

基于多分支瓶颈结构的轻量化图像分类算法研究

苏田田,王慧敏,张小凤

(陕西师范大学物理学与信息技术学院,陕西西安 710119)

摘要: 传统卷积神经网络存在着参数量大、训练耗时长、轻量级模型的识别准确度不足的问题。本文提出了一种基于ResNet网络的多分支结构轻量化网络(Residual multi-branch structured Network, RemulbNet),通过在残差结构的主干中使用多分支结构增加特征多样性,利用变体的深度可分离卷积缩减模型参数量,采用Mish激活函数增加网络的非线性表达能力,在有效减少模型体积的情况下,提升网络的分类准确率。利用图像识别数据库,对网络性能进行测试。研究表明,对于5分类花卉识别问题,RemulbNet相比ResNet网络识别准确率提高3.9%,模型参数量减小71%,模型体积减小77%,缩短了约40%训练耗时;与轻量级网络(MobileNet v2和ShuffleNet v2)相比,RemulbNet在识别准确度、模型体积、训练时长和不同的图像分类数据集上都表现出优良的性能。

关键词: 轻量化网络;多分支瓶颈结构;Mish激活函数;深度可分离卷积;图像分类;卷积神经网络

基金项目: 国家自然科学基金(No.11874252)

中图分类号: TP391

文献标识码: A

文章编号: 0372-2112(2023)05-1319-08

电子学报URL: <http://www.ejournal.org.cn>

DOI: 10.12263/DZXB.20220920

Research on Lightweight Image Classification Algorithm Based on Multi-Branch Bottleneck Structure

SU Tian-tian, WANG Hui-min, ZHANG Xiao-feng

(School of Physics and Information Technology, Shaanxi Normal University, Xi'an, Shaanxi 710119, China)

Abstract: The traditional convolutional neural networks have many problems, such as large number of parameters, long training time, and insufficient recognition accuracy of the lightweight models. Based on ResNet network, a lightweight network named RemulbNet (Residual multi-branch structured Network) with multi-branch structure, which increases feature diversity with multi-branch structure in the backbone of the residual structure, reduces the number of model parameters with the depth-separable convolution of variants, and also increases the nonlinear expression capability of the network with Mish activation function. These measures can effectively reduce the model volume and improve the classification accuracy of the network. Using the image recognition database, the network performance is tested. For 5 categories of flower identification, RemulbNet improves the recognition accuracy by 3.9%, reduces the number of model parameters by 71%, reduces the model volume by 77%, and shortens the training time by about 40% compared with the ResNet network. Facing different image classification datasets, RemulbNet also shows excellent performance in terms of recognition accuracy, model volume, training time compared with the lightweight networks (MobileNet v2 and ShuffleNet v2).

Key words: lightweight networks; multi-branch bottleneck structure; mish activation function; deep separable convolution; image classification; convolutional neural network

Foundation Item(s): National Natural Science Foundation of China (No.11874252)

1 引言

卷积神经网络(Convolutional Neural Network, CNN)作为一种优良的机器学习方法,能够摒弃数据中的冗余信息,保留数据的本征信息,具有自主学习、权值共享、分类效果好等优势,在图像分类、目标检测、自然语

言处理等领域有着广泛的应用。

神经网络源于日本学者Fukushima Kunihiko提出的神经认知机模型(neocognitron)^[1],其在工程上的开创性研究应用,为人工智能技术的发展做出重要贡献。1989年Lecun Y, Bottou L提出了LeNet^[2]是最早的卷积神经网络,

其基本单元奠定了 CNN 的基础。而 2012 年提出的 AlexNet^[3,4]通过 ReLU 函数、数据增强和 dropout 等方法有效提高了 CNN 的性能。作为 2014 年 ILSVRC (ImageNet Large-Scale Visual Recognition Challenge) 竞赛的冠军, GoogLeNet^[5]的核心思想是 Inception 结构,通过使用不同大小的卷积核拓宽网络宽度,增加特征的多样性,并将结果进行拼接以降低网络的参数量和运算量。同年的 VGG-Net^[6]则是对网络深度的探索,虽然在图像识别方面略差于 GoogLeNet,但小卷积核的使用和加深网络方面给人以启示。2015 年,ResNet^[7]的残差结构减轻了网络的退化,使网络可以尽可能加深。同时,该网络还采用了批归一化 (Batch Normalization, BN)^[8]对批数据进行处理,该结构不但加快了网络的收敛速度而且增加了网络稳定性,缓解了梯度消失问题。后续各种改进的网络虽然在性能上有所提升,但体积都过于繁重,不便于在移动设备安装。

为了将深度学习模型部署在移动设备上,网络的轻量化成为研究热点。2016 年提出的 MobileNet^[9]使用倒残差结构和深度可分离卷积,在保证精度的同时减少模型的参数量与计算量。同年提出的 ShuffleNet^[10]是在 MobileNet 的基础上将逐点卷积进行分组操作以及 channel shuffle^[10,11]操作,在不影响精度的前提下使通道充分混合以降低计算量。2019 年的 EfficientNet^[12],采用倒置线性瓶颈模块, Swish 激活函数,每个模块都加入压缩-扩张 (Squeeze-and-Excitation, SE)^[13]机制,使网络具备较高的计算效率。

随着机器学习、人工智能等学科的蓬勃发展,轻量化网络的应用将更加广泛。但现有的模型中存在着模型体积较小的网络的准确率不高,而识别准确率高的模型体积庞大、训练耗时长等问题。如何设计出训练速度快、体积小、准确率高的网络成为 CNN 在实际应用中要解决的关键问题。本文基于 ResNet 网络的残差结构,提出了一种轻量化的网络 RemulbNet,该网络通过在 ResNet 的瓶颈 (bottleneck) 结构中使用多分支结构增加特征多样性,利用变体的深度可分离卷积缩减体积,采用 Mish^[14]函数增加网络非线性表达能力,在有效提升网络性能的同时提高计算效率。

2 网络模型

2.1 理论基础

2.1.1 ResNet

ResNet 的核心是残差模块,其结构如图 1 所示。当输入为 x 时,经过卷积输出的特征记为: $H(x)=F(x)$ 。引入残差结构后,学到的特征为 $H(x)=F(x)+x$ 。由于快捷连接的加入,当 $F(x)$ 为 0 时,卷积层只做恒等映射,这使得网络的性能不会下降。实际中 $F(x)$ 不为 0,便可以在上层的基础上不断学习,因此残差模块具有良好的学习性能。其快捷

连接的加入可以缓解网络退化问题,使网络深度增加。但 ResNet 的弊端在于网络深度的增加会导致运算激增、模型体积庞大,训练中会消耗更多的时间。

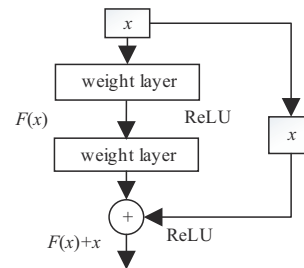


图 1 残差模块

2.1.2 Inception 结构

Inception v1 结构是 GoogLeNet 的核心模块,其结构主要由 1×1 卷积、 3×3 卷积、 5×5 卷积和 3×3 最大池化组成,四个部分处理结果在通道上进行拼接。Inception^[15] 结构的核心思想是利用不同大小的卷积核提取图像的信息,增加特征的多样性,提高网络的适应能力,并通过对特征进行融合,以更好的表征图像。GoogLeNet 网络虽然贡献了可以通过多分支处理和多样的特征矩阵拼接解决模型体积压缩和加速训练的问题,但也存在准确率不高的问题。

2.1.3 深度可分离卷积

深度可分离卷积 (Depthwise Separable Convolution, DSC)^[16] 可分为逐通道卷积 (Depthwise Convolution, DWC) 和逐点卷积 (Pointwise Convolution, PWC) 两步操作。首先,对尺寸为 $i \times j$ 的输入特征图 (Feature Map, FM),当通道数为 m 时的特征矩阵为 A ,使用大小为 $k \times l$ 的卷积核进行卷积,在逐通道卷积中,每一个通道都会相应的分配一个卷积核,接下来使用大小为 1×1 、通道数为 M 的卷积核进行逐点卷积。深度可分离卷积相比常规卷积的优势在于其能够在精度损失不多的情况下大幅减少网络的参数量和计算量,从而提高模型运行速率,减少训练时间。

深度可分离卷积的计算公式如下:

$$DSC(W_p, W_d, y)_{(i,j)} = PWC(W_p, DWC(W_d, y)_{(i,j)}) \quad (1)$$

$$DWC(W, y)_{(i,j)} = \sum_{k,l} W_{(k,l)} * y_{(i+k,j+l)} \quad (2)$$

$$PWC(W, y)_{(i,j)} = \sum_m W_{(m)} * y_{(i,j,m)} \quad (3)$$

其中, $W_{(k,l)}$ 为大小为 $k \times l$ 的卷积核, y 为输入特征图, i, j 为输入特征图尺寸, K, L 为输出特征图尺寸, W_d 表示逐通道卷积, W_p 表示逐点卷积, $W_{(m)}$ 代表大小为 1×1 、通道数为 m 的卷积, $W_{(k,l)} * y_{(i+k,j+l)}$ 表示 $k \times l$ 的卷积核与 $i \times j$ 的输入特征图 y 进行卷积。

2.2 RemulbNet 网络

受到残差学习、Inception 结构及深度可分离卷积等思想的启发,本文在 ResNet 的基础上提出了一种轻量化网络,简称 RemulbNet. 该网络主要从整体构建、模块设计和激活函数选择三个方面对网络进行了设计.

2.2.1 轻量化设计

RemulbNet 的整体框架如图 2 所示,网络对经预处理后的彩色图片的处理步骤如下:

Step1:将 3 通道的图像(尺寸为 224×224)输入到 7×7 的卷积层及 3×3 的池化层进行噪声滤除和维度变换,得到 56×56 的 64 通道的特征图.

Step2:对特征图使用一个多分支残差模块进行特征提取,得 56×56 的 256 通道的特征图.

Step3:经过一个残差模块,处理得到 28×28 的 512 通道的特征图.

Step4:再经过两个残差模块处理得到 14×14 的 1024 通道的特征图.

Step5:利用一个残差模块处理得到 7×7 的 2048 通道的特征图.

Step6:通过池化层和全连接层进行特征组合,对图像进行分类.

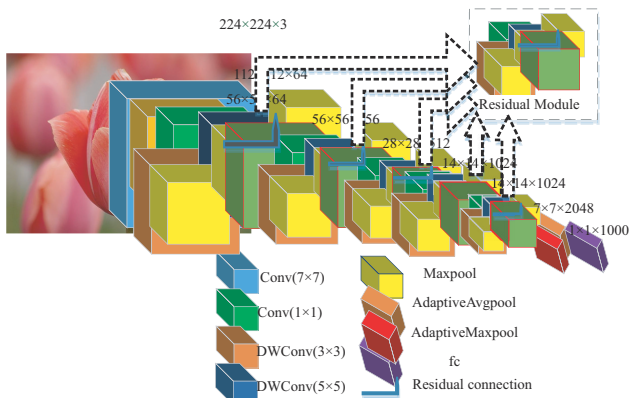


图 2 RemulbNet 网络

从图 2 所示的整体框架可知,RemulbNet 在宏观的网络计算层数配置上,卷积层的总数为 22,其多分支残差模块的配置分别为 1, 1, 2, 1, 该配置与 ResNet50 网络的残差模块设置 3, 4, 6, 3 类似,但结构更加简单,有效减小了参数量、运算量和训练耗时,提升了网络性能.

2.2.2 多分支残差模块

多分支残差模块作为 RemulbNet 的核心,由残差学习,深度可分离卷积以及多分支^[17]结构组成,结构如图 3 所示. 首先,对输入使用 1×1 的卷积进行降维;接下来运用多分支结构对数据进行处理. 其中,分支一使用大小为 3×3 卷积,之后使用 3×3 的最大池化;分支二使用 5×5 卷积核,再进行池化;分支三仅进行池化操作. 将

上述各分支所得的特征矩阵求和,使用逐点卷积进行特征合并和维度变换. 最后,引入残差学习将上述数据与下采样后的原始特征叠加在一起,得到多分支残差特征矩阵,作为下一级的输入.

多分支逐通道卷积的计算公式为

$$DWC(W, y)_{(i,j)} = \sum_{k,l}^{K,L} Pool(W_{(k,l)} * y_{(i+k,j+l)}) + \sum_{k,l}^{K,L} Pool(W_{(k,l)} * y_{(i+k,j+l)}) + Pool(y_{(i+k,j+l)}) \quad (4)$$

其中,Pool 为池化操作,其余参数含义与普通深度可分离卷积相同.

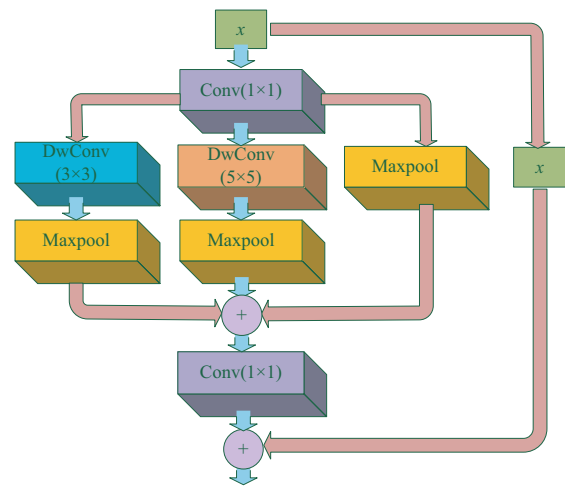


图 3 多分支残差模块

由于多分支残差模块的特征提取策略为卷积加池化,该结构不仅保证了特征提取的高效性,又保留了数据的主要特征,同时简化了网络复杂度,减少了网络计算量. 虽然其多分支结构与 Inception 以及深度可分离卷积的思想相同,但其使用的是各分支处理的结果之和,即多分支逐通道卷积,在增强特征多样性的同时,有效降低了模型的参数量与运算量,使得每一个维度的特征矩阵都具有全面的特征信息,有利于网络学习能力的提升,在反向传播过程中更有利于误差的传递. 此外,多分支残差模块在分支中使用较大的卷积核,能够增大感受野,显著提升网络性能.

2.2.3 激活函数的选择

激活函数作为卷积神经网络中的基础元素,具有提高模型鲁棒性,缓解梯度消失问题和增强非线性的特性. 在传统网络中,通常使用 ReLU 作为激活函数,

$$ReLU(x) = \max(0, x) \quad (5)$$

其中, $\max(a, b)$ 表示取 a, b 之中的最大值. ReLU 函数具有收敛速度快、计算复杂度低、能有效缓解梯度消失问题等优点,但该函数不具连续性,在负值域的参数会

“失活”,不会被学习更新.在本模型中,选择 Mish 函数作为激活函数,其定义如下:

$$\text{Mish}(x)=x \times \text{Tanh}[\log(1+e^x)] \quad (6)$$

其中, Tanh 代表双曲正切函数. Mish 具有与 ReLU 函数相同的上无界下有界特性,可以避免梯度饱和问题,能加速训练过程.同时, Mish 函数光滑且具有非单调性,在负值域能以较小的数值稳定网络梯度流,具有较好的泛化能力和优化能力,可以很好的提高网络性能.

RemulbNet 网络继承了残差网络的优良特性,充分发挥了 Inception 的优点,同时嵌入深度可分离卷积,很好契合了三种网络的优良特性.该模型不仅解决了残差网络体积庞大的问题,而且改善了经典轻量级网络准确率不足和训练耗时长的的问题,是一种准确率高的轻量化网络.

RemulbNet 网络的具体流程如算法 1 所示.

算法 1 基于多分支瓶颈结构的轻量级图像分类算法

输入:图片数据集

输出:图片的类别与准确率

1. 对图片进行预处理;
2. 从数据中选取 N 个样本,输入网络进行前向传播,得到各层及全连接层输出;
3. 根据全连接层输出与样本标签,利用网络的反向传播,计算各层误差项;
4. 依据网络各层的输出项与误差项,更新权值;
5. 按照设定迭代次数,循环步骤 2~4,保存模型,完成训练;
6. 加载模型,对测试集进行分类.

3 实验与结果分析

3.1 仿真实验条件

表 1 所示为本文所用的图像数据集,其均可在网站和鲸社区图像数据集模块(<https://www.heywhale.com/home/dataset>)中获取,当中牛津花卉、昆虫和车辆数据集由原数据集中挑选而出.在实验中使用程序将数据集以 8:1:1 划分为训练集、验证集、测试集(CIFAR-10 除外).鉴于部分数据集的数量较少,直接传入网络会出现过拟合现象,影响识别准确率,所以在训练前进行图像预处理.本文采用

的预处理(Pre-Processing, PP)方法包括:随机大小裁剪;随机旋转;随机水平翻转等,以增强网络的泛化能力.

实验环境为:CentOS 操作系统, Intel(R) Xeon(R) E5-2640 v3 处理器, CPU 逻辑核数为 32, 内存 125.65 GB, 主频 2.6 GHz, GeForce RTX 3090 (24 GB 显存) 显卡. Python 版本为 3.9.1, Pytorch 版本为 1.10.0, CUDA 版本为 11.3. 使用 Adam 优化器, 交叉熵损失函数, 学习率设置为 0.000 1, batchsize 设置为 32 张, 迭代次数为 200.

表 1 数据集信息

数据集	类别	总数/张
常见花卉	5	3 670
牛津花卉	17	2 391
昆虫	3	5 969
车辆 ^[18]	5	4 028
CIFAR-10	10	60 000

3.2 实验结果分析、讨论

为了探究 RemulbNet 的性能,将分别从 RemulbNet 模型与经典网络性能的对比、不同数据集实验对比以及学习率、激活函数和优化器选择对于模型性能影响的讨论等几个方面对网络的性能进行评价.实验结果采用去掉最小、最大值后取平均的统计方法,且每个网络的评价指标都是通过十组以上的重复实验得到.

3.2.1 消融实验

首先,以 5 分类花卉识别为例,对 RemulbNet 网络进行消融实验,其中各模型的残差模块如图 4 所示,实验结果如表 2 所示.

由表 2 可以看出,相比传统的 ReLU,选择 Mish 为激活函数,能将准确率提高约 2.3%,且对网络结构没有影响;在 ResNet50 中使用深度可分离卷积如图 4(b)和 4(c),可以减少约 44% 的参数量,约 6.32% 的训练耗时,从而加速训练;使用 5×5 卷积核如图 4(d)可以获得广阔的感受野,提升 2% 的准确率;而采取数据预处理,可以在不改变网络结构的条件下,加强网络的泛化能力,提高约 3%~5% 的准确率. RemulbNet 通过使用 Mish 函数、多分支处理、逐通道卷积融合等技术,比 ResNet50 提高约

表 2 消融实验

网络	准确率/%	参数量/MB	体积/MB	训练时间/s
ResNet50	85.1	25.56	90.02	2 502
ResNet50(Mish 函数)	87.4	25.56	90.02	2 592
ResNet50(DWConv(3×3))	84.9	14.27	46.97	2 344
ResNet50(DWConv(3×3)+Maxpool)	84.1	14.27	46.98	2 381
ResNet50(conv(5×5))	87.1	45.67	166.77	2 728
ResNet50(PP)	89.7	25.56	90.02	2 497
RemulbNet	89.0	7.50	20.98	1 495
RemulbNet(DWConv+Maxpool)	86.4	7.50	20.98	1 433
RemulbNet(DWConv+Avgpool)	87.9	7.50	20.98	1 397

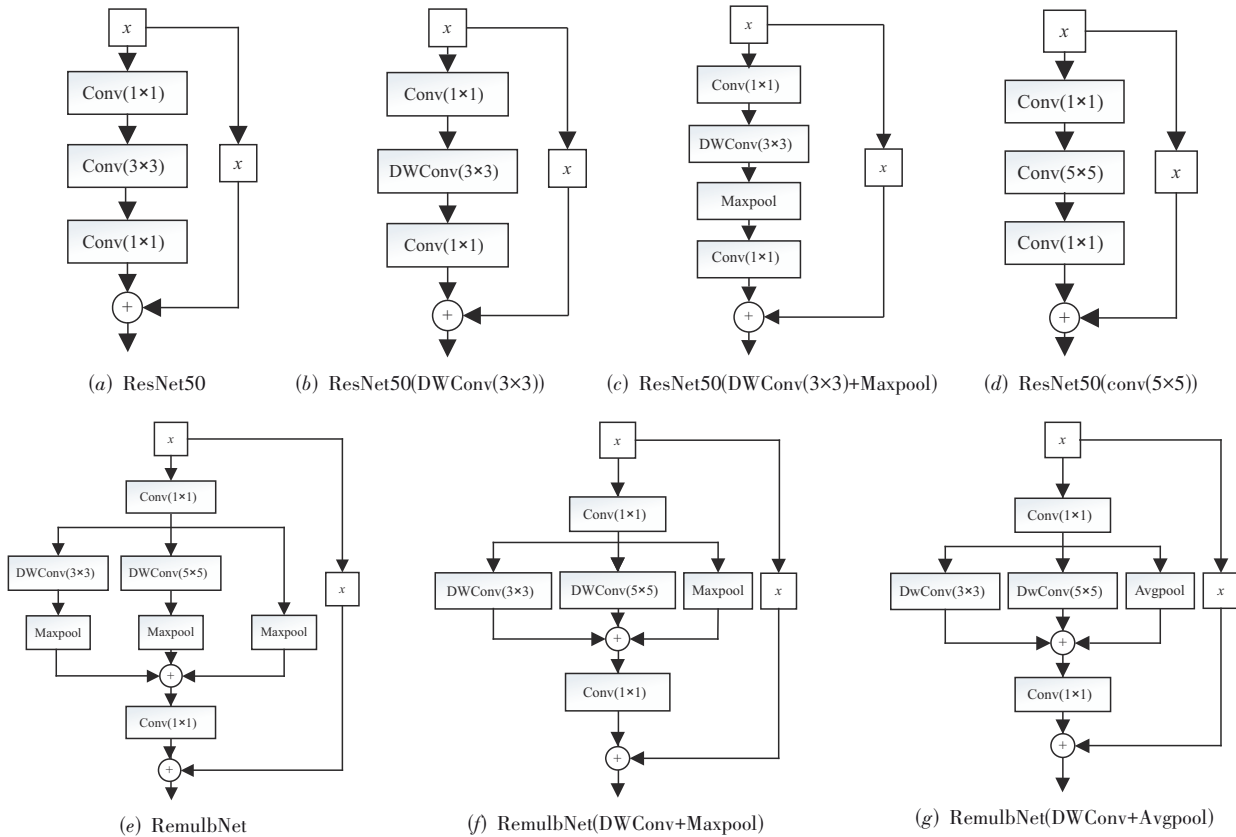


图4 不同结构的残差模块示意图

1.3%~3.9%的准确率,减少约71%的参数量,约40%~44%的训练时间,大幅提高了网络的综合性能。

针对 RemulbNet 的多分支残差模块有三种结构,通过花卉分类的实验证明,图 4(e)所示的卷积加池化的结构中,使用步长为 1 的卷积可以更全面的提取特征,池化对重要信息可以很好的保留. 通过实验证明在堆叠式结构中使用卷积加池化的组合具有良好的性能. 而在图 4(f)和 4(g)中使用卷积与池化的组合,由于卷积中的步长可变,导致步长较大时信息遗漏,因而分类结果稍有降低.

3.2.2 不同网络模型的对比

为了验证 RemulbNet 模型在准确率、体积和训练耗时方面的性能,使用花卉 5 分类数据集,将 RemulbNet 与 ResNet50、EfficientNet v2^[19]以及经典的轻量型网络 (MobileNet v2^[20,21]和 ShuffleNet v2^[22])进行了对比,结果如表 3 和图 5 所示.

由表 3 及图 5 可以看出,在训练中 RemulbNet 收敛速度较快,训练损失明显低于其他网络,损失的最小值为 0.07;在验证集上,本模型的识别精度可以很快达到稳定阶段且与准确率为 91%的 EfficientNet v2 网络表现相当;在测试集上,该网络的准确率比 ResNet50 高出 3.9%,比 ShuffleNet v2 高出 7.9%,也比其他的网络高出

表3 花卉识别实验(5类)

网络	准确率/%	参数量/MB	体积/MB	训练时间/s
RemulbNet(PP)	91.6	7.50	20.98	1 536
RemulbNet	89.0	7.50	20.98	1 495
ResNet50	85.1	25.56	90.02	2 502
MobileNet v2	81.6	3.51	8.74	2 089
ShuffleNet v2	81.1	2.28	4.97	2 284
EfficientNet v2	86.6	21.46	77.85	4 881

约 2%~7%. 在模型体积方面,本模型相比 ResNet50 体积缩减了 77%,比 EfficientNet v2 体积减少 73%,但还是略逊于其他的轻量型网络. 在训练耗时上,本模型相比 ResNet50、ShuffleNet v2 和 MobileNet v2 网络节约了 28%~40%的训练时间,比 EfficientNet v2 网络缩减了约 69%的训练耗时. 即,在综合性能方面,RemulbNet 模型优于其他对比的网络.

3.2.3 不同数据集实验对比

使用 17 分类花卉、昆虫、车辆、CIFAR-10 数据集,对 RemulbNet 的性能进行了进一步的测试,并与 ResNet、EfficientNet v2、MobileNet v2 和 ShuffleNet v2 网络进行了比较,结果如表 4 所示.

由表 4 可以看出,各个网络对于花卉 17 分类均可以

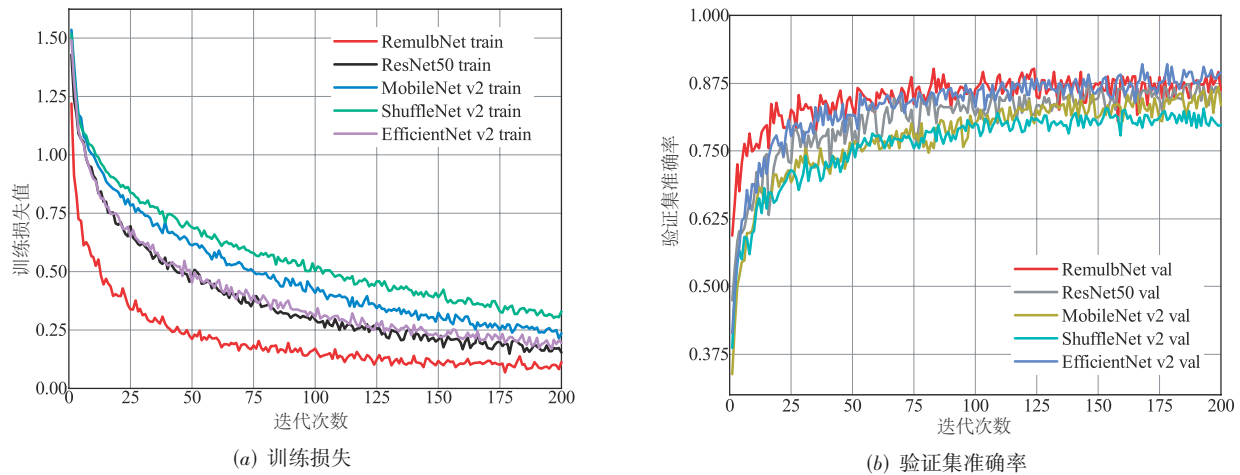


图5 网络模型对比

表4 不同数据集实验

数据集	网络	准确率/%	参数量/MB	体积/MB	训练时间/s
17分类 花卉	RemulbNet(PP)	98.7	7.50	21.08	1 183
	RemulbNet	95.3	7.50	21.08	1 175
	ResNet50	90.6	25.56	90.11	1 723
	MobileNet v2	94.0	3.51	8.80	1 465
	ShuffleNet v2	90.6	2.28	5.01	1 552
	EfficientNet v2	94.8	21.46	77.91	3 229
昆虫 (迭代:100)	RemulbNet(PP)	98.0	7.50	20.97	1 546
	RemulbNet	97.8	7.50	20.97	1 544
	ResNet50	97.0	25.56	90.00	1 873
	MobileNet v2	96.6	3.51	8.73	1 646
	ShuffleNet v2	95.6	2.28	4.96	1 777
	EfficientNet v2	97.5	21.46	77.84	3 904
车辆	RemulbNet(PP)	96.0	7.50	20.98	1 738
	RemulbNet	94.0	7.50	20.98	1 711
	ResNet50	94.8	25.56	90.02	2 678
	MobileNet v2	94.0	3.51	8.74	2 291
	ShuffleNet v2	87.5	2.28	4.97	2 495
	EfficientNet v2	95.5	21.46	77.85	5 214
CIFAR-10 (迭代:100)	RemulbNet	94.6	7.668	20.99	3 474
	ResNet50	94.6	25.55	90.03	4 848
	MobileNet v2	93.2	3.51	8.76	2 653
	ShuffleNet v2	91.5	2.28	4.99	3 733
	EfficientNet v2	92.2	21.46	77.87	7 776

达到90%以上的准确率,但RemulbNet网络的准确率最高,训练时间最短,相比其他网络的耗时缩短了约19.8%~63.6%。同样的,在昆虫图像识别结果中,RemulbNet与EfficientNet v2网络在精度上表现相当且显著高于其他网络。在参数量、训练耗时方面,RemulbNet优于EfficientNet v2和ResNet50网络,具有较高的计算

效率。在车辆识别数据集上,RemulbNet的准确率比ShuffleNet v2网络高出6.5%,而训练耗时仅为ShuffleNet v2的68.6%。对CIFAR-10数据集的实验结果表明,RemulbNet与ResNet50网络的准确率相当,但参数量仅为ResNet50的30%,训练耗时为ResNet50的28.3%,即在保证准确率不下降的条件下具有更高的计算效率。同时,对比EfficientNet v2、MobileNet v2和ShuffleNet v2网络,RemulbNet在准确率和训练耗时方面具有明显的优势。

通过以上实验可以看出,对于大多数图像数据集,RemulbNet在准确率方面都具有较好的性能,明显高于MobileNet v2和ShuffleNet v2网络;训练耗时明显低于其他网络,在数据量较大时,优势则更加明显;在参数量和网络体积方面,RemulbNet明显低于EfficientNet v2和ResNet50网络。

3.2.4 模型参数对RemulbNet性能的影响

最后,为了更进一步探究RemulbNet的性能,使用常见花卉数据集对学习率、激活函数以及不同的优化器对网络性能的影响进行了分析,实验结果如表5、表6和图7所示。

当使用ReLU、PReLU、ELU、Mish、Hardswish这些函数时,RemulbNet的准确率分别为90%、90.5%、90.7%、90.1%、91%,虽然有所起伏但对网络其他性能影响不大,Mish的识别准确率为91.0%,高于其他函数。为进一步分析激活函数的影响,绘出各激活函数的曲线如图6所示。其中, x 为自变量, y 为各函数的值。从图中可知,在正值域五种函数都比较类似,但是在负值域,ReLU输出为零,PReLU无下界且做线性输出,ELU有界但有较多的参数,Hardswish具有非线性但曲线不平滑,而Mish函数平滑、非线性输出且存在下界,Mish函数的这一特性有效提高了计算效率。

表5为不同学习率对模型的性能影响。当学习率逐渐增大时,准确率增加,训练时间变化不大。当 $lr=$

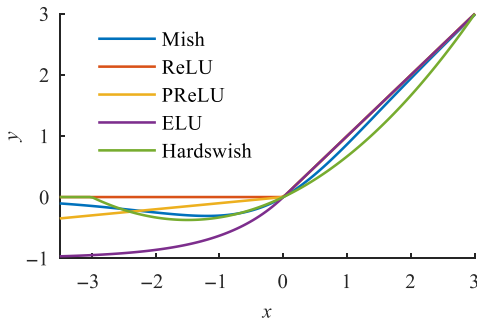
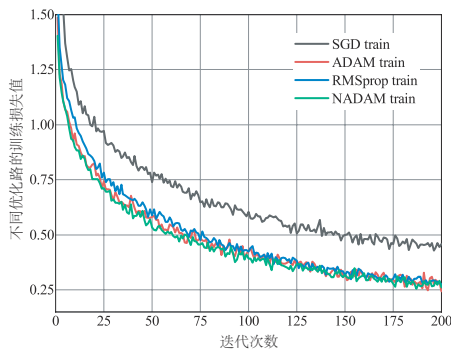


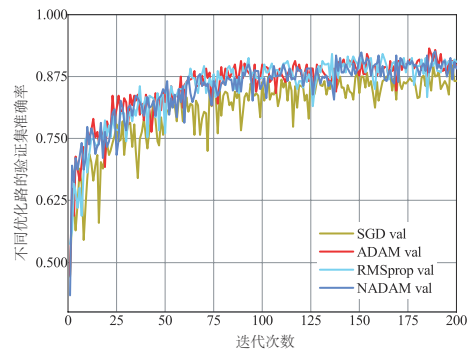
图6 激活函数示意图

0.001时, RemulbNet 识别准确率最高为 92.9%, 即使用较大的学习率可以提升网络的性能, 而且不会影响网络在训练上的耗时。

通过表6及图7可以看出, 在收敛速度方面, SGD 优化器的训练损失高且存在严重的振荡现象; RMSprop 在训练中收敛速度和精度与 Adam 相似, 但验证集上有振荡;



(a) 训练损失



(b) 验证集准确率

图7 优化器对比图

表5 学习率对性能的影响

学习率	准确率/%	训练时间/s
0.000 1	91.6	1 536
0.000 5	92.0	1 567
0.001	92.9	1 587

表6 优化器对比

优化器类型	准确率/%	训练时间/s
SGD	88.4	1 764
RMSprop	89.0	1 553
Adam	91.6	1 536
NAdam	91.3	2 007

Adam 系列优化器具有准确率高, 稳定的优点, 其验证准确率最高可达 93.2%, 而训练损失仅有 0.224; NAdam 的训练时间有所增加. 在综合考虑准确率与耗时的情况下, Adam 性能良好, 超参数设置较少、振荡小、适用性较好。

4 总结

本文提出一种多分支轻量化网络 RemulbNet. 该模型通过采用多分支结构、变体的深度可分离卷积和替换激活函数等策略, 在提高识别准确率的基础上减少了模型参数量、运算量以及训练时间. 通过使用多种数据集进行训练与测试, 实验结果表明相比 ResNet50、EfficientNet v2 以及经典的轻量化网络, 本模型具有优良的性能. RemulbNet 在诸多方面的性能都有提升, 但池化的加入导致了部分信息缺失, 准确率的提升有限, 且体积和训练时长方面也有待压缩. 希望在未来的研究中能够解决以上问题, 进一步提升网络性能。

参考文献

[1] FUKUSHIMA K. Neocognitron: A self-organizing neural network model for a mechanism of pattern recognition unaffected by shift in position[J]. Biological Cybernetics,

1980, 36(4): 193-202.
 [2] LECUN Y, BOTTOU L. Gradient-based learning applied to document recognition[J]. Proceedings of the IEEE, 1998, 86(11): 2278-2324.
 [3] KRIZHEVSKY A, SUTSKEVER I, HINTON G E. ImageNet classification with deep convolutional neural networks[C]//Proceedings of the 25th International Conference on Neural Information Processing Systems-Volume 1. Lake Tahoe: MIT Press, 2012: 1106-1114.
 [4] TAJBAKSH N, SHIN J Y, GURUDU S R, et al. Convolutional neural networks for medical image analysis: Full training or fine tuning?[J]. IEEE Transactions on Medical Imaging, 2016, 35(5): 1299-1312.
 [5] SZEGEDY C, LIU W, JIA Y, et al. Going deeper with convolutions[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Boston: IEEE, 2015: 1-9.
 [6] SIMONYAN K, ZISSERMAN A. Very deep convolution-

- al networks for large-scale image recognition[EB/OL]. (2014-09-04)[2022-08-05]. <https://arxiv.org/abs/1409.1556>.
- [7] 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. Piscataway: IEEE, 2016: 770-778.
- [8] IOFFE S, SZEGEDY C. Batch normalization: Accelerating deep network training by reducing internal covariate shift[C]//International Conference on Machine Learning. New York: ACM, 2015: 448-456.
- [9] HOWARD A G, ZHU M, CHEN B, et al. Mobilenets: Efficient convolutional neural networks for mobile vision applications[EB/OL]. (2017-04-17)[2022-08-05]. <https://arxiv.org/abs/1704.04861>.
- [10] ZHANG X, ZHOU X, LIN M, et al. Shufflenet: An extremely efficient convolutional neural network for mobile devices[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Salt Lake City, UT, USA: IEEE, 2018: 6848-6856.
- [11] GOMES R, ROZARIO P, ADHIKARI N. Deep learning optimization in remote sensing image segmentation using dilated convolutions and ShuffleNet[C]//2021 IEEE International Conference on Electro Information Technology (EIT). Piscataway: IEEE, 2021: 244-249.
- [12] TAN M, LE Q. Efficientnet: Rethinking model scaling for convolutional neural networks[C]//International Conference on Machine Learning. Long Beach, California, USA: ACM, 2019: 6105-6114.
- [13] HU J, SHEN L, SUN G. Squeeze-and-excitation networks [C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE, 2018: 7132-7141.
- [14] BOCHKOVSKIY A, WANG C Y, LIAO H Y M. Yolo4: Optimal speed and accuracy of object detection[EB/OL]. (2020-04-23) [2022-08-05]. <https://arxiv.org/abs/2004.10934>.
- [15] SZEGEDY C, IOFFE S, VANHOUCHE V, et al. Inception-v4, inception-ResNet and the impact of residual connections on learning[C]//Thirty-first AAAI Conference on Artificial Intelligence. San Francisco: AAAI Press, 2017: 4278-4284.
- [16] KAISER L, GOMEZ A N, CHOLLET F. Depthwise separable convolutions for neural machine translation[EB/OL]. (2017-06-09) [2022-08-05]. <https://arxiv.org/abs/1706.03059>.
- [17] 伍邦谷, 张苏林, 石红, 等. 基于多分支结构的不确定性局部通道注意力机制[J]. 电子学报, 2022, 50(2): 374-382.
- WU B G, ZHANG S L, SHI H, et al. Multi-branch structure based local channel attention with uncertainty[J]. Acta Electronica Sinica, 2022, 50(2): 374-382. (in Chinese)
- [18] TAFAZZOLI F, FRIGUI H, NISHIYAMA K. A large and diverse dataset for improved vehicle make and model recognition[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE, 2017: 874-881.
- [19] TAN M, LE Q. Efficientnetv2: Smaller models and faster training[EB/OL]. (2021-04-01)[2022-08-05]. <https://arxiv.org/abs/2104.00298>.
- [20] SANDLER M, HOWARD A, ZHU M L, et al. MobileNetV2: Inverted residuals and linear bottlenecks[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE, 2018: 4510-4520.
- [21] 孟球, 徐磊, 郭嘉阳. 一种基于改进的MobileNetV2网络语义分割算法[J]. 电子学报, 2020, 48(9): 1769-1776.
- MENG L, XU L, GUO J Y. Semantic segmentation algorithm based on improved MobileNetV2[J]. Acta Electronica Sinica, 2020, 48(9): 1769-1776. (in Chinese)
- [22] MA N, ZHANG X, ZHENG H T, et al. Shufflenet v2: Practical guidelines for efficient cnn architecture design [C]//European Conference on Computer Vision. Munich, Germany: Springer, 2018: 122-138.

作者简介



苏田田 男, 1996年出生于甘肃省. 2019年在西北师范大学获得学士学位. 现为陕西师范大学研究生, 主要研究方向为计算机视觉和深度学习.

E-mail: sutcarl@hotmail.com



王慧敏 女, 1997年出生于陕西省. 2019年在陕西师范大学获得学士学位. 现为陕西师范大学研究生, 主要研究方向为图像分类.

E-mail: 1164994333@qq.com

张小凤(通讯作者) 女, 1971年出生于陕西省. 现为陕西师范大学物理学与信息技术学院教授, 博士生导师. 主要研究方向为超声工程、智能信号处理与检测.

E-mail: xiaofengzhang71@snnu.edu.cn