

改进型的 MobileNet 的轻量级 人脸表情识别方法

吴振荣，邱卫根，张立臣
(广东工业大学计算机学院，广州 510006)

摘要：为了解决目前轻量级卷积神经网络 MobileNet 应用于人脸表情识别准确率不高、实时性差和时空负载大等问题，文中提出了一种改进型的 MobileNet 的轻量级人脸表情识别方法。该方法在 MobileNet X 的基础上，引入 SE 注意力模块并针对表情图像的特点对深度卷积层和网络结构进行优化，避免了信息丢失和神经元“坏死”问题，提高了模型的人脸表情识别率。与 MobileNet X 模型相比，改进后的网络模型复杂度低、识别精度高。在 Fer2013 人脸表情数据集上的实验证明，文中方法得到了 73.54% 的识别率，较其他表情识别方法在识别率和时间效率上都有一定提高。

关键词：人脸表情识别；卷积神经网络；轻量级；注意力机制；深度可分离卷积

中图分类号：TP319.4 **文献标识码：**A **文章编号：**1009-2552(2024)08-0044-07

DOI：10.13274/j.cnki.hdzj.2024.08.007

Improved mobileNet for lightweight facial expression identification

WU Zhen-rong, QIU Wei-gen, ZHANG Li-chen

(School of Computers, Guangdong University of Technology, Guangzhou 510006, China)

Abstract: In order to solve the problems of low accuracy, poor real-time performance and large space-time load in the application of lightweight convolutional neural network MobileNet to facial expression recognition, this paper proposes an improved MobileNet lightweight facial expression recognition method. Based on MobileNet X, this method introduces the SE attention module and optimizes the deep convolution layer and network structure according to the characteristics of the expression image, avoiding the problem of information loss and neuron “necrosis”, and improving the recognition rate of the model’s facial expression. Compared with MobileNet X model, the improved network model has low complexity and high recognition accuracy. The experiment on Fer2013 facial expression dataset shows that the recognition rate of this method is 73.54%, which is higher than other facial expression recognition methods in recognition rate and time efficiency.

Key words: facial expression recognition; Convolutional Neural Network (CNN); lightweight; attention mechanism; depthwise separable convolution

0 引言

近年来，人们对于人机交互的要求不断提高，让机器根据不同的情感做出不同的反应也是未来人工智能的重要发展方向^[1]。但现有的人脸表

基金项目：国家自然科学基金项目(61873068)

作者简介：吴振荣(1998-)，男，在读硕士研究生，研究方向为计算机视觉、图像处理、深度学习。

情识别模型存在识别精度低、泛化能力不足以及实时性差等问题。

人脸表情识别重点在于特征提取,目前已出现两类特征提取方法。一类是传统提取表情特征的方法;另一类是基于深度学习的表情特征提取方法。后者能同时进行特征提取和表情分类,避免了部分意想不到的关键点和特征无法被提取出来的问题。何晓云等人^[2]提出注意力机制和视频表情识别相结合的方法,使用 VGGNet16 作为骨架网络,对侧输出层的特征图进行加权融合;何超等人^[3]提出在 AlexNet 基础上引入 UCNN 的方法,有效地降低了网络的参数量;Ni 等人^[4]提出利用多尺度特征和注意力机制融合,并用改进的深度可分离卷积降低网络参数的方法。但这些方法在实际应用于表情识别时依然存在识别率低、泛化能力差和识别速度慢等问题。

针对上述问题,本文提出一个改进型的 MobileNet 的轻量级人脸表情识别方法 Mobile-SENet 模型。该模型引入 SE 注意力模块^[5],与深度可分离卷积层模块进行融合,立足人脸表情识别的特点,优化了网络结构。

1 相关工作

1.1 深度可分离卷积

传统标准卷积和深度可分离卷积的过程,如图 1 所示,其中 ReLU (Rectified Linear Unit) 表示激活函数,BN (Batch Normalization) 表示批量归一化。

图 1(a)为传统标准卷积。假设输入特征图 (Input Features) 的尺寸是 $H \times W$,通道数为 N ,标准卷积核的尺寸是 $D \times D$,通道数为 N ,其个数为 K 。若卷积后输出与输入尺寸一致,传统标准卷积计算量(FLOPs, F)大小为:

$$F = H \times W \times D \times D \times N \times K \quad (1)$$

图 1(b)为深度可分离卷积。假设输入特征图尺寸为 $H \times W$,通道个数为 N ,卷积核大小为 $D \times D$,通道数为 1,其个数为 N 。若输出特征图 (Output Features) 的尺寸和输入特征图尺寸一致,其计算量为:

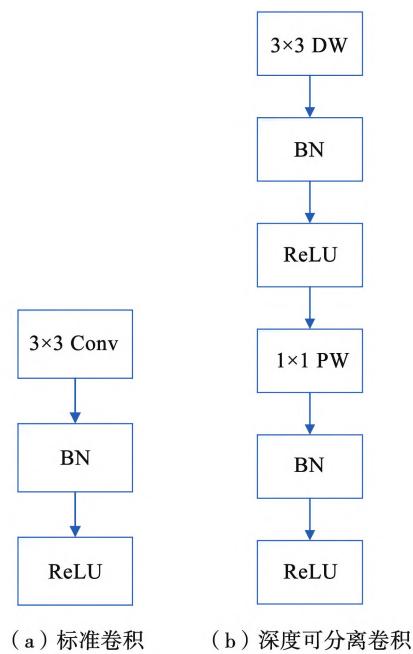


图 1 标准卷积与深度可分离卷积

$$F_w = H \times W \times D \times D \times 1 \times N \quad (2)$$

点卷积能有效解决深度卷积无法改变通道数的问题。其本身就具有通道融合的能力,较好地弥补了深度卷积无法有效利用同一位置在不同通道之间的信息缺陷,其计算量为:

$$F_p = H \times W \times 1 \times 1 \times N \times K \quad (3)$$

深度可分离卷积计算量与传统标准卷积计算量相比:

$$\frac{(F_w + F_p)}{F} = \frac{1}{K} + \frac{1}{D^2} \quad (4)$$

由式(4)可以看出,若网络所使用的卷积核尺寸为 3×3 ,那么深度可分离卷积的计算量大约是传统标准卷积计算量的 $1/9$,在保证精度损失不大的情况下,大大减少了网络的参数量和计算量。因此,深度可分离卷积也成为了轻量化网络的常用组件之一。

1.2 SE 注意力模块

针对人脸表情识别数据和 MobileNet 网络结构的特点,本文使用 SE 注意力模块对 MobileNet 模型进行优化。SE 模块由 Hu 等人提出,图 2 为 SE 模块的基本结构。

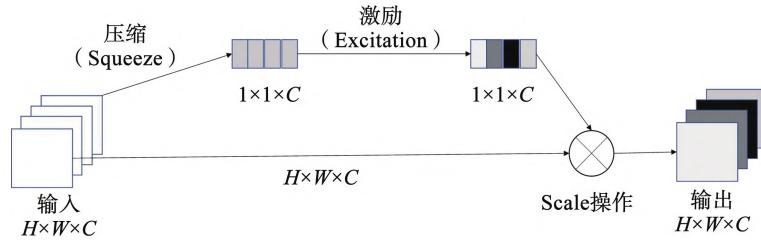


图 2 SE 模块的基本结构

该模块对输入特征图的操作主要分为三部分,分别是压缩(Squeeze)操作、激励(Excitation)操作以及特征图重标定(Scale)操作。

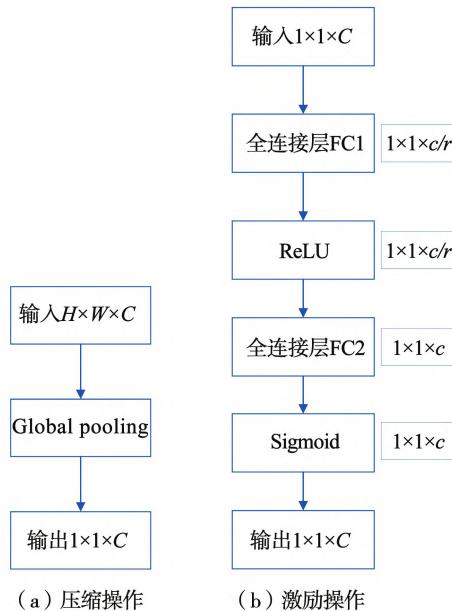


图 3 Squeeze 操作与 Excitation 操作

Squeeze 操作如图 3 所示。特征图 \mathbf{U}_c 通过全局平均池化将每个通道的二维特征 ($H \times W$) 压缩为 1 个实数,把特征图从 $[h, w, c] = [1, 1, c]$,这些数据从某种意义上来说是具有全局感受野的,最后将这些数据按照通道顺序串联起来形成一个特征向量 \mathbf{Z}_c ,每个通道如公式(5)所示:

$$\mathbf{Z}_c = F_{sq}(\mathbf{U}_c) = \frac{1}{H \times W} \sum_{i=1}^H \sum_{j=1}^W U_c(i, j), \mathbf{Z} \in R^c \quad (5)$$

其中, H, W 表示输入特征图的长和宽, U_c 表示上层卷积层第 c 个卷积核卷积后的输出。

如图 3 所示,Excitation 操作由“FC1 层 – ReLU 激活 – FC2 层 – Sigmoid 激活”组成,将上一步

Sequeeze 操作后得到的特征向量 \mathbf{Z} 通过第一个全连接层 \mathbf{W}_1 ,将特征向量的维度降至原来的 $1/r$ 。通过 ReLU 激活函数整合不同通道的特征,用全连接层 \mathbf{W}_2 将特征向量维度升回原始的维度,最后采用 Sigmoid 激活函数转化为一个 0~1 的归一化相关性权重 \mathbf{S} 。公式如(6)所示:

$$\begin{aligned} \mathbf{S} &= F_{ex}(\mathbf{Z}, \mathbf{W}) = \sigma(g(\mathbf{Z}, \mathbf{W})) \\ &= \sigma(\mathbf{W}_2 \delta(\mathbf{W}_1 \mathbf{Z})) \end{aligned} \quad (6)$$

其中, $\mathbf{W}_1, \mathbf{W}_2$ 表示两个全连接层; δ 表示 ReLU 函数; σ 表示 Sigmoid 函数; \mathbf{Z} 表示 Sequeeze 后的特征向量。

Scale 操作是将得到的权重 \mathbf{S} 与输入特征图进行乘法运算生成加权后的特征图,最终实现了在通道维度上对输入特征图权重的重新校准。公式如(7)所示:

$$\bar{\mathbf{X}}_c = F_{scale}(\mathbf{U}_c, \mathbf{S}_c) = \mathbf{S}_c \otimes \mathbf{U}_c \quad (7)$$

其中, S 表示 Excitation 后的权重向量; U 表示原特征图; c 表示第 c 个通道; X 表示最终的特征图。

SE 模块本质上是根据 loss 对特征权重进行重新校准,从而提高了网络模型对不同通道特征的敏感性。为了增强 SE 模块的非线性和通道之间的相关性,SE 模块特地采用了两个全连接层,并在两个全连接层之间设置 bottleneck 结构,有效地减少了模块本身带来的参数量和计算量。

1.3 H-Swish 激活函数

ReLU6 与 ReLU 函数图像如图 4 所示。相比于 ReLU,ReLU6 激活函数计算量小、收敛速度快,能够有效避免梯度消失和过拟合的问题,公式如(8)所示。但缺点也非常明显,由 ReLU6 的函数图像可以得知,当输入激活函数的值小于 0 时,对应的神经元的回传梯度会被置零,从而不会更新参数,使该神经元再也不能激活,即神经元“坏死”。

$$\text{ReLU6} = \min(6, \max(0, x)) \quad (8)$$

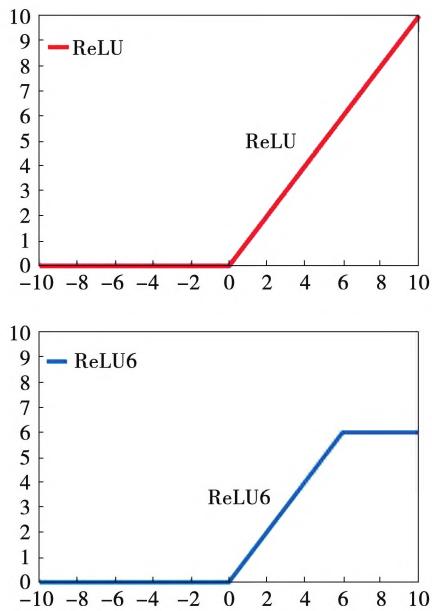


图 4 ReLU6 与 ReLU 函数图像

在 2017 年, Google 等人提出 Swish 激活函数代替 ReLU 激活函数, 用于缓解神经元“坏死”的问题。公式如式(9) – (10) 所示, 函数图像如图 5 所示。Swish 激活函数可以看作是介于线性和非线性之间的平滑函数。但 Swish 激活函数只有在深层网络中才能发挥作用, 计算量也很大。

$$\text{Sigmoid}(x) = \sigma(x) = \frac{1}{1 + e^{-x}} \quad (9)$$

$$\text{Swish}(x) = x \cdot \sigma(\beta x) \quad (10)$$

Howard 等人^[7]改用近似函数来逼近 Swish, 把 Swish 中的 Sigmoid 函数替换成 ReLU6 函数, 并称为 H-Swish。该函数不仅解决了神经元“坏死”的问题, 而且计算量小, 并不会对模型带来太大的计算负担, 公式如式(11) 所示, 函数图像如图 6 所示。

$$\begin{aligned} H - \text{Swish}(x) &= x \cdot \frac{\text{ReLU6}(x+3)}{6} \\ &= \begin{cases} 0 & \text{if } x \leq -3 \\ x & \text{if } x \geq +3 \\ x \cdot (x+3)/6 & \text{otherwise} \end{cases} \end{aligned} \quad (11)$$

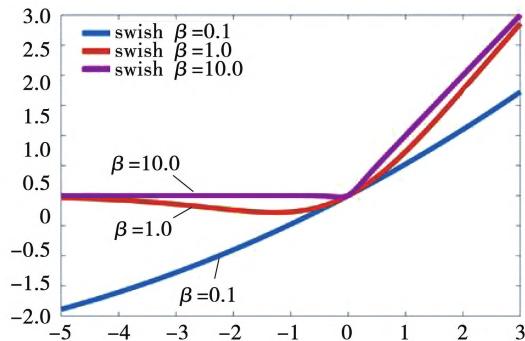


图 5 Swish 函数图像

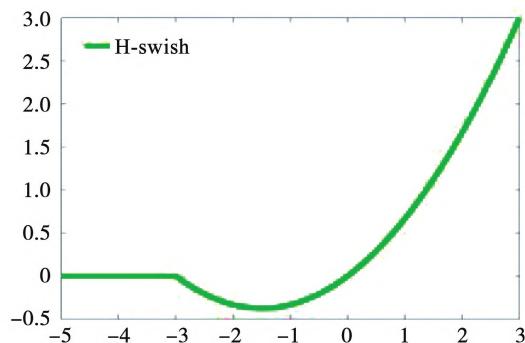


图 6 H-Swish 函数图像

2 改进方法

2.1 改进的深度可分离卷积

如图 7(a) 所示, MobileNet 的深度可分离卷积层, 采用了“深度卷积 – 点卷积”结构。MobileNet 中引入了 ReLU 非线性激活函数, 但此函数会带来一定的信息损耗, 特别是当通道数非常少的时候, 这种信息的损耗最为明显, 几乎是不可逆转的。而人脸表情图像多是单通道的, 这就导致在模型训练过程中部分信息丢失。

因此, 本文提出一种改进的深度可分离卷积层。如图 7(b) 所示, 通过在 3×3 深度卷积后引入 SE 注意力模块, 提高了网络模型对不同通道特征的敏感性。使用 H-Swish 激活函数替代 ReLU6 激活函数, 避免了神经元“坏死”的问题, 提高了网络的特征提取能力。将最后一个 1×1 点卷积的 ReLU6 激活函数替换成线性激活函数, 可在一定程度上避免信息丢失。若输入数据的维度较低, 那么深度卷积就只能在通道数较少的特征图中提取特征, 反而会导致识别率降低。因此, 需

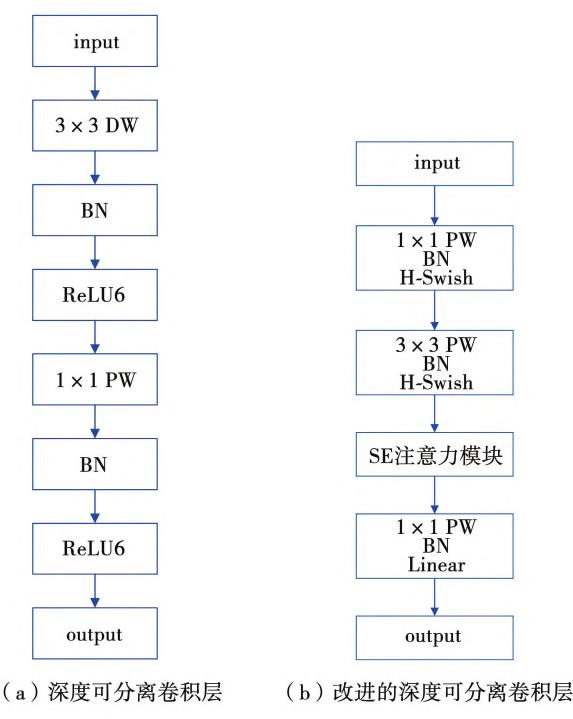


图 7 深度可分离卷积层结构

要在每个深度卷积前,配备一个点卷积专门用来升维,使其能在高维提取特征。改进后的深度可分离卷积层使用的是顺序级联结构,而非残差网络结构,因为后者更为复杂,会导致网络在表情识别中的表现变差。

2.2 网络结构调整

MobileNet 网络模型是针对 ImageNet 数据集分类设计的,本文只需要对 7 种表情数据进行分类,分类任务相对简单。由于表情图像的尺寸都比较小,把 MobileNet 的输入尺寸由 $224 \times 224 \times 3$ 改为 $42 \times 42 \times 1$,能有效地减少网络的参数量。为了提高模型的特征提取能力,把部分深度可分离卷积层替换为改进后的深度可分离卷积层。在一部分深度可分离卷积层中引入 SE 模块,提高网络模型对不同通道特征的敏感性。其网络结构如表 1 所示。

表中 SE 列代表是否在深度可分离卷积层中加入 SE 注意力模块。在 Mobile-SENet 网络中,首先让输入的特征图通过一个大小为 3×3 、步数为 2 的标准卷积层,然后通过 3 个不包含 SE 注意力模块的深度可分离卷积层、6 个包含 SE 注意力模

表 1 Mobile-SENet 网络结构

Type/Stride	Filter Shape	Input Size	SE	
Conv/s2	$3 \times 3 \times 1 \times 32$	$42 \times 42 \times 1$	-	
Conv dw/s1	$3 \times 3 \times 32 \text{ dw}$	$21 \times 21 \times 32$	-	
Conv/s1	$1 \times 1 \times 32 \times 64$	$21 \times 21 \times 32$	-	
Conv dw/s2	$3 \times 3 \times 64 \text{ dw}$	$21 \times 21 \times 64$	-	
Conv/s1	$1 \times 1 \times 64 \times 128$	$10 \times 10 \times 64$	-	
Conv dw/s1	$3 \times 3 \times 128 \text{ dw}$	$10 \times 10 \times 128$	-	
Conv/s1	$1 \times 1 \times 128 \times 128$	$10 \times 10 \times 128$	-	
×5	Conv dw/s1	$3 \times 3 \times 128 \text{ dw}$	$10 \times 10 \times 128$	✓
	Conv/s1	$1 \times 1 \times 128 \times 128$	$10 \times 10 \times 128$	✓
	Conv dw/s2	$3 \times 3 \times 128 \text{ dw}$	$10 \times 10 \times 128$	✓
	Conv/s1	$1 \times 1 \times 128 \times 256$	$5 \times 5 \times 128$	✓
	Conv dw/s1	$3 \times 3 \times 256 \text{ dw}$	$5 \times 5 \times 256$	-
Conv/s1	$1 \times 1 \times 256 \times 256$	$5 \times 5 \times 256$	-	
AvgPool/s1	Pool 6×6	$5 \times 5 \times 256$	-	
FC/s1		$1 \times 1 \times 256$	-	
SoftMax/s1		$1 \times 1 \times 7$	-	

块的深度可分离卷积层和 1 个不包含 SE 注意力模块的深度可分离卷积层,依次经过平均池化层、全连接层进行特征提取。为了避免网络过拟合和提高模型的泛化能力,加入了 Dropout 层。最后将提取后的特征输入给 SoftMax 层,用于实现表情图像的分类。在深度可分离卷积层中,在每个卷积层后都加入一个 BN 层以加快网络训练速度和收敛速度,之后还采用 H-Swish 激活函数为网络添加非线性因素。鉴于最后一个点卷积的特征图通道数都比较小,采用线性函数进行输出,避免信息丢失。为了提高网络模型对不同通道特征的敏感性,在深度卷积后引入了 SE 注意力模块。

3 实验与结果

3.1 实验环境

本文实验在 Ubuntu 20.04 LTS 64 位操作系统下进行,显卡型号为 NVIDIA GeForce RTX 2070 Super、显存 16GB, CPU 为 i7-10750H、主频 2.60GHz, 使用 Pytorch1.7.1 深度学习框架和 CUDA10.1 搭建网络模型。

实验中采用随机梯度下降优化器,其中参数设置为 momentum = 0.9, weight_decay = 0.0001,

学习率初始值设置为 0.01,采用 RLRP 的学习率衰减策略,当验证精度连续 5 个 epoch 保持稳定时,则将学习率降低至原来的 0.75 倍,一共进行 300 个迭代 epoch 训练,dropout = 0.2,批处理大小为 64。

3.2 数据集及预处理

本文提出的方法在 Fer2013 表情数据集中进行实验。Fer2013 表情数据集包含 28709 张训练图像、3589 张验证集图像和 3589 张测试集图像,共 35887 张表情图像。每张图像的大小为 48×48 ,均为灰度图像。这些表情图像包含“生气”“厌恶”“恐惧”“开心”“伤心”“惊讶”和“正常”七种基本表情,如表 2 所示。Fer2013 数据集是利用 184 个与情感相关的关键词在 Google Search API 中搜索,经过对齐调整大小后获得,最终被用于表情识别大赛中。由于该数据集的图像都是源自网络,其中难免会存在一些与人脸表情无关的图像,人眼在该数据集上的识别率为 $(65 \pm 5)\%$,给模型训练带来了挑战。

在训练表情识别网络前,为了增强训练的数据量,防止网络过拟合,进一步提高模型泛化能力,在训练阶段,使用 MATLAB 进行数据增强,把人脸表情图像随机裁剪为 10 个 42×42 像素的图像,对裁剪后的每个图像数据进行水平镜像操作,最后使得数据集扩大了 20 倍。效果如图 8 所示。

表 2 FER2013 数据集样本数量分布

类别	训练集(张)	测试集(张)
生气	3995	491
厌恶	436	55
恐惧	4097	528
开心	7215	879
伤心	4830	594
惊讶	3171	416
正常	4965	626

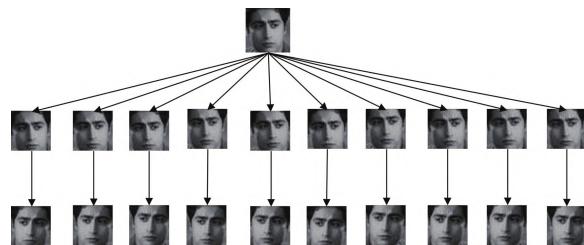
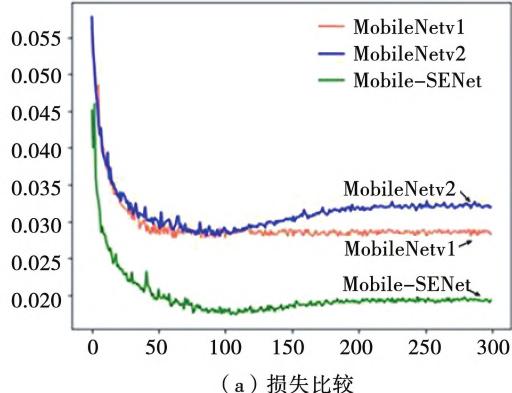


图 8 数据增强效果图

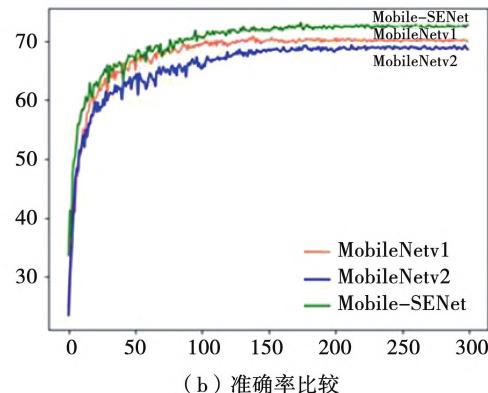
3.3 实验结果及分析

如图 9 所示,通过对 MobileNetv1、MobileNetv2 和本文提出的改进模型 Mobile-SENet 在数据增强后的 Fer2013 数据集上的损失(Loss)和准确度(Accuracy)的实验结果进行比较,可以看出,Mobile-SENet 模型的收敛速度明显优于 MobileNetv1 和 MobileNetv2,且更稳定,其准确率也高于其他模型,说明 Mobile-SENet 网络具有良好的特征提取能力。



— MobileNetv1
— MobileNetv2
— Mobile-SENet

(a) 损失比较



— MobileNetv1
— MobileNetv2
— Mobile-SENet

(b) 准确率比较

图 9 原始模型与改进模型比较

为了验证本文方法的有效性,本文从参数量、计算量、准确率和处理速度四个方面与其他现有

公开的人脸表情识别方法在 Fer2013 数据集上进行对比实验,识别结果如表 3 所示。

表 3 不同表情识别方法之间的对比

Methods	#Params/M	#FLOPs/M	Accuracy/%	Cpu time/ms
LBP-VGG ^[8]	138.3	16	68.85	565.214
FER-SoC ^[9]	102.4	9.6	66.0	531.642
gACNN ^[10]	62.3	0.727	68.2	341.533
Dam-CNN ^[11]	21.8	0.242	66.20	257.842
KLS-Net ^[12]	1.95	0.03	69.21	200.153
SGE-Net ^[13]	16.32	0.215	70.12	211.452
LA-Net ^[14]	15.03	0.283	70.25	261.235
本文方法	1.41	0.03	73.54	124.167

在模型参数量方面,本文方法的参数量远低于 LBP-VGG 的 138.3M、FER-SoC 的 102.4M 和 gACNN 的 62.3M。相较于参数量较小的 Dam-CNN、SGE-Net 以及 LA-Net,本文方法的参数量不到其的十分之一。

在浮点数运算方面,本文方法仅为 LBP-VGG 和 FER-SoC 的千分之一左右,约为 gACNN 的 1/24,即使是相比于计算量较小的 Dam-CNN、SGE-Net 以及 LA-Net,也仅为其中的 1/10 左右。

在准确率方面,相比于近期提出的 KLS-Net 和 SGE-Net,本文方法的准确率分别提升了 4.33% 和 3.42%;相比于 LBP-VGG、FER-SoC、gACNN 和 Dam-CNN,本文方法的准确率也分别提升了 4.69%、7.54%、5.34% 和 7.34%。虽然 LA-Net 和 SGE-Net 的识别准确率均与本文方法接近,但本文方法的参数量和计算量相比于前两者压缩了约 11 倍。

在处理时间方面,本文方法相比于 LBP-VGG 和 FER-SoC 压缩了约 4 倍;相比于 Dam-CNN、SGE-Net、LA-Net、KLS-Net、SGE-Net 压缩了约 2 倍。

综合来看,本文方法在准确率和处理时间之间达到了一个较好的平衡,在参数量和计算量不大、处理时间短的情况下,提高了表情识别的准确度,验证了此方法的有效性。

4 结束语

本文提出一种改进 MobileNet 的轻量级人脸

— 50 —

表情识别方法 Mobile-SENet,用于解决目前已知的人脸表情识别模型参数量大、识别率低,在计算资源有限的移动端和嵌入式设备难以部署和实际应用的问题。在 Mobile-SENet 网络模型中,引入 SE 注意力模块,在 3×3 的深度卷积前加上一个 1×1 的点卷积扩大输入的通道数,提高深度卷积的特征提取能力和网络模型的人脸表情识别率。以 H-swish 激活函数代替 ReLU6,最后以线性输出机制,有效地降低了 ReLU6 带来信息丢失率,避免了神经元“坏死”的问题。实验结果表明,与现有的多种表情识别方法相比,本文方法在参数量、计算量和准确率之间达到了一个较好的平衡,能满足部署到移动端设备中轻量化和高识别率的需求。下一步工作将会继续探索网络结构和分类算法,增强网络对细微表情的特征提取能力,提高表情间的识别和分类能力,进一步实现表情识别的实时化和智能化。

参 考 文 献:

- [1] 韩杰. 基于注意力融合卷积神经网络的人脸表情识别[D]. 哈尔滨:哈尔滨工业大学,2021.
- [2] 何晓云,许江淳,史鹏坤,等. 基于注意力机制的视频人脸表情识别[J]. 信息技术,2020,44(2):103–107.
- [3] 何超,侯明. 基于改进卷积神经网络的人脸表情识别方法[J]. 信息技术,2022(5):107–111,117.
- [4] Ni J, Zhang X, Zhang J. Multiscale feature fusion attention lightweight facial expression recognition[J]. International Journal of Aerospace Engineering,2022,2022:1–15.
- [5] Hu J, Shen L, Sun G. Squeeze-and-excitation networks[C]. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition,2018:7132–7141.
- [6] Goodfellow I J, Erhan D, Carrier P L, et al. Challenges in representation learning: A report on three machine learning contests[C]. International Conference on Neural Information Processing. Springer, Berlin, Heidelberg, 2013: 117–124.
- [7] Howard A, Sandler M, Chu G, et al. Searching for mobilenetv3[C]. Proceedings of the IEEE/CVF International Conference on Computer Vision, 2019: 1314–1324.

(下转第 58 页)

- difference coarray based MIMO radar array optimization with its application for DOA estimation [J]. *Multidimensional Systems and Signal Processing*, 2017, 28(4): 25 – 32.
- [3] Wu W, Cao Y, Wu C L, et al. Space-time coding waveform design for MIMO radar transmitbeamforming with minimum MFL [J]. *Digital Signal Processing*, 2019, 89: 124 – 130.
- [4] Shaikh A H, Dang X, Ahmed T, et al. New transmit-receive array configurations for the MIMO radar with enhanced degrees of freedom [J]. *IEEE Communications Letters*, 2020(99): 1.
- [5] Hongwu Liu, Guowei Yang, Gaojun Song. MIMO radar array synthesis using QPSO with normal distributed contraction-expansion factor [J]. *Procedia Engineering*, 2011, 15(2): 2449 – 1453.
- [6] 和洁, 冯大政, 孟超, 马仑. 采用遗传算法的 MIMO 雷达 L 型阵列稀布优化 [J]. *电讯技术*, 2015, 55(8): 872 – 878.
- [7] 庞育才, 刘松. 基于改进人工蜂群算法的 MIMO 雷达稀疏阵列优化 [J]. *系统工程与电子技术*, 2018, 40(5): 1026 – 1030.
- [8] 陆嘉胜, 张贞凯. 基于蚁狮算法的 MIMO 雷达阵列方向图优化设计 [J]. *电光与控制*, 2021, 28(4): 77 – 81.
- [9] 梁浩, 李小波, 徐旭宇. 基于改进自适应遗传算法的 MIMO 雷达阵列优化 [J]. *微波学报*, 2013, 29(4): 12 – 19.
- [10] 张筱, 魏钰洁, 姬伟杰, 等. 基于改进粒子群算法的 MIMO 雷达布阵优化 [J]. *指挥控制与仿真*, 2016, 38(5): 75 – 78.
- [11] 张伟, 何子述, 李军. MIMO 雷达稀疏阵优化设计 [J]. *系统工程与电子技术*, 2013, 35(2): 299 – 303.
- [12] 唐军奎, 刘峥, 谢荣, 等. MIMO 雷达稀疏阵列优化设计方法 [J]. *系统工程与电子技术*, 2022, 44(12): 3661 – 3666.
- [13] Heidari A A, Mirjalili S, Faris H, et al. Harris hawks optimization: algorithm and applications [J]. *Future Generation Computer Systems*, 2019, 97: 849 – 872.
- [14] 单斌斌, 李华, 谷瑞政, 等. 基于牛牛须搜索算法的短期风电功率组合预测 [J]. *科学技术与工程*, 2022, 22(2): 540 – 546.

(责任编辑:温长英)

(上接第 50 页)

- [8] Zeng N, Zhang H, Song B, et al. Facial expression recognition via learning deep sparse autoencoders [J]. *Neurocomputing*, 2018, 273: 643 – 649.
- [9] Vinh P T, Vinh T Q. Facial expression recognition system on SoC FPGA [C]. 2019 International Symposium on Electrical and Electronics Engineering (ISEE). IEEE, 2019: 1 – 4.
- [10] Li Y, Zeng J, Shan S, et al. Occlusion aware facial expression recognition using CNN with attention mechanism [J]. *IEEE Transactions on Image Processing*, 2018, 28(5): 2439 – 2450.
- [11] Xie S, Hu H, Wu Y. Deep multi-path convolutional neural network joint with salient region attention for facial expression recognition [J]. *Pattern Recognition*, 2019, 92: 177 – 191.
- [12] 刘劲, 罗晓曙, 徐照兴. 权重推断与标签平滑的轻量级人脸表情识别 [J]. *计算机工程与应用*, 2024, 60(2): 254 – 263.
- [13] 刘劲, 罗晓曙, 徐照兴. 空间分组增强注意力的轻量级人脸表情识别 [J]. *计算机工程与应用*, 2023, 59(22): 233 – 241.
- [14] Ma H, Celik T, Li H C. Lightweight attention convolutional neural network through network slimming for robust facial expression recognition [J]. *Signal, Image and Video Processing*, 2021, 15(7): 1507 – 1515.

(责任编辑:崔兴国)