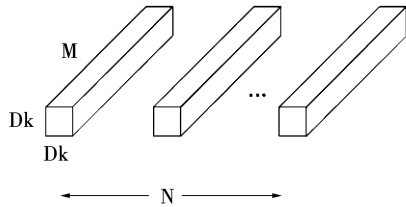


图2 传统卷积图

Fig.2 Traditional convolution graph

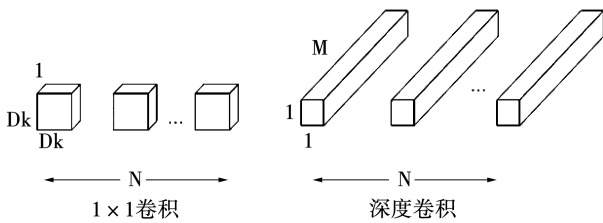


图3 可分离卷积图

Fig.3 Separable convolution graph

2 基于 CBAM-DSC 网络的表情识别方法

2.1 特征提取模块

特征提取部分将深度可分离卷积与传统卷积结合使用,网络前端依旧使用传统卷积操作,网络主体部分使用 Inception 网络结构。对图像进行多条分

支处理后汇合到一起传输到下一层,使得网络学习到的图像信息更丰富。同时在网络模型的首尾处卷积层后添加卷积注意力机制模块,其中先进行通道注意力后进行空间注意力,增强有效特征抑制无效的特征。

使用的 Inception 网络结构是对原有 Inception v1 结构的改进。其中 Inception v1 模块里有4条并行的通道,通道1和2先使用 1×1 卷积核改变特征图的通道数,这样可以降低网络复杂度,然后分别使用大小为 3×3 、 5×5 的卷积核对待特征图进行提取同时改变通道数,通道3使用最大池化和卷积核一起调整通道数,通道4单独使用小卷积核改变特征图的通道数,在提取特征过程中都使用合适的填充,使得4条通道的特征图尺寸相同,最后将提取到的多个特征图在通道维度上进行拼接后作为后面网络层的输入。Inception 网络结构如图4所示。

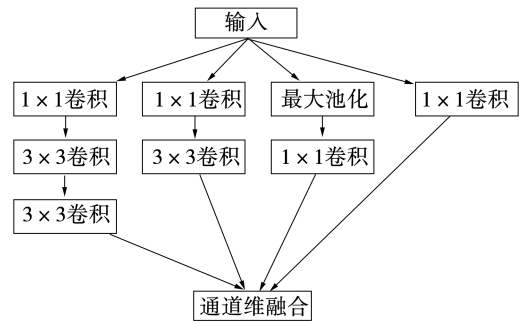


图4 Inception 网络结构

Fig.4 Inception network structure

将通道1的 5×5 卷积层替换为两个 3×3 卷积层,保留了原有卷积的提取效果并且降低了参数量,深度的增加提高了非线性特征的表达能力,在卷积层后加入 BN 层。

2.2 CBAM-DSC 网络结构

CBAM-DSC 模型的网络结构,在浅层特征使用传统卷积进行特征提取。两个 Inception 模块在其中添加了深度可分离卷积,然后在首尾处各添加一个 CBAM 模块。其中,深度可分离卷积在 Inception 模块中的通道1中两次,通道2中使用一次,最后将 Inception 模块的4个通道的结果融合后是一个 $6 \times 6 \times 512$ 的特征图,在经过注意力机制增强关键信息的提取,进行池化操作调整尺寸,用全局平均池化代替全连接层实现降维生成特征向量,减少了参数量避免出现过拟合,最后将特征向量传入到 SoftMax 进行类别的预测,输入大小为 $48 \times 48 \times 1$,其中网络结构中的参数配置,以及输出的大小如表1所示。

表1 CBAM-DSC 网络配置

Tab.1 Configuration of the CBAM-DSC

主网络	分支网络	卷积核	输出大小
Conv-1	—	$3 \times 3 \times 64$	$48 \times 48 \times 64$
MaxPool-1	—	—	$24 \times 24 \times 64$
Conv-2	—	$3 \times 3 \times 128$	$24 \times 24 \times 128$
CBAM	—	—	$24 \times 24 \times 128$
MaxPool-2	—	—	$12 \times 12 \times 128$
Inception1-1	Conv-3	$1 \times 1 \times 16$	$12 \times 12 \times 16$
—	S-Conv-1	$3 \times 3 \times 32$	$12 \times 12 \times 32$
—	S-Conv-2	$3 \times 3 \times 64$	$12 \times 12 \times 64$
Inception1-2	Conv-4	$1 \times 1 \times 32$	$12 \times 12 \times 32$
—	S-Conv-3	$3 \times 3 \times 64$	$12 \times 12 \times 64$
Inception1-3	MaxPool-3	—	$12 \times 12 \times 128$
—	Conv-5	$1 \times 1 \times 64$	$12 \times 12 \times 64$
Inception1-4	Conv-6	$1 \times 1 \times 64$	$12 \times 12 \times 64$
Concat	—	—	$12 \times 12 \times 256$
MaxPool-4	—	—	$6 \times 6 \times 256$
Inception2-1	Conv-7	$1 \times 1 \times 32$	$6 \times 6 \times 32$
—	S-Conv-4	$3 \times 3 \times 64$	$6 \times 6 \times 64$
—	S-Conv-5	$3 \times 3 \times 128$	$6 \times 6 \times 128$
Inception2-2	Conv-8	$1 \times 1 \times 64$	$6 \times 6 \times 64$
—	S-Conv-6	$3 \times 3 \times 128$	$6 \times 6 \times 128$
Inception2-3	MaxPool-5	—	$6 \times 6 \times 256$
—	Conv-9	$1 \times 1 \times 128$	$6 \times 6 \times 128$
Inception2-4	Conv-10	$1 \times 1 \times 128$	$6 \times 6 \times 128$
Concat	—	—	$6 \times 6 \times 512$
Conv-11	—	$3 \times 3 \times 512$	$6 \times 6 \times 512$
CBAM	—	—	$6 \times 6 \times 512$
MaxPool-6	—	—	$3 \times 3 \times 512$
Global-Pool	—	—	$1 \times 1 \times 512$
FC	—	—	$1 \times 1 \times 7$

2.3 CBAM-DSC 算法流程

CBAM-DSC 网络中的表情识别学习过程如算法1所示。其中 t 为迭代次数,初始值设为0。

算法1 CBAM-DSC 算法

输入:预处理后的人脸图像

输出:表情类别

1: $t \leftarrow t + 1$;

2:将人脸图像进行2次卷积操作和1次最大池化后进入CBAM模块运算。

3:经两个 Inception 模块的卷积操作和深度可分离卷积运算后调用CBAM模块。

4:将特征向量传入全局平均池化进行融合降维。

5:将结果与真实标签输入到损失函数中,

$$L_s = \sum_{i=1}^m \log \frac{e^{W_{yi}^T + b_{yi}}}{\sum_{j=1}^n e^{W_{ji}^T + b_{ji}}} \text{ 计算得到损失。}$$

6: $f(x) = \frac{1}{n} \sum_{i=1}^n f_i(x)$ 计算反向传播误差,通过

$x = x - \eta \nabla f_i(x)$ 调整参数并更新 x 。

7:输出表情类别。步骤3中假设输入为 $H \times W \times M$,卷积核为 $K \times K$,数量为 N ,输出特征图尺寸与输入保持一致。则传统卷积计算量的计算如式(3)所示:

$$Cv = K^2 \times H \times W \times M \times N \quad (3)$$

深度可分离卷积独立的将特征图的每一个通道进行分组再将每组的通道对应进行卷积,卷积后的通道数与输入通道数相同,得到 M 组 $H \times W$ 大小的特征图,将其组合为 $H \times W \times M$ 大小的卷积结果。计算量如式(4)所示。

$$DC = D_k^2 \times 1 \times D_i^2 \times M + M \times N \times D_i \times D_i \quad (4)$$

两种方式的计算量比值如式(5)所示。

$$\frac{DC}{Cv} = \frac{D_k^2 \times D_i^2 \times M + M \times N \times D_i^2}{D_k^2 \times M \times D_i^2 \times N} = \frac{1}{N} + \frac{1}{D_k^2} \quad (5)$$

由式(5)可知,再卷积核尺寸为 3×3 的卷积层中,深度可分离卷积的参数计算量是传统卷积的 $1/9$,优化了CBAM-DSC网络的计算量,提高网络模型的整体计算效率。

3 实验结果与分析

3.1 实验环境

本文提出的网络模型在GPU版本的Pytorch深度学习框架上进行实验,通过使用RAF-DB和CK+两种数据集进行训练和验证,对于效果最好的模型参数进行保存用于测试。

3.2 数据集介绍

RAF-DB数据集由29672个真实世界图像组成,数据集中的图片非常多样,每张图片有多个注释器进行标记。本次实验使用其中的一个子数据集。该数据集的7类表情分别为:惊讶、恐惧、厌恶、快乐、悲伤、生气、正常。为了保证RAF-DB数据集与

CK + 数据集的统一,分别删去 RAF-DB 中的正常类和 CK + 中的轻蔑类,保留 6 个相同的类别。

选取的 RAF-DB 数据集一共包括 10 000 多张图片,因为恐惧类的图片较少,快乐类的图片过多,对恐惧类图片进行处理,通过不同角度旋转增加恐惧类的图片数量,RAF-DB 数据库中的 6 种表情示例如图 5 所示。



图 5 RAF-DB 数据集六种表情示例

Fig. 5 Examples of six expressions in the RAF-DB dataset

CK + 数据集属于小样本数据集,由 327 个表情图片组成。实验中使用删减后的 6 个类别进行实验。CK + 数据库中的 6 种表情示例如图 6 所示。



图 6 CK + 数据集六种表情示例

Fig. 6 Examples of six expressions in the CK + dataset

在实验之前,由于数据集图片大小不同,人脸位置不同,所以将两个数据集中的图片进行人脸对齐并且灰度化,所有样本图片的大小统一为 48×48 的灰度图像,为了防止过拟合,对图片适当进行数据增强,将图片进行水平镜像和小幅度转动操作,数据集图片数量较增强前扩大 1.5 倍。

3.3 分析与对比

对于 RAF-DB 数据集,将数据增强后的 15 000 张图片分成两个集合,选取 13 000 张作为训练集,2 000 张作为验证集。网络训练过程中 epoch 设为 150,64 张图片为一批进行训练,学习率设置为 0.01,当迭代次数超过 30 后学习率每 5 轮衰减一次,每次衰减 5%,激活函数选取 ReLU 函数,通过随机梯度下降进行网络参数的优化。为了防止过拟合,CK + 数据集采用交叉验证法,将数据增强后的 460 张图片分成 4 份训练集和 1 份验证集,每个网络训练 60 个 epoch,同样选取 ReLU 函数作为激活函数,通过随机梯度下降进行网络参数的优化。

在网络模型的训练过程中将验证数据集损失值最低的模型作为最终的训练模型进行保存。将训练好的整体模型使用测试数据进行测试,分别在 RAF-DB 和 CK + 数据库上进行测试实验,表 2 为在 RAF-

BD 和 CK + 数据集上得到的表情识别率,针对 6 种不同的表情类别分别得到其对应的准确率。通过表中数据可知,在 RAF-DB 数据集中,网络模型 CBAM-DSC 识别率最高的是快乐类(Happy),与其他的类别相差较为明显,悲伤类别的识别率最低,6 种类别的平均识别率为 87.14%。在 CK + 数据集中六种类别的识别率相差较小,并且较为平均,平均识别率为 97.58%,整体识别率很高。通过实验结果可知,通过添加卷积注意力机制模块保证了较高的识别率。

表 2 数据集上的准确率

Tab. 2 Accuracy on both datasets

表情	RAF-DB 准确率/%	CK + 准确率/%
愤怒	90.16	97.34
厌恶	85.09	97.87
恐惧	89.23	98.64
快乐	94.11	98.53
悲伤	80.16	95.23
惊讶	84.11	97.89
平均	87.14	97.58

表 3 为在 RAF-DB 数据集上进行消融实验,其中 CNN 为在 Inception 模块中未使用深度可分离卷积以及模块前后未加入注意力机制的结果,CBAM-CNN 为在 Inception 模块前后加入注意力机制的结果,CBAM-DSC 为本文网络。

三种网络进行实验得到训练时间、参数量以及识别结果的数据对比,结果显示添加注意力机制以及使用可分离卷积均产生更好的效果,比传统卷积网络的效果更好,参数量减少的同时准确率得到提升。

表 3 消融实验

Tab. 3 Ablation experiments

	参数量	训练时间/min	准确率/%
CNN	2 737 342	130	80.43
CBAM-CNN	2 853 600	150	84.27
CBAM-DSC	2 563 248	110	87.14

图 7 为 CBAM-DSC 网络模型在两个数据集上进行测试实验得到的混淆矩阵,其中矩阵的行表示真实类,列对应的是预测类。

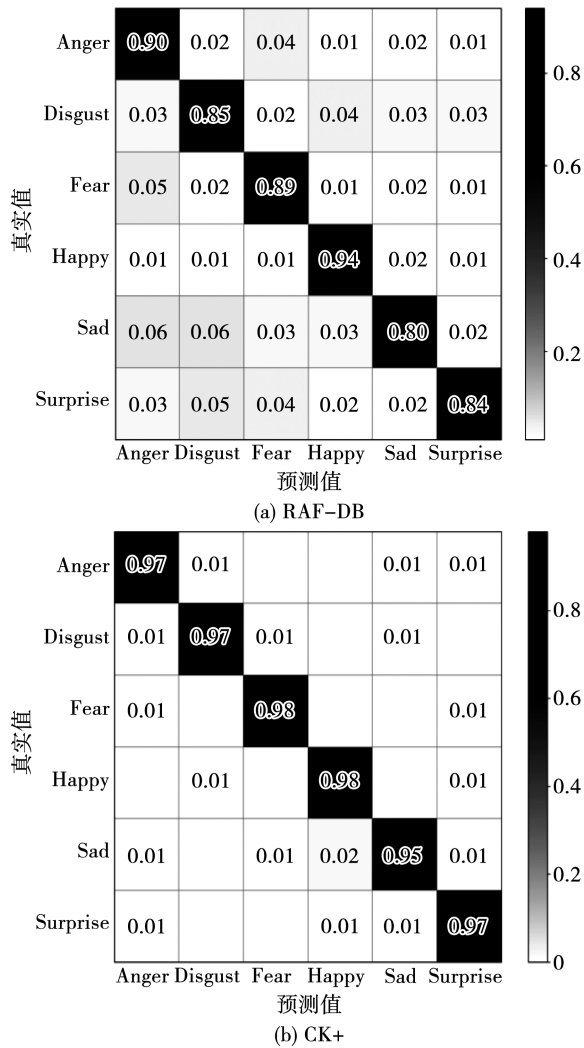


图7 6种表情的混淆矩阵

Fig.7 Confusion matrix of six expressions

在 RAF-BD 中因为各类表情之间差距不是很明显导致提取到的特征存在相似处,进而影响分类导致整体识别率较 CK+ 偏低,各类表情之间存在一定程度的混淆,识别时部分图片分到错误的类别中,从测试结果可知 Happy 表情因为和其他表情特征较容易区分,所以混淆的概率偏低,整体识别率最高。而 Sad 与其他表情存在较多相似特征,导致 Sad 表情识别过程中容易出错,识别率为 6 类之中的最低值。

在 CK+ 数据集上中每个表情都与其他个别表情存在轻微混淆,各类表情之间提取到的特征比较容易区分,所以整体识别率很高,Happy 和 fear 表情混淆程度最小,识别率最高,与 RAF-BD 数据集上的混淆矩阵一样,Sad 表情更容易因为混淆出现错误,不过差距相对较小。

为了验证 CBAM-DSC 网络模型的有效性,在 RAF-BD 和 CK+ 数据库上的实验结果与一些表情识别经典算法的识别率进行比较,表 4 为在 RAF-BD 和 CK+ 数据集上与几种模型准确率的对比。参与 RAF-BD 对比实验的算法模型有: ShuffleNet^[16]、ShuffleNetV2^[17]、MobileNetV2^[18] 和 SqueezeNeXt^[19]; 参与 CK+ 对比实验的算法模型有: Attention Net^[20]、MRE-CNN^[21]、Xception^[22]、Resnet18 + CBAM^[23]、Xception + CBAM。

表4 模型在两个数据集上的平均识别率

Tab.4 Average recognition rate of the model over two data sets

模型	RAF-BD 准确率/%	CK+ 准确率/%
ShuffleNef	78.65	—
ShuffleNetV2	78.85	—
MobileNetV2	80.14	—
SqueezeNeXt	82.40	—
Attention Net	—	90.30
MRE-CNN	—	95.79
Resnet + CBAM	—	96.80
Xception	—	94.50
xception + CBAM	—	94.90
本文方法	87.14	97.58

从实验结果可知,CBAM-DSC 在 CK+ 上实验得到的识别率较高,平均识别率为 97.58%。而在 RAF-BD 数据集上的识别率达到 87.14%,相较于其他模型识别率更高,提升幅度较大,高出 4%,在 CK+ 数据集上识别率高出 1%,可知在使用深度可分离卷积代替传统卷积并且添加了注意力机制对识别效果有很好的提升。整个网络在减少参数和计算量的情况下保证的较高的识别率。

4 结 论

本文针对人脸表情网络模型参数复杂和计算性能偏低等问题,提出一种基于 CBAM-DSC 网络的表情识别方法。通过卷积注意力机制进行多方位卷积,增强了网络提取有效信息的能力;深度可分离卷积分开提取通道维和空间维上的信息,提高了计算性能。试验结果表明,本文提出的基于 CBAM-DSC

网络的表情识别方法提高了准确率并降低了计算量。接下来的工作要继续扩大类间差距,使得在相似表情的识别上取得更好的效果。

参考文献:

- [1] 文元美, 欧阳文, 凌永权. 一种面向表情识别的 ROI 区域二级投票机制[J]. 计算机应用研究, 2019, 36 (9): 2861.
WEN Yuanmei, OUYANG Wen, LING Yongquan. A ROI Region Two-level Voting Mechanism for Expression Recognition[J]. Computer Application Research, 2019, 36 (9): 2861.
- [2] JUNG H, LEE S, YIM J, et al. Joint Fine-tuning in Deep Neural Networks for Facial Expression Recognition [C]//IEEE International Conference on Computer Vision. IEEE, 2015: 2983.
- [3] 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]//7th ACM International Conference on Multimodal Interaction, 2015: 427.
- [4] 张鹏, 孔韦韦, 滕金保. 基于多尺度特征注意力机制的人脸表情识别[J]. 计算机工程与应用, 2022, 58 (1): 182.
ZHANG Peng, KONG Weiwei, TENG Jinbao. Facial Expression Recognition Based on Multi-scale Feature Attention Mechanism[J]. Computer Engineering and Application, 2022, 58 (1): 182.
- [5] 程学军, 邢萧飞. 利用改进型 VGG 标签学习的表情识别方法[J]. 计算机工程与设计, 2022, 43 (4): 1134.
CHENG Xuejun, XING Xiaofei. Expression Recognition Method Based on Improved VGg Tag Learning[J]. Computer Engineering and Design, 2022, 43 (4): 1134.
- [6] 李春虹, 卢宇. 基于深度可分离卷积的人脸表情识别[J]. 计算机工程与设计, 2021, 42 (5): 1448.
LI Chunhong, LU Yu. Facial Expression Recognition Based on Deep Separable Convolution[J]. Computer Engineering and Design, 2021, 42 (5): 1448.
- [7] 魏赞, 李栋. 结合改进卷积神经网络与自编码器的表情识别[J]. 小型微型计算机系统, 2022, 43 (2): 387.
WEI Yun, LI Dong. Expression Recognition Based on Improved Convolutional Neural Network and Self Encoder [J]. Small Microcomputer System, 2022, 43 (2): 387.
- [8] 亢洁, 李思禹. 基于注意力机制的人脸表情识别迁移学习方法[J]. 计算机工程与设计, 2021, 42 (3): 797.
KANG Jie, LI Siyu. Transfer Learning Method of Facial Expression Recognition Based on Attention Mechanism [J]. Computer Engineering and Design, 2021, 42 (3): 797.
- [9] 韦赛远, 林丽媛, 张怡然. 基于 M-Xception 网络的戴口罩人脸表情识别[J]. 天津科技大学学报, 2021, 36 (3): 72.
WEI Saiyuan, LIN Liyuan, ZHANG Yiran. Facial Expression Recognition with Masks Based on M-Xception Network[J]. Journal of Tianjin University of Science and Technology, 2021, 36 (3): 72.
- [10] 关小蕊, 程卫月, 张雪琴, 等. 基于 GCSLBP 的 DBN 人脸识别算法[J]. 哈尔滨理工大学学报, 2022, 27 (2): 69.
GUAN Xiaorui, CHENG Weiyue, ZHANG Xueqin, et al. DBN Face Recognition Algorithm Based on GCSLBP [J]. Journal of Harbin University of Technology, 2022, 27 (2): 69.
- [11] 牛瑞华, 杨俊, 邢瀚馨, 等. 基于卷积注意力模块和双通道网络的微表情识别算法[J]. 计算机应用, 2021, 41 (9): 2552.
NIU Ruihua, YANG Jun, XING Lianxin, et al. Micro Expression Recognition Algorithm Based on Convolutional Attention Module and Dual Channel Network [J]. Computer Applications, 2021, 41 (9): 2552.
- [12] 张为, 李璞. 基于注意力机制的人脸表情识别网络[J]. 天津大学学报(自然科学与工程技术版), 2022, 55 (7): 706.
ZHANG Wei, LI Pu. Facial Expression Recognition Network Based on Attention Mechanism [J]. Journal of Tianjin University (Natural Science and Engineering Technology Edition), 2022, 55 (7): 706.
- [13] 林克正, 许松婷, 李骛, 等. 利用关键特征点的多信息融合人脸识别[J]. 哈尔滨理工大学学报, 2017, 22 (3): 19.
LIN Kezheng, XU Songting, LI Ao, et al. Multi Information Fusion Face Recognition Using Key Feature Points [J]. Journal of Harbin University of Technology, 2017, 22 (3): 19.
- [14] 李靖宇, 程卫月, 李子翔, 等. 超分辨率重建的微小人脸表情识别算法[J]. 哈尔滨理工大学学报, 2022, 27 (3): 52.
LI Jingyu, CHENG Weiyue, LI Zixiang, et al. Micro Face Recognition Algorithm for Super Resolution Reconstruction [J]. Journal of Harbin University of Science and Technology, 2022, 27 (3): 52.
- [15] 罗岩, 冯天波, 邵洁. 基于注意力及视觉 Transformer 的

- 野外人脸表情识别[J]. 计算机工程与应用, 2022, 58 (10): 200.
- LUO Yan, FENG Tianbo, SHAO Jie. Field Facial Expression Recognition Based on Attention and Visual Transformer[J]. Computer Engineering and Application, 2022, 58 (10): 200.
- [16] ZHANG X, ZHOU X, LIN M, et al. ShuffleNet: An Extremely Efficient Convolutional Neural Network for Mobile Devices[C]//2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). IEEE, 2018: 6848.
- [17] MA N, ZHANG X, ZHENG H T, et al. ShuffleNet V2: Practical Guidelines for Efficient CNN Architecture Design [C]//European Conference on Computer Vision. Springer, Cham, 2018: 116.
- [18] SANDLER M, HOWARD A, ZHU M, et al. MobileNetV2: Inverted Residuals and Linear Bottlenecks[C]//2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2018: 4510.
- [19] GHOLAMI A, KWON K, WU B, et al. SqueezeNext: Hardware-Aware Neural Network Design [C]//2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), 2018: 1638.
- [20] FERNANDEZ P D M, PEA F A G, REN T I, et al. FERAtt: Facial Expression Recognition With Attention Net [C]//2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW). IEEE, 2020.
- [21] BOROUMAND M, CHEN M, FRIDRICH J. Deep Residual Network for Steganalysis of Digital Images[J]. IEEE Transactions on Information Forensics and Security, 2018, 14(5): 1181.
- [22] CHOLLET F. Xception: Deep Learning with Depth Wise Separable Convolutions[C]//IEEE Conference on Computer Vision and Pattern Recognition. Hawaii: IEEE, 2017: 1800.
- [23] SANGHYUN W, JONGCHAN P, JOON-YOUNG L, et al. CBAM: Convolutional Block Attention Module[C]//European Conference on Computer Vision. Munich: Springer, 2018: 3.

(编辑:温泽宇)