

基于改进ShuffleNetV2的图像分类算法

张浩

长江大学电子信息与电气工程学院, 湖北 荆州

收稿日期: 2023年11月11日; 录用日期: 2023年12月12日; 发布日期: 2023年12月19日

摘要

对ShuffleNetV2的轻量级分类网络进行了研究, 提出了一种改进图像分类方法, 来提高图像分类的准确性。相比于传统的卷积神经网络, 可以提高模型的分类准确度和速度。首先, 改进的模型引入 6×6 的深度分离卷积来替代原来 3×3 的深度分离卷积, 来增加分类的准确度, 其次, 改进的模型应用了扁平层对数据进行降尺度和增加全连接层来提升学习能力。此外, 改进的模型还对比了其他传统CNN模型的图像分类表现, 选取cifar10/100数据集进行实验, 实验结果表明改进的模型在准确度上提高了2%以上。

关键词

ShuffleNetV2, 图像分类, 深度分离卷积, Cifar数据集

Image Classification Algorithm Based on Improved ShuffleNetV2

Hao Zhang

School of Electronic Information and Electrical Engineering, Yangtze University, Jingzhou Hubei

Received: Nov. 11th, 2023; accepted: Dec. 12th, 2023; published: Dec. 19th, 2023

Abstract

This study investigates the lightweight classification network of ShuffleNetV2 and proposes an improved method for image classification to enhance accuracy. In comparison to traditional convolutional neural networks, the improved model achieves higher classification accuracy and faster processing speed. Firstly, the enhanced model introduces 6×6 depthwise separable convolution to replace the original 3×3 depthwise separable convolution, thereby increasing classification accuracy. Secondly, the improved model incorporates a flatten layer for data dimensionality reduction and adds a fully connected layer to enhance learning capabilities. Additionally, the improved model compares the image classification performance with other traditional CNN models,

conducting experiments on the cifar10/100 datasets. The experimental results demonstrate an increase in accuracy of over 2% for the improved model.

Keywords

ShuffleNetV2, Image Classification, Depthwise Separable Convolution, Cifar Datasets

Copyright © 2023 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

当涉及图像分类任务时，深度学习的快速发展已经为我们提供了一系列强大的神经网络体系结构，其中之一便是 ShuffleNetV2。在图像数据迅速增加的背景下，尤其是在移动设备、物联网和医疗图像等领域，对高效且准确的图像分类算法的需求变得日益紧迫。在这一情况下，ShuffleNetV2 以其卓越的性能和高效的特点引起了广泛的关注。作为轻量级深度神经网络的代表，ShuffleNetV2 充分发挥了在计算和存储效率方面的优势。通过引入组卷积和通道重排机制，ShuffleNetV2 有效减少了参数数量和计算负担，与传统卷积神经网络相比，更具优势。这使得在不牺牲模型准确性的前提下，ShuffleNetV2 能够在移动设备等资源受限的环境中高效运行[1]。

ShuffleNet 是由微软亚洲研究院的研究团队于 2017 年提出的，其核心创新是引入了通道重排操作，以解决低复杂度下通道信息传递的问题。在 ShuffleNet 的基础上，微软亚洲研究院的研究团队于 2018 年推出了 ShuffleNetV2，ShuffleNetV2 在保留通道重排操作的基础上，使用更小的卷积核和更多的层次来提高特征提取能力，同时通过逐点卷积和深度可分离卷积等操作来减少计算量。

本论文的目标在于探究 ShuffleNetV2 模型在图像分类任务中的应用。我们将深入研究 ShuffleNetV2 的核心机制和创新点，剖析其在图像分类领域的优越性能。我们将详细介绍 ShuffleNetV2 的网络结构，包括深度可分离卷积、通道重排等关键设计，并展示如何通过替换深度分离卷积来提升其在图像分类任务中的性能，并加入扁平层使数据进行缩小。通过一系列的实验和与其他经典 CNN 模型的对比，我们将充分验证改进模型在 cifar 数据集上的表现[2] [3]。

2. 网络结构

ShuffleNetV2 是一种轻量级的卷积神经网络架构，旨在实现高效的模型设计，以适应移动设备和嵌入式系统等资源有限的环境。ShuffleNetV2 的核心思想在于引入了深度可分离卷积和通道重排(Channel Shuffle)机制，以减少参数数量和计算复杂度，同时保持模型性能。

2.1. 输入层

ShuffleNetV2 的输入层通常指的是网络模型的第一层，它用来接收输入图像并进行初始的特征提取。ShuffleNetV2 的输入层包括以下几个重要组成部分：卷积层，用来对输入图像进行卷积操作，这个卷积层也包含了一些超参数，如卷积核的大小、步幅等，可以通过改变这些参数来初步决定对图像的特征提取。批量归一化层，通常在卷积层后会添加一个批量归一化层，作用是加速网络的训练和提高模型的泛化能力。激活函数，在经过上面两层之后，会用激活函数来引入非线性特性，通常使用的激活函数有 ReLU (Rectified Linear Unit)、Leaky ReLU 等，具体通过设计的模型来进行选择。

ShuffleNetV2 的输入层通常会根据具体的网络配置而有所不同,但总体思路是通过这些部分来处理输入图像,提取特征,然后将特征传递给网络的后续层进行更深层次的特征学习和分类。网络的输入层的设计对模型的性能和效率都有重要影响,因此需要根据具体任务和条件进行合理的设计和调整[4]。

2.2. 深度可分离卷积

深度可分离卷积是卷积神经网络中常用的一种方法,其作用是减少模型的参数量和计算量,同时保持良好的性能,深度可分离卷积有两个步骤:深度卷积和逐点卷积。

深度卷积的作用是首先对输入特征图的每个通道进行独立卷积操作,这意味着每个输入通道都有一个对应的卷积核,用于提取该通道的特征,这一步骤仅在通道内进行卷积,不跨通道。它还可以减少计算量和减少参数,这使得其在有限的条件中也能表现的不错。

对于输入张量(I),每个通道执行独立深度卷积,表示为

$$D(I) = [I_1 \times K_1, I_2 \times K_2, \dots, I_n \times K_n]$$

这里的 I_1, I_2, \dots, I_n 是输入通道, K_1, K_2, \dots, K_n 是对应的深度卷积核。

逐点卷积的作用是深度可分离卷积的第二个部分,它在每个位置上将深度卷积的输出特征图与一个小的卷积核相卷积。逐点卷积引入了非线性特性,通过激活函数对每个点进行处理。

深度卷积的输出张量与逐点卷积核相乘并相加,表示为

$$P(D(I)) = \sum (D(I) \times W_i)$$

这里的 W_i 是逐点卷积核,对应输出通道。

深度可分离卷积的这两个部分结合在一起,可以在几乎不损失性能的情况下,显著减少模型的参数量和计算量。ShuffleNetV2 将深度可分离卷积作为其主要卷积部分,用于替代传统的卷积层,从而实现了轻量级和高效的网络结构,这使得其在条件有限的环境中的理想选择[5]。

2.3. 通道重排

通道重排用于改善模型的性能,减少计算复杂度,并加速模型的训练,其主要思想是将输入特征图中的通道重新排列成多个小的子组,用这些子组进行并行处理。

通道重排的步骤是先进行分组,这些小组有相同数量的通道,然后对每个小组内的通道进行重排,通常会将通道交叉排列,确保每个小组都包含输入特征图的不同信息。重排后,每个小组进行并行处理,互不干扰,这使得资源得到有效的利用,提高计算的效率。在小组并行处理后,可选择不同的小组间进行信息交流,这种组间交流可以提升特征的表达能力。

通道重排是 ShuffleNetV2 中的一个关键创新,使得其在计算复杂程度和速度方面优势较为明显,让其在资源有限的环境中成为不错的选择[6] [7] [8]。

2.4. ShuffleNetV2 单元

ShuffleNetV2 单元是由 ShuffleNetV1 经过改进后得到的,ShuffleNetV2 单元如图 1 所示。

由图 1(a)可以看出,输入的特征图一分为二,且左右各有总量的 1/2,右边分支包括 3 个连续的卷积,步长都是 1,左边分支不做卷积,输入和输出的通道相同。 1×1 卷积属于普通卷积, 3×3 卷积属于深度分离卷积,两个分支的输出通过 concat 方式合并通道,紧接着把 concat 处理的结果通过通道重排结合在一起,以便在组之间进行信息的交换和融合,确保跨组的信息进行传递。图 1(b)中,输入的特征图并不会再对通道进行划分,而是直接进入左右两个分支,每个分支都有步长为 2 的 3×3 深度分离卷积,最后

通过 `concat` 方式合并通道，这种方式的好处是降低计算量，通道的数量翻倍，还能够加强提取特征的能力，最后经过通道重排，使信息进行交换和融合。

ShuffleNetV2 单元的结构允许模型以更高效的方式进行特征学习，同时保持了模型的表达能力[5] [9]。

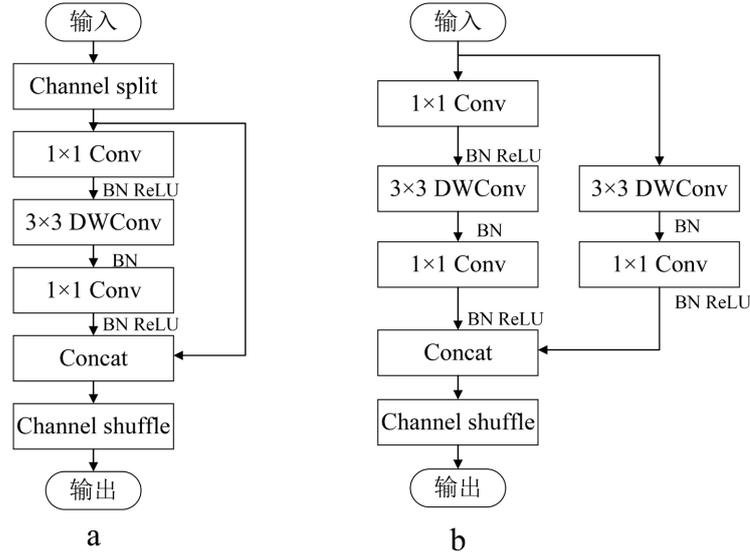


Figure 1. ShuffleNetV2 unit
图 1. ShuffleNetV2 单元

3. 改进的 ShuffleNetV2 模型

3.1. 不同改进模型的对比

可分离卷积的计算在原始模型中的占比较小，所以替换成 6×6 的可分离卷积后，可以捕获更多的特征，这有助于提高模型对图像的感知能力。增加卷积核的尺寸不仅可以改善通道重排的效果，更有助于维护特征之间的关联性。此外，采用更大的卷积核并没有显著增加计算成本，却能提升模型的准确率，使其表现优于原始模型。所以在不违反轻量化网络规则的前提下，替换可分离卷积是效果比较好的方法。

如果在原始模型中添加注意力机制，在模型的参数量上比替换可分离卷积的多，但准确率确比替换深度可分离卷积差不多，所以在保证改进模型的计算量以及运行时长等方面，选择替换可分离卷积作为本文实验的主要研究方向[10]。

3.2. 扁平层和全连接层

扁平层的主要作用是将多维的输入数据转换成一维向量，在深度卷积神经网络等架构中，卷积和池化层减小了特征图的空间维度，但保留了通道维度，而扁平层的作用就是将最后一个卷积层或池化层的输出扁平化，以便输入到全连接层。

全连接层通常是卷积神经网络的最后一层，作用是将扁平化的特征向量映射到输出类别的概率，全连接层上的每个神经元与上一层的所有神经元相连接，这表明全连接层的每个神经元都对上一层所有特征有权重，这增加了模型的学习能力，使模型更加全面和准确。

全连接层对于前一次输出向量： $X = [x_1, x_2, \dots, x_n]$

前一层有 n 个神经元，全连接层有 m 个神经元，权重矩阵为 W ，偏置向量为 b ，则全连接层的计算可以表示为：

$$Y = WX + b$$

其中, $Y = [y_1, y_2, \dots, y_m]$ 是全连接层的输出向量[1]。

3.3. 改进模型图像分类的流程

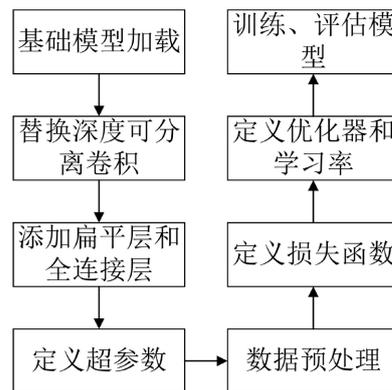


Figure 2. Workflow of model image classification
图 2. 模型图像分类的流程

如图 2 所示, 首先进行原始模型的加载, 在原始的模型上进行改进, 进行替换深度分离卷积和添加扁平层和全连接层。定义随机种子和超参数, 以确保结果的重复性, 创建数据预处理和加载器, 并将数据集加载到训练和测试数据加载器中。用交叉熵损失函数来计算预测和实际标签之间的误差, 来减少模型误分类的概率, 优化器用来更新神经网络的权重, 以减小损失函数的数值, 使用随机梯度下降优化器, 用来更新神经网络的学习率和定义训练动量。

进行模型训练, 将标签和数据移到设备上, 执行前向传播, 计算模型的输出和损失, 执行反向传播并执行优化步骤, 随后卷积层的输出被展平成一维向量, 该向量通过全连接层, 并在测试集上评估模型, 计算模型在测试集的准确率。

4. 实验与结果分析

4.1. 实验环境和参数设置

本实验采用的操作系统为 Windows 系统, CUDA 版本为 11.1, Python 版本为 3.8, Pytorch 版本为 2.0, GPU 为 NVIDIA GTX1650。

在本实验的训练过程中, 将 SGD 作为优化器, 且带训练动量, 并且设置为 0.9, 训练权重的衰减设置为 $5 \times e^{-4}$, 使用 GPU 进行训练。训练周期为 100 个 epoch, 前 70 个周期的学习率是 0.01, 70 到 90 周期的学习率是 0.001, 90 到 100 周期的学习率是 0.0001。为增加实验数据的准确性, 对数据集训练三次, 取平均值。

4.2. 实验数据

为了评估改进 ShuffleNetV2 模型在预测数据上的效果, 本文采用 cifar10 和 cifar100 图像分类数据集进行实验, cifar10 数据集里面有 10 个类别, 共有 60,000 张彩色图像, 图像的尺寸是 32×32 , 训练集有 50,000 张图像, 每类有 5000 张; 测试集有 10,000 张图像, 每类有 1000 张。cifar100 数据集里面有 100 个类别, 共有 60,000 张彩色图像, 图像的尺寸也是 32×32 , 有 50,000 张训练集和 10,000 张测试集, 每个类别包含 600 张图像。

4.3. 实验结果分析

第一个实验对比模型之间的参数数量，以 ShuffleNetV2 标准模型在 cifar 数据集上的参数量为基准，通过对比改进的 ShuffleNetV2 模型参数数量来进行对比实验，结果如表 1 所示，通过参数数量的对比，改进的 ShuffleNetV2 模型的参数量只提高了约 4%。

Table 1. Comparison of parameter quantities between improved ShuffleNetV2 model and original model

表 1. 改进 ShuffleNetV2 模型和原始模型的参数量对比

模型	参数量
ShuffleNetV2	2,278,604
ShuffleNetV2-s	2,380,648

在第二个实验中，我们通过对不同模型在 cifar10 和 cifar100 数据集上的图像分类准确率，以 ShuffleNetV2 原始模型为基准进行改进。通过比较改进后的 ShuffleNetV2 模型与传统卷积神经网络的准确率，我们使实验条件更加全面，以全面评估模型性能的提升效果。结果如表 2 所示，从表中可以看出，改进的 ShuffleNetV2 模型在 cifar10 数据集上的准确率提高了约 2%，在 cifar100 数据集上的准确率提高了约 5%；相较于传统的 CNN 模型，改进的 ShuffleNetV2 模型准确率得到了提升且运行的时长也缩短了。

Table 2. Comparison of accuracy among various models on cifar10/100 datasets

表 2. 各模型在 cifar10/100 数据集上准确率对比

模型	Cifar10 准确率(%)	Cifar100 准确率(%)
ShuffleNetV2	88.42	70.82
ShuffleNetV2-s	90.67	76.59
传统 CNN	84.47	53.69

对于原始的 ShuffleNetV2 模型准确率较低、学习能力较差等方面的不足，通过替换其深度卷积层和加入全连接层使其原来不足的方面得到了改进，增加网络提取特征的能力和学习能力。

5. 结论

为了克服原始 ShuffleNetV2 模型准确率较低的问题，本研究采用了替代深度分离卷积、添加扁平层和全连接层的策略，提出了一种改进 ShuffleNetV2 模型用于图像分类的方法。改进后的模型在参数量和准确率之间取得了平衡，仅略微增加了参数量，却显著提升了准确率。通过在 cifar10 和 cifar100 数据集上进行图像分类实验证实了这一改进的有效性。相较于传统卷积神经网络，不仅准确率得到显著提升，而且在运行时间上也表现出明显的优势。改进 ShuffleNetV2 模型引入的扁平层和全连接层还增强了模型的学习能力，从而提升了模型的综合性能。未来希望能在不增加大量计算的情况下，使得模型图像分类的精度更近一步，让轻量级 ShuffleNetV2 模型可以应用到更多条件有限的环境中。

参考文献

- [1] Zhang, J., Li, C. and Guo, B. (2023) Improved ShuffleNetV2 for Garbage Image Classification. *Advances in Intelligent Information Hiding and Multimedia Signal Processing*, **341**, 291-299. https://doi.org/10.1007/978-981-99-0605-5_28
- [2] Liu, X., Wu, Z.Z., Wu, Z.J., Zou, L., Xu, L.X. and Wang, X.F. (2020) Lightweight Neural Network Based Garbage Image Classification Using a Deep Mutual Learning. *Parallel Architectures, Algorithms and Programming*, **1362**, 212-223.

- [3] Zhang, X., Zhou, X., Lin, M. and Sun, J. (2017) Shufflenet: An Extremely Efficient Convolutional Neural Network for Mobile Devices. *Computer Vision and Pattern Recognition*.
- [4] Gu, Y.Q. and Ge, B. (2021) Research on Lightweight Convolutional Neural Network in Garbage Classification. *IOP Conference Series: Earth and Environmental Science*, **781**, Article ID: 032011.
- [5] 袁硕, 刘玉敏, 安志伟, 王硕昌, 魏海军. 基于改进 ShuffleNetV2 网络的岩石图像识别[J]. 吉林大学学报(信息科学版), 2023, 41(3): 450-458.
- [6] 周楠, 欧阳鑫玉. 卷积神经网络的发展[J]. 辽宁科技大学学报, 2021, 44(5): 349-356.
- [7] 严春满, 王铖. 卷积神经网络模型发展及应用[J]. 计算机科学与探索, 2021, 15(1): 27-46.
- [8] 刘丽, 赵凌君, 郭承玉, 等. 图像纹理分类方法研究进展和展望[J]. 自动化学报, 2018, 44(4): 584-607.
- [9] 赵毓, 任艺平, 朴欣茹, 郑丹阳, 李东明. 基于改进 ShuffleNet V2 的轻量级防风药材道地性识别[J]. 智慧农业(中英文), 2023, 5(2): 104-114.
- [10] 蒋博文. 基于改进 ResNet 模型的图像分类方法[J]. 现代信息科技, 2022, 6(12): 83-85.