

DOI:10.19651/j.cnki.emt.2415816

结合超轻量级双注意力模块的 ShuffleNetV2 面部表情识别*

林恩惠 王凡 谭晓玲

(重庆三峡学院电子与信息工程学院 重庆 404100)

摘要: 针对面部表情识别领域中难以同时实现低参数量与高准确率的挑战,提出了一种结合注意力机制的 ShuffleNetV2 网络的面部表情识别方法。该方法基于 ShuffleNetV2 架构,通过微调模型将 Relu 激活函数替换为 PRelu 激活函数,进一步提升了模型的特征捕获与分类能力。此外,本文创新性地引入了一种超轻量级双注意力模块 LDAM,该模块结合了 DCAM 注意力机制与空间注意力机制,并通过捷径连接技术集成到优化后的 ShuffleNetV2 模型中,以增强模型对细节特征的识别能力及分类效果。在 FER2013 和 CK+ 两大公认的面部表情识别数据集上的实验结果显示,本方法分别达到了 69.12% 和 94.77% 的识别准确率,同时保持了低至 1.25 的模型参数量。这一成果不仅展示了在保持模型轻量化的同时提升识别性能的可能性,而且通过实验验证了所提出方法的高效性和实用性。

关键词: 面部表情识别方法的改进;激活函数;空间注意力机制;轻量化模型;超轻量级双注意力模块

中图分类号: TN911.73 **文献标识码:** A **国家标准学科分类代码:** 510

Facial expression recognition of ShuffleNetV2 combined with ultra-lightweight dual attention modules

Lin Enhui Wang Fan Tan Xiaoling

(College of Electronics and Information Engineering, Chongqing Three Gorges University, Chongqing 404100, China)

Abstract: In the field of facial expression recognition, low parameters and high accuracy are difficult to complete together. A facial expression recognition method based on ShuffleNetV2 network combined with attention mechanism is proposed in this study. Based on ShuffleNetV2 architecture, this method further improves the feature capture and classification capability of the model by fine-tuning the model and replacing the Relu activation function with the PRelu activation function. In addition, this paper introduces an ultra-lightweight dual attention module LDAM, which combines DCAM attention mechanism and spatial attention mechanism, and integrates it into the optimized ShuffleNetV2 model by shortcut connection technology, so as to enhance the model's recognition ability and classification effect of detailed features. Experimental results on FER2013 and CK+, two widely recognized facial expression recognition datasets, show that the proposed method achieves recognition accuracy of 69.12% and 94.77%, respectively, while maintaining a model parameter count as low as 1.25. This result is not only demonstrated in the possibility of maintaining the lightweight model while improving the recognition performance, but also verifies the efficiency and practicability of the proposed method through experiments.

Keywords: improvement of facial expression recognition method; activation function; spatial attention mechanism; lightweight model; lightweight dual attention module

0 引言

在计算机硬件快速更新和人工智能技术持续进步的背景下,人们对于人机交互智能化的需求日益增长^[1-2]。面部表情识别技术,作为人机交互领域的一个重点研究方向,已

广泛应用于智能驾驶、现代医疗等多个领域^[3-7]。面部表情识别技术的核心,在于开发出具有高准确性的模型,这是衡量算法优劣的关键标准。

近年来,众多创新的模型相继出现。张鹏等^[8]尝试将注意力机制与空洞卷积技术结合起来,目的是提取面部表

收稿日期:2024-04-11

* 基金项目:重庆市重点实验室开放基金(ZD2020A0302)项目资助

情的多尺度特征,虽然这种方法在提取复杂面部表情特征方面有其独到之处,但在众多实际应用场景中,尤其是那些表情识别场景中,其效果并不尽如人意。刘娟等^[9]提出一种基于邻近平滑二值模式的面部表情识别,达到了较高的识别率,却无法降低模型的复杂度。吴昊等^[10]提出一种融合双编码局部二值模式算子和绝对梯度直方图算子的人脸表情识别方法,提升表情识别的准确性,而其在训练过程中产生了庞大的参数量并且提高了深度模型的复杂性,这是一个致命的缺点。孙静等^[11]提出了 HOAG 算法来提取面部表情,同时将其与改进后的局部二值模式算法结合,进一步提高了面部表情识别的准确率,却没有考虑到模型的参数量和计算量的问题。尽管这些方法在面部表情识别的准确性上取得了良好的效果,但它们通常受限于计算资源较为丰富的环境,难以直接应用于计算能力有限的移动设备上。

另一方面,Zhu 等^[12]将 ResNeXt 网络与金字塔卷积和注意力机制相结合,提升了网络在自然环境下的表情识别效果,然而却没考虑到由于模型变得更加轻量化,其识别效果并不理想。况朝青等^[13]利用 SqueezeNet 网络轻量化的特点,提出了一种基于边缘计算中极端姿态和表情的人脸识别,尽管在一定程度上简化了模型,但同时也导致了表情识别的准确性下降。夏亮等^[14]提出了使用 MobileNetV2 替换 VGG16 模型,有效地降低模型的规模和参数的数量,却难以达到高识别率和低参数量的平衡。

理想情况下,一个高效的模型应在保证准确率的同时,考虑到模型轻量化,避免模型参数庞大、训练周期长和反应慢等问题。为了有效的解决这些问题,本研究提出了一种结合超轻量级双注意力模块的 ShuffleNetV2 面部表情识别,研究工作如下:

1) 本文选择轻量化的 ShuffleNetV2 网络模型为基础模型,对其通过精细地调整,替换原有的 Relu 激活函数为 PRelu,使其应用于面部表情识别中,提高分类效果。

2) 引入了一种超轻量级双注意力模块(lightweight dual attention model, LDAM),该模块通过双池化通道注意力模块(dualpooling channel attention module, DCAM)注意力机制进行调优,找到一个最佳参数,结合空间注意力机制,并通过捷径连接技术集成。

3) 将 LDAM 注意力模块融入到改进的 ShuffleNetV2 网络模型中,增强网络的判别性特征,进一步提高面部表情识别的准确性。

总之,本文提出的方法,不仅在优化网络运行速度方面取得了显著成果,还在大幅降低参数数量的同时,保持了模型对面部表情的高准确识别率,展现出其在面部表情识别技术中的巨大潜力和实用价值。

1 本文方法

1.1 ShuffleNetV2 网络

本文选取 ShuffleNetV2^[15]网络为基础模型,该网络克

服了 ShuffleNet^[16]引入分组卷积和通道洗牌来优化性能与计算效率之间的平衡问题。ShuffleNetV2 引入了通道分离操作,既加速了网络处理速度,又实现了特征的有效重用。ShuffleNetV2 的核心改进之一是对通道分割和融合策略的优化,使得信息在网络中的流动更加高效。这种设计不仅提高了特征的利用效率,也进一步降低了模型的参数量和计算需求。因此,ShuffleNetV2 不仅在轻量级网络中展现了卓越的性能和速度,而且提供了更为优异的效率和准确性,特别是在移动设备和边缘计算场景下的应用。一个高效的网络设计必须遵守以下 4 个准则:1) 使用平衡性的卷积;2) 减少分组卷积的使用;3) 减少网络分支;4) 降低元素级操作次数。根据这 4 条准则设计出的 ShuffleNetV2 网络单元结构如图 1 所示。

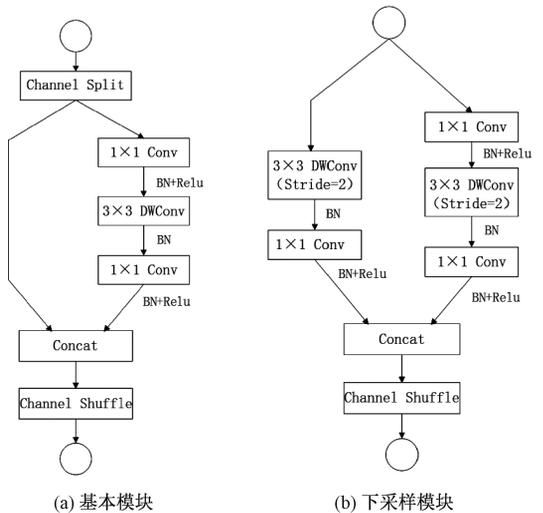


图 1 ShuffleNetV2 网络单元结构

1.2 DCAM 注意力模块

由于在对神经网络训练过程中并没有考虑到在特征图的通道与权值是需要直接对应的,这一点在很多模型中并没有体现出来,而 DCAM^[17]正是做到了这关键的一点。为了进一步提取面部表情的深层特征^[18],DCAM 注意力模块首先通过全局平均池化(global average pooling, GAP)和最大池化(max pooling, MAX)这两种方法对特征图进行压缩,使得特征图在通道重要性的表达更加突出。接着为了得到局部相邻通道间的交互信息,在此基础上,通过一维卷积,得到局部相邻通道间的交互信息,把权值向量提取出来。并且将这两部分通道特征进行相加,接着通过 Sigmoid 激活即式(1)中的 σ ,生成通道注意力加权系数 $N(H)$,将它与原输入特征 H 相乘就能得到含有更多通道关键信息的特征图 H' ,如式(1)和(2)所示。

$$N(H) = \sigma(A1B)(\text{AvgPool}(H) + A1B(\text{MaxPool}(H))) = \sigma(W_1(H_{\text{avg}}^C) + W_2(H_{\text{max}}^C)) \quad (1)$$

$$H' = N(H) \otimes H \quad (2)$$

其中,式(1)中 A1B 代表一维卷积, W_1 和 W_2 是分别

表示卷积的权重, σ 表示 Sigmoid 激活函数以确保得到的通道注意力加权系数处于 0~1, 从而可以作为加权系数来调整每个通道的重要性。

这种设计使得模块能够在保持计算效率的同时, 有效地提取并强调那些对面部表情识别最关键的信息。通过这种方式, DCAM 注意力模块显著提升了特征表达的能力, 使得模型能够更准确地识别面部表情。具体来说, 通过 GAP 和 MAX 池化压缩特征图并突出通道的重要性, 然后通过一维卷积处理这两种池化的结果以提取局部通道间的交互信息, 最后将这些信息综合起来, 并通过 Sigmoid 激活得到最终的通道注意力加权系数。这个过程不仅增强了模型对重要特征的捕捉能力, 而且也提高了面部表情识别的准确性, 进一步证明了在深度学习模型中加入注意力机制的有效性。

1.3 LDAM 注意力模块

LDAM 注意力模块采取了一种综合的策略来增强模型对特征的识别和处理能力, 先后通过 DCAM 和空间注意力模块对特征进行强化和提取。这种双重注意力机制使得模型能够综合通道和空间信息, 实现对输入特征的全面优化。

在这个过程中, 为了增强模型的稳定性并防止深层网络中常见的梯度相关问题, LDAM 在两个关键点上引入了捷径连接。图 2 为 LDAM 的结构示意图, 一是在 DCAM 模块的输入与输出之间, 二是在空间注意力模块中。捷径连接的引入, 使得前向传播过程中的信号可以直接传递, 避免了梯度在反向传播过程中的消失或扩散, 从而确保了即使在网络结构深入时, 模型的性能也不会受到梯度消失或爆炸的负面影响。

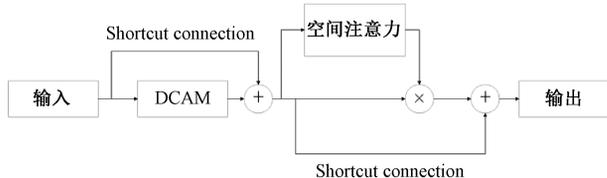


图 2 LDAM 结构示意图

这种设计思路不仅提高了特征增强的效果, 而且通过捷径连接保证了信息在深层网络中的有效流动, 增强了网络的训练稳定性。这对于面部表情识别而言, 能够显著提升模型的性能和可靠性。通过这样的结构设计, LDAM 能够充分利用深度学习模型的强大能力, 同时避免了深层网络训练过程中常见的问题。

1.4 改进后 ShuffleNetV2 模型

将 ShuffleNetV2 模型中的 Relu 激活函数替换为 PRelu(parametric relu) 激活函数, 是一种旨在增强网络特征提取和分类能力的改进。PRelu 作为 Relu 的一个变体, 引入了可学习的参数, 允许激活函数在训练过程中自适应地调整其形状。这一改进的动机源于 Relu 激活函数在处

理负输入时会遇到的问题——Relu 对于任何负输入值都输出零, 这在某些情况下会导致神经元的激活完全消失, 也就是所谓的“死神经元”问题, 从而影响模型学习的效率和效果。PRelu 激活函数通过引入一个小的斜率参数 α 来允许负输入值有一个非零的输出, 这个参数 α 是在网络训练过程中学习得到的。这意味着网络可以自适应地调整对负输入的处理方式, 从而避免了 Relu 中的梯度饱和问题。图 3 所示为 Relu 激活函数和 PRelu 激活函数二者在定义上的区别。如图 3(a) 所示, 可以看到, Relu 激活函数在负数区域的斜率是零, 而在正数区域的斜率是一。相比之下, 如图 3(b) 所示, 当 $f(y) < 0$ 时 PRelu 激活函数的斜率明显小于 $f(y) > 0$ 时的斜率, 但却不会为 0, 由于斜率较小, 可以避免产生 Dead Relu 问题。

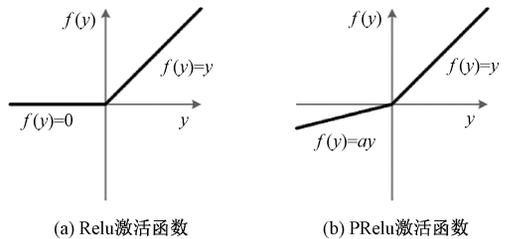


图 3 Relu 和 PRelu 的区别

具体而言, Relu 函数和 PRelu 函数表达式分别如式(3)、(4)所示。

$$Relu(y) = \begin{cases} y, & y > 0 \\ 0, & y \leq 0 \end{cases} \quad (3)$$

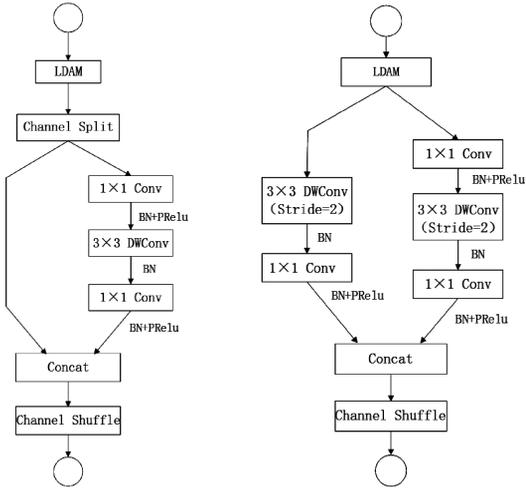
$$PRelu(y_i) = \begin{cases} y_i, & y_i > 0 \\ \alpha_i y_i, & y_i \leq 0 \end{cases} \quad (4)$$

由式(4)可知当参数为 0 时, 那么 PRelu 就会转化为原始函数 Relu 激活函数。

然后在 ShuffleNetV2 网络模型中引入注意力机制, 并对其进行微调, 图 4(a) 所示为改进后 ShuffleNetV2 的基本模块, 将输入特征先经过 LDAM 通道, 增强关键性通道特征的表达, 有效提升网络模型对面面部表情特征的识别能力, 然后对其特征图做通道分离(channel split), 并将其输入维度平均分为两个支路, 一条支路通过通道分离后仍保持不变; 而另一条分支则会先后通过 3 个卷积, 先经过一个步距为 1 的 1×1 的普通卷积, 在此基础上, 引入 BN 正则化与 PRelu 激励函数, 避免训练参数的过度拟合, 提高网络的非线性度, 紧接着经过一个 3×3 的深度可分离卷积, 并且同时加入 BN 正则化; 最后再经过一个 1×1 的普通卷积, 加入 BN 正则化和 PRelu 激活函数。然后把左右两条分支的结果通过 Concat 连接, 把其结果通过 Channel Shuffle 通道重排, 从而使通道间传输的信息得到更好的融合。

图 4(b) 所示为改进后 ShuffleNetV2 的下采样模块的单元结构, 与基础模块不同的是, 它是直接通过 LDAM 注意力模块而没有进行 Channel Split 这个操作, 而后将通道平均分成两条支路, 其中一条支路先后通过一个步距为 2

的 3×3 的卷积和一个 1×1 的普通卷积;而另一条分支则通过两个 1×1 的普通卷积和一个 3×3 的步距为 2 的深度可分离卷积。最后将这两条支路形成的结果通过通道重排的方式来实现不同组之间的信息融合。



(a) 基本模块

(b) 下采样模块

图 4 改进后 ShuffleNetV2 网络单元结构

2 实验

2.1 实验准备

为了验证本研究通过引入 LDAM 注意力机制的 ShuffleNetV2 轻量化的网络模型的性能,本研究的实验环境如下:具有 Core Inter i5-12500H 的 CPU,64 G 内存的 windows 操作系统,GPU 为 NVIDIA GeForce RTX 3050,安装有 Python 语言编译软件 Pycharm,Pytorch 深度学习环境的笔记本电脑。在训练中采用 Adam 优化器进行优化,初始学习率设定为 0.001,Batch Size 为 32,训练轮数为 300。

2.2 实验数据集及数据增强

本文采用的两种数据集为国内外公开的扩展 Cohn-Kanade(extended cohn-kanade,CK+)人脸表情数据集和 FER2013 人脸表情数据集。这两种数据集是大多数研究人员在面部表情识别中最常用的数据集,就 CK+数据集而言,这个数据集被设计用于支持面部表情识别等相关研究,涵盖了 7 种基本的表情类型。CK+数据集中 7 种对应标签的部分样例如图 5 所示。FER2013 人脸表情数据集是由 35 886 张人脸表情图片构成的,FER2013 数据集的特点是来源于现实世界的场景,图片多样性较高,包括不同的光照条件、面部遮挡和背景噪声等,这给表情识别带来了额外的挑战。不同的数据集对实验产生不同的结果,这两种数据集所处的环境是不同的,并且对应的 7 种标签也有所差异,本研究选取这两种不同的数据集,以确保训练出的模型能够在真实环境中取得良好的表现。FER2013 数据集中 7 种对应标签的部分样例如图 6 所示。



图 5 CK+数据集样例



图 6 FER2013 数据集样例

由于数据集的多样性和复杂性,在训练过程中有时会出现过拟合的现象^[19]。过拟合通常是由数据集分布不均或数据量不足引起的。为了解决这一问题,本研究采用了数据增强技术对 CK+数据集进行预处理,以增加模型的泛化能力,包括增强对比度和随机翻转等,这些技术能够有效地扩充数据集样本,增加数据的多样性,从而提高模型在面对未见过的数据时的鲁棒性和准确性。通过这些预处理和数据增强策略,提高面部表情识别模型的性能,尤其是在处理如 FER2013 这样复杂多变的数据集时。

2.3 实验步骤

针对面部表情识别任务进行了深入的探讨与实验验证,主要实验流程分为以下 4 个步骤,实验流程如图 7 所示。

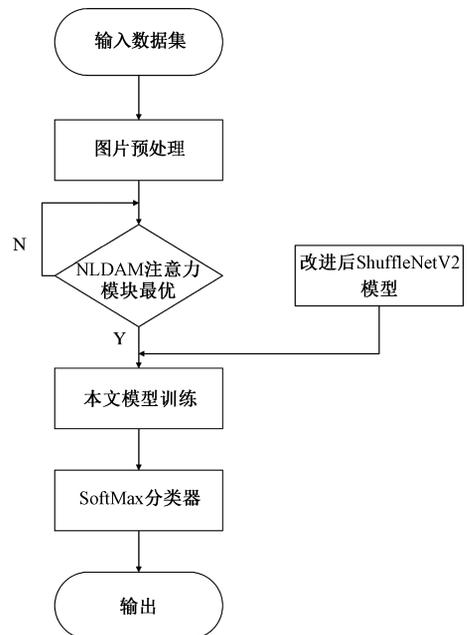


图 7 实验流程

首先,对实验数据集中的图片进行预处理,这一步骤非常重要。在预处理过程中,将所选数据集中的图像转换为灰度图像,然后利用多任务级联卷积网络(multi-task cascaded convolutional networks, MTCNN)来检测图像中的单一人脸,去除非人脸区域,最终将人脸剪裁为尺寸为

48×48 的灰度图像。将 CK+数据集通过增强对比度和随机翻转的方法扩充数据集至原来的 5 倍,以便后续的模式训练和评估。

其次,通过对超轻量化的 LDAM 注意力模型进行训练,并持续调整模型参数,以达到在 DCAM 注意力模块中找到最佳参数的目标,进而实现对图片特征进行整体加强的目的。

然后,对基础模块 ShuffleNetV2 网络模型进行微调,对参数初始化,降低损失函数。将 LDAM 注意力模块引入改进后的 ShuffleNetV2 网络模型中,并利用训练集对该注意力模型进行训练,以便模型更好地学习面部表情之间的关系。

最后,记录每次模型训练的数据并保存效果最好的模型。通过 SoftMax 分类器进行表情分类,获得面部表情的识别率。

2.4 实验结果与分析

为了确保实验结果的可靠性和准确性,采用了十折交叉法^[20]验证的方法来评估模型性能,这种方法通过将数据集均匀分成十份并轮流使用其中一份作为测试集,其余作为训练集,减少实验误差并提高评估的准确性。

为了验证本文所提模型的可靠性和准确性,与其他流行的网络模型如 VGG19、ResNet18、MobileNetV2 以及 ShuffleNetV2 进行了对比实验。通过在公开的 FER2013 和 CK+数据集上进行测试,本研究展示了各个模型在面部表情识别任务上的性能表现,包括识别准确率和模型的参数量,如表 1 所示。实验结果表明,在面部表情识别的准确率方面,本研究提出的模型在 FER2013 数据集上的表现略低于 ResNet18 模型,这是因为 ResNet18 模型有 18 个卷积层,比本模型包含更多卷积核和神经元,可以捕捉更复杂的特征,所以在有大量遮挡的 FER2013 数据集中识别率略高于本模型,但这个优点也使得 ResNet18 模型占用了大量的计算机资源,复杂度更高。同时本模型的识别率高于 VGG19、MobileNetV2 以及原版的 ShuffleNetV2 模型。特别是在 CK+数据集上,提出的模型相比于上述 4 种模型都展现出了显著的性能提升,准确率较 VGG19、ResNet18、MobileNetV2 和原版 ShuffleNetV2 分别高出了 2.59%、4.46%、0.12%和 5.48%。值得注意的是,本文提出的模型在参数量方面表现出极大的优势,远小于 VGG19 模型,并且在保持参数量低的同时,还能够实现较高的识别准确率。这一结果强调了将注意力机制整合进轻量级网络结构的有效性,不仅显著提升了模型在面部表情识别任务上的性能,也保持了模型的高效性和实用性。

图 8、9 分别展示了本研究提出的方法在 FER2013 数据集上和 CK+数据集上针对训练集与验证集的识别准确率表现。如图 8 所示,实线代表的是本文模型在 FER2013 训练集的训练,虚线是在验证集上的训练。在训练的早期阶段,模型的准确率增长迅速,随着训练次数的增加,图中

表 1 模型准确率和参数量比较

模型名称	FER2013%	CK+%	参数量/M
VGG19	68.53	92.18	77.23
Resnet18	69.75	90.31	42.83
MobileNetV2	62.50	94.65	8.65
ShuffleNetV2	60.32	89.29	1.4
本文	69.12	94.77	1.25

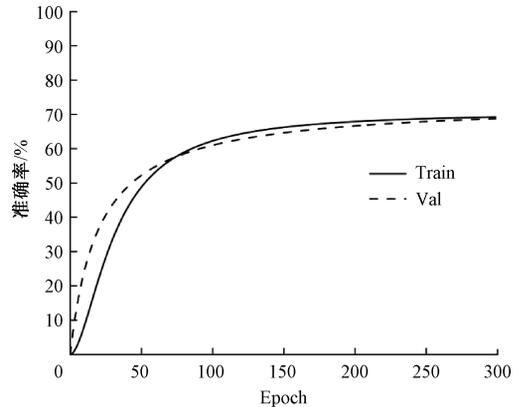


图 8 在 FER2013 数据集上的训练结果

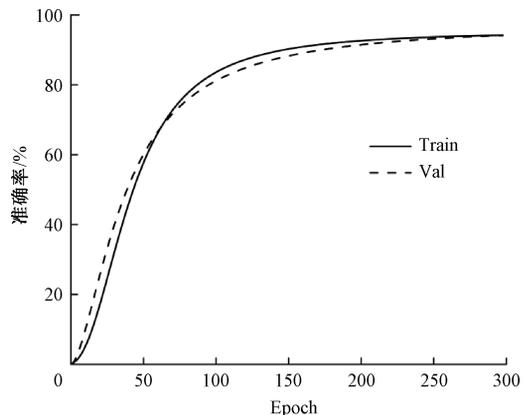


图 9 在 CK+数据集上的训练结果

的实现和虚线都显示出识别准确率的稳步提升;大约在 100 个 epoch 左右,增长速度开始减缓,但总体呈现增长趋势;而在接近 250 个 Epoch 时,模型的准确率稳定在一个相对平稳的水平。如图 9 所示,对 CK+数据集而言,模型在训练初期,识别率同样增长速度较快;当训练至第 100 个 Epoch 左右时,模型的增长渐渐变得缓慢;直至第 250 个 Epoch 后,模型的识别准确率变得很平稳。

总体而言,本文提出的结合 LDAM 注意力机制的改进 ShuffleNetV2 模型在综合性能上表现最优,证明了在设计面部表情识别模型时,通过精心选择和优化网络结构与机制,即使在参数量较低的情况下,也能够达到很高的识别准确率,展现了在资源受限环境下进行高效面部表情识别的潜力。

2.5 消融实验

为了进一步验证所提出方法的鲁棒性^[21],特别是注意力机制和 PRelu 激活函数在改善面部表情识别性能方面的作用。本文进行了消融实验。一共分为 3 个步骤:1)基线模型:使用未经修改的 ShuffleNetV2 网络模型,不加入任何注意力机制,也不对网络结构进行调整。这一步设置为控制组,以评估后续改进对性能的影响。2)激活函数优化:在 ShuffleNetV2 网络模型基础上,将 Relu 激活函数替换为 PRelu 激活函数。这一步骤旨在测试 PRelu 激活函数在提升模型捕获非线性特征方面的效果。3)加入 LDAM 注意力机制:在步骤 2)的基础上,进一步引入 LDAM 注意力模块,增强模型对重要特征的关注意度,从而提高识别准确率。

由表 2 可以看出,对 ShuffleNetV2 网络模型进行的微调,特别是将 Relu 激活函数替换为 PRelu 激活函数,虽然在参数量上没有显著变化,但识别准确率却实现了轻微的提升。这表明使用 PRelu 激活函数在面部表情识别任务中确实能够带来性能上的优化,可能由于 PRelu 提供了更好的模型非线性表达能力,以及对负输入值的有效处理能力。

表 2 消融实验识别结果

方法	CK+%	FER2013%	参数量/M
方法(1)	89.29	60.32	1.4
方法(2)	89.37	60.43	1.4
方法(3)	94.77	69.12	1.25

更进一步,在方法(2)基础上引入 LDAM 注意力机制后,模型在 CK+数据集上的识别准确率提高了显著的 8.4%,在 FER2013 数据集上也提高了 8.69%,同时模型的参数量还减少了 0.15。这一结果证明了 LDAM 注意力机制不仅在提高模型的识别准确率方面起到了关键作用,同时还实现了模型轻量化的目标。通过精确地关注重要的特征并降低对不重要特征的依赖,LDAM 机制有效地提升了模型的性能,同时减少了不必要的参数量,进一步提高了计算效率。

3 结 论

本文提出的一种结合了 LDAM 注意力机制的 ShuffleNetV2 网络模型为面部表情识别任务带来了显著的进步。通过精心调整 ShuffleNetV2,替换原有的 Relu 激活函数为 PRelu,本研究成功地增强了模型的性能。此外,引入的 LDAM 注意力机制进一步提升了模型对输入特征的处理能力,特别是在通道维度和空间维度上的判别性特征增强,使得模型能够更加细致和全面地理解输入数据。

与现有的几种流行模型——VGG19、ResNet18、MobileNetV2 以及原版 ShuffleNetV2——进行对比后,本

研究所提出的模型在 FER2013 和 CK+数据集上分别达到了 69.12%和 94.77%的识别准确率,同时保持了较低的参数量(1.25 M)。这一结果不仅展示了本文模型在准确率上的优势,也证明了其在参数上的优越性,特别是在考虑到模型轻量化对于资源受限环境的重要性时,本文的方法显示了其在实际应用中的巨大潜力。

同时本研究仍然存在一些不足之处,例如对一些在非正面和复杂环境下拍摄的表情图片识别方面存在挑战。未来研究可探讨如何提升这类表情图片的识别率,以满足日常生活中多变的面部表情需求。

参考文献

- [1] 徐桂芝,赵阳,郭苗苗,等.基于深度分离卷积的情绪识别机器人即时交互研究[J].仪器仪表学报,2019,40(10):161-168.
- [2] 张波,兰艳亭,鲜浩,等.基于通道注意力机制的人脸表情识别机器人交互研究[J].电子测量技术,2021,44(11):169-174.
- [3] 黄浩,葛洪伟.强化类间区分的深度残差表情识别网络[J].计算机科学与探索,2022,16(8):1842-1849.
- [4] 洪惠群,沈贵萍,黄风华.表情识别技术综述[J].计算机科学与探索,2022,16(8):1764-1778.
- [5] 韩凌,李栋,王增霖,等.基于人脸表情与脑电信号的情绪识别系统[J].国外电子测量技术,2023,42(10):190-195.
- [6] 孙耀尧,李秀茹,王松林.基于改进 YOLOv5 的学生面部表情识别[J].齐鲁工业大学学报,2023,37(1):28-35.
- [7] SAJJAD M, ULLAH F U M, ULLAH M, et al. A comprehensive survey on deep facial expression recognition: challenges, applications, and future guidelines[J]. Alexandria Engineering Journal, 2023, 68: 817-840.
- [8] 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):182-189.
- [9] 刘娟,胡敏,黄忠.基于区域 NSBP 特征的加权证据融合表情识别[J].电子测量与仪器学报,2020,34(11):132-139.
- [10] 吴昊,胡敏,高永,等.融合 DCLBP 和 HOAG 特征的人脸表情识别方法[J].电子测量与仪器学报,2020,34(2):73-79.
- [11] 孙静.基于多特征融合的人脸表情识别方法研究[D].沈阳:沈阳工业大学,2022.
- [12] ZHU X, HE Z, ZHAO L, et al. A cascade attention based facial expression recognition network by fusing

- multi-scale spatio-temporal features [J]. Sensors, 2022, 22(4): 1350.
- [13] 况朝青,贺超,王均成,等. 基于边缘计算中极端姿态和表情的人脸识别[J]. 电子技术应用, 2021, 47(6): 30-34.
- [14] 夏亮. 基于轻量级卷积神经网络的人脸检测研究[D]. 广州:广州大学, 2023.
- [15] YANG H J, LIU J X, MEI G M, et al. Research on real-time detection method of rail corrugation based on improved ShuffleNet V2[J]. Engineering Applications of Artificial Intelligence, 2023, 126, DOI: 10.1016/J.ENGAPPAL.2023.106825.
- [16] SUN W, FU B, ZHANG Z. Maize nitrogen grading estimation method based on UAV images and an improved shufflenet network [J]. Agronomy, 2023, 13(8):1974.
- [17] 史志博. 基于特征融合与深度学习的表情识别技术研究[D]. 北京:北京建筑大学, 2021.
- [18] UMER S, ROUT R K, PERO C, et al. Facial expression recognition with trade-offs between data augmentation and deep learning features[J]. Journal of Ambient Intelligence and Humanized Computing, 2022,13(2): 1-15.
- [19] 程换新,成凯,蒋泽芹. 基于注意力机制的 CNN 人脸表情识别[J]. 电子测量技术, 2021, 44(10):128-132.
- [20] 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.
- [21] 张晋婧,刘双峰,丰雷,等. 融合注意力机制的人脸识别算法研究 [J]. 国外电子测量技术, 2023, 42(2): 107-113.

作者简介

林恩惠, 硕士研究生, 主要研究方向为模式识别、图像处理和表情识别。

E-mail:360573386@qq.com

谭晓玲(通信作者), 教授, 主要研究方向为模式识别、图像处理和表情识别。

E-mail:tan-xiao-ling@163.com

王凡, 硕士研究生, 主要研究方向为模式识别、图像处理和文本识别。

E-mail:2389369690@qq.com