

基于深度学习的中草药图像分类算法的研究与设计

杨瑞林

(北京邮电大学计算机学院(国家示范性软件学院), 北京 100876)

摘要: 中草药图像的分类识别在医学、生活等方面有着重要的研究和应用价值。许多中草药有着相似的类间特征, 给分类任务带来了干扰。为了准确、快速的对中草药图像进行分类, 本文基于深度学习方法设计并实现了一种中草药图像分类的方法。本文基于轻量级分类网络 ShuffleNet V2, 使用数据增强方法扩展数据集, 增加分类网络的鲁棒性; 使用 ACNet 单元替代 ShuffleNet V2 普通单元中的传统卷积, 进一步强化了特征提取能力。本文方法在中草药数据集上进行实验, 发现准确率相对原网络都有提升, 说明方法改善是有效的。同时, 分类算法准确率在 94.89%, 可以满足中草药图像分类的使用需求。

关键词: 计算机应用; 深度学习; 中草药图像分类; ShuffleNet V2; 数据增强

中图分类号: TP399

Research and design of Chinese herbal medicine image classification algorithm based on deep learning

Yang Ruilin

(School of Computer Science (National Pilot Software Engineering School), Beijing University of Posts and Telecommunications, Beijing 100876)

Abstract: The classification of Chinese herbal medicine image has important research and application value in medicine and life. Many Chinese herbs have similar interclass characteristics, which presents a barrier to the task of classification. In order to classify Chinese herbal images accurately and quickly, this paper designed and implemented a classification method of Chinese herbal images based on deep learning method. The method in this paper is based on ShuffleNet V2. The data enhancement method is used to expand the dataset and increase the robustness of the classification network. ACNet unit is used to replace the traditional convolution, which further strengthens the ability of feature extraction. The method in this paper was tested on the dataset of Chinese herbal medicine. The accuracy was improved compared with the original network, indicating that the improved method is effective. At the same time, the accuracy of the classification algorithm is 94.89%, which can meet the needs of Chinese herbal medicine image classification.

Keywords: Computer application; Deep learning; Chinese herbal medicine image classification; ShuffleNet V2; Data enhancement

0 引言

随着现代医学的迅速发展, 中医治疗再次引起了人们的重视。特别是在新冠疫情防控中, 中医成为一种重要的辅助治疗方式。医学领域也在逐步探索中医在未来疾病防控中发挥更建设性的作用^[1]。中草药是中医治疗的必须药物, 在使用时需要准确的了解其种类和药性。但学习中草药需要长时间的积累, 能分辨一些稀少的、难以辨认的中草药的老一辈人逐渐减少, 在此背景下, 对中草药图像进行准确的分类识别是很有必要的。在对中草药拍摄的图像分类过程中, 存在着一些难点。一是中草药种类数量繁杂, 许多不同种类药物有着相似的表面特

作者简介: 杨瑞林 (1997-), 男, 硕士研究生, 主要研究方向: 计算机视觉. E-mail: rlyang@bupt.edu.cn

征, 在分类识别时容易混淆。二是中草药植物都有不同的生长阶段, 但在某些阶段植物特征大都相似, 比如在生长阶段绿色特征较多, 在采摘后处理之后以干枯的灰黄色特征较多, 这些都给分类识别带来了很大的困难。因此, 研究中草药的分类识别对人们生产生活具有重大意义。

深度学习的快速发展, 在处理图像领域的各种任务上都有了明显突破, 为解决上述问题提供了理论基础。本文为了解决中草药图像的分类识别问题, 在现有的轻量级分类识别网络 ShuffleNet V2^[2]的基础上, 进行了网络调整和数据增强, 进一步提高了网络在中草药分类中的性能, 基本满足了用户对中草药图像分类的需求。

1 相关工作

图像分类任务是计算机视觉中的基本任务之一。图像分类主要分为两个阶段, 特征提取阶段和分类阶段。前一阶段主要是对图像的特征进行描述, 形成特征向量; 后一阶段主要是依据提取的特征进行分类。图像分类算法按照出现的时间可以分为基于传统机器学习算法的图像分类和基于深度神经网络的图像分类。

基于机器学习的图像分类算法主要包括 K 近邻、决策树、支持向量机等。第一阶段使用图像描述算子进行特征提取, 如纹理、颜色、边缘特征; 第二阶段使用分类器进行分类。K 近邻分类是依据与预测图像特征最接近的 K 个训练图像特征对应类别标定来判定的, 主要通过调整 K 的大小和 K 个结果的不同权值来实现。决策树是将特征属性与类别的关系通过树形数据结构进行整理对应, 最后根据预测图像特征在特征树上进行解析, 得到预测的类别。支持向量机是将图像特征映射到高维空间中, 在高维空间中拟合出各类别特征的边界线。

基于深度学习的分类网络近年来有很多, 如 AlexNet、VGG、ResNet、DenseNet、MobileNet 系列、ShuffleNet 系列。AlexNet 是早期的图像分类网络, 在医学图像的分类任务中已有应用^[3]。AlexNet 在网络结构中引入了非线性单元, 并使用数据增强、dropout 等技巧减轻过拟合。VGG 系列探索了网络深度对网络性能的影响, 在医学图像分类中也表现较好^[4]。虽然近年来出现了许多新的网络结构, 但 VGG 仍然常被用于提取图像特征。ResNet 残差网络^[5]解决了深度网络深度太多产生的梯度爆炸和梯度消散问题。但随着网络深度的越来越深, 网络越来越复杂, 需要的计算时间也越来越多, 无法满足低延迟等应用对网络快速响应的需求。MobileNet 网络^[6]解决了这个问题, 进行了模型压缩。使用了深度可分离卷积代替传统的卷积, 减少了参数量。ShuffleNet^[7]在 MobileNet 的基础上, 对冗余的 1*1 卷积又进行了改进; 进行了通道打乱和逐点群卷积操作, 进一步压缩了模型, 提高了分类精度。此后的 MobileNet V2^[8]和 ShuffleNet V2 有对 V1 版本网络进行了微调, 实现了更轻量级的网络。整个深度学习分类网络的发展历程都在精度和速度两个方面进行着调整, 新的轻量级网络已在保证精度的前提下, 运算速度更快、访存开销更小。

2 算法设计与实现

2.1 问题描述

算法要解决的问题是根据输入的 RGB 中草药图片，经过一些列的网络计算，最终预测出中草药的类别。



图 1 中草药图片示例

Fig. 1 Examples of pictures of Chinese herbs

几种中草药示例如图 1 所示，从左向右依次是金银花、槐花、百合。从图中可看出，一些中草药具有相似的颜色、纹理特征，如金银花和槐花，都是棕、褐、黄色的，且都有晒干后的植物纹理特征，给分类识别带来了干扰，在分类算法中需要能解决这些相似特征带来的影响。

2.2 算法模型设计

本文在轻量级分类网络 ShuffleNet V2 的基础上，使用数据增强扩展了数据的多样特征，使用 ACNet 单元替换 ShuffleNet V2 基本单元的普通卷积，ACNet 在卷积过程中引入了横向和纵向的卷积，强化了对图片的特征提取能力。

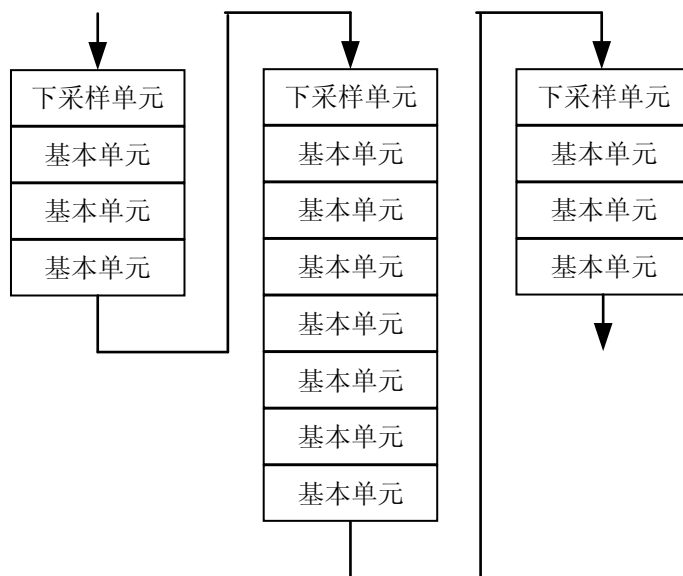


图 2ShuffleNet V2 的组成单元示意图

Fig. 2A constituent unit example of shufflenet V2

ShuffleNet V2 经过实验论证，发现了可进一步轻量化的几条准则。卷积层的通道数量

80

85

90

95 设置上,当输入通道数和输出通道数相同时,内存访问损耗时间最短。网络中的分组操作过多也会增加模型运行时间,因此需要减少分组数量。同时,网络中的分支数量越少,模型运行速度越快,因此需要降低网络中的分支数。此外,元素级操作也会降低每秒的浮点运行操作数,所以网络中尽量减少元素级操作。ShuffleNet V2 组成模块如图 2 所示,每一部分由一个下采样单元和多个基本单元构成。基本单元和下采样单元内部的通道特征打乱是为了在不增加参数的情况下,使通道充分融合。在其他参数设置上,ShuffleNet V2 也尽量满足上述四条原则,使得网络运行速度进一步加快。

100 本文使用了数据增强扩展数据量。使用的方法包括光线、模糊、裁剪、变换系列。光线变化包括亮度、对比度、饱和度、明度,模糊主要包括高斯模糊和椒盐模糊,裁剪幅度较小,裁剪上限为 20%,变换主要包括镜像、缩放、旋转三大类。每张图片在数据增强时,随机从上面的方法中挑出 3~4 种进行数据增强。由于增强时人手动设置的参数无法完全合理估计,因此先将数据增强方法参数阈值设置的较大,最后人工筛选特别不合理的图片,如太模糊的、太亮或太暗的,筛选出去。

105 ACNet 的核心是非对称卷积块 ACBlock^[9],ACBlock 构造了一组非对称卷积,在不增加额外计算量的同时可以提高准确率,本文使用 ACBlock 替代 ShuffleNet V2 中的普通卷积操作,改善网络的分类识别性能。ACBlock 基本原理如图 3 所示,考虑到卷积核十字位置包含的信息更多,增加使用横向和纵向的卷积核构成卷积组,在提取特征时能够提取到更多的信息熵。最后将非对称卷积块的输出合并再进行后续运算。

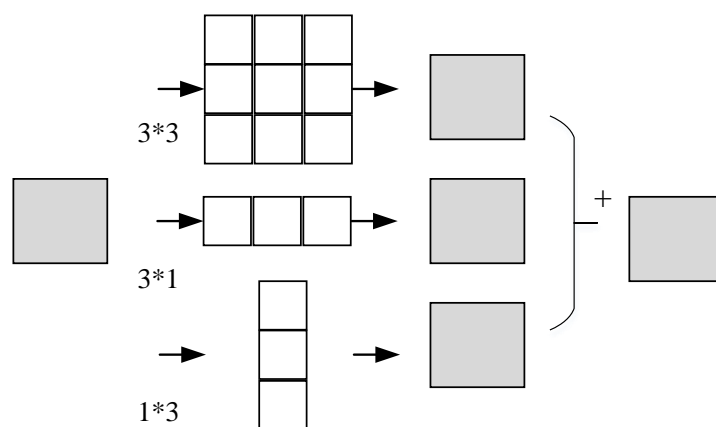


图 3ACBlock 基本构成示意图

115 Fig. 3Basic composition diagram of ACBlock

在本文中,使用 ACBlock 替代了 ShuffleNet V2 中各基本单元的卷积操作,在不增加额外计算量的同时,提升了网络的信息提取和信息表达能力。最终的中草药分类识别网络在保证检测速度的同时,具有更高的精度。

2.3 算法模型实现

120 图 4 是数据增强前后的对比图片。每一行第一张是原始图片,后面几张是使用不同数据增强技术之后的图片效果。上下两行分别是百合和金银花的增强效果。由图可见,图片增强后,原有的基本特征继续保留,同时也衍生出各种环境条件下的效果,在扩展数据量的同时也有助于增加模型的鲁棒性。

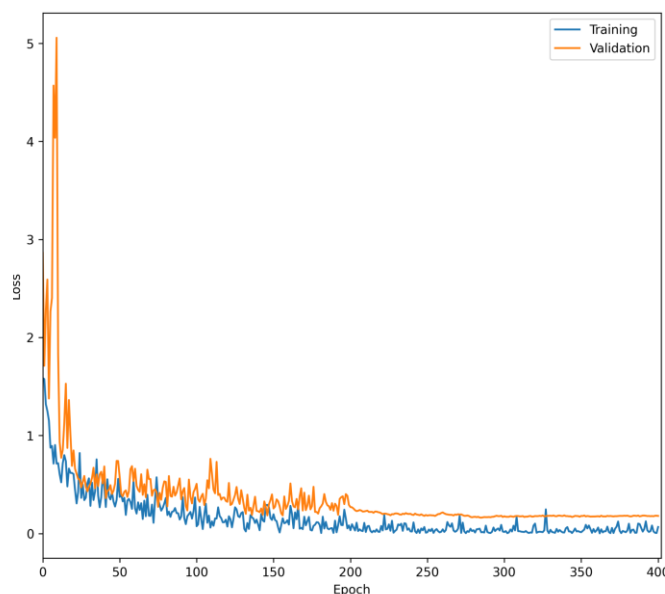


125

图 4 数据增强前后示例

Fig. 4 Examples before and after data enhancement

模型训练 loss 曲线如图 5 所示, 训练过程中逐渐降低学习率。在训练的早期过程中, loss 迅速的下降, 之后 loss 下降趋势变缓, 不断趋于收敛。



130

图 5 中草药分类模型的 loss 曲线图

Fig. 5 Loss curve of Chinese herbal medicine classification model

网络模型训练参数配置如表 1 所示。其中, groups 是 ShuffleNet V2 网络中执行 channel shuffle 操作时的分组数量。

135

表 1 模型训练参数设置

Tab. 1 Model training parameter setting

参数	参数值	参数	参数值
初始 learning rate	1e-3	训练 epoch	400
batch size	64	动量参数	0.9
groups	2	衰减因子	4e-5

3 实验结果与分析

为验证设计算法在中医药图像分类中的实际效果, 进行测试。本文中使用的数据集来自 AI studio 中草药数据集, 共包括 5 种中草药。其中, 百合 180 张, 党参 190 张, 枸杞 185

140 张，槐花 167 张，金银花 180 张。数据增强后，训练测试数据共 2489 张。将数据集图片按照 15%：85% 的比例划分测试集和训练集，在测试集上进行测试。

3.1 模型评估指标

对于图像分类问题，常用准确率和错误率进行评估。公式如下：

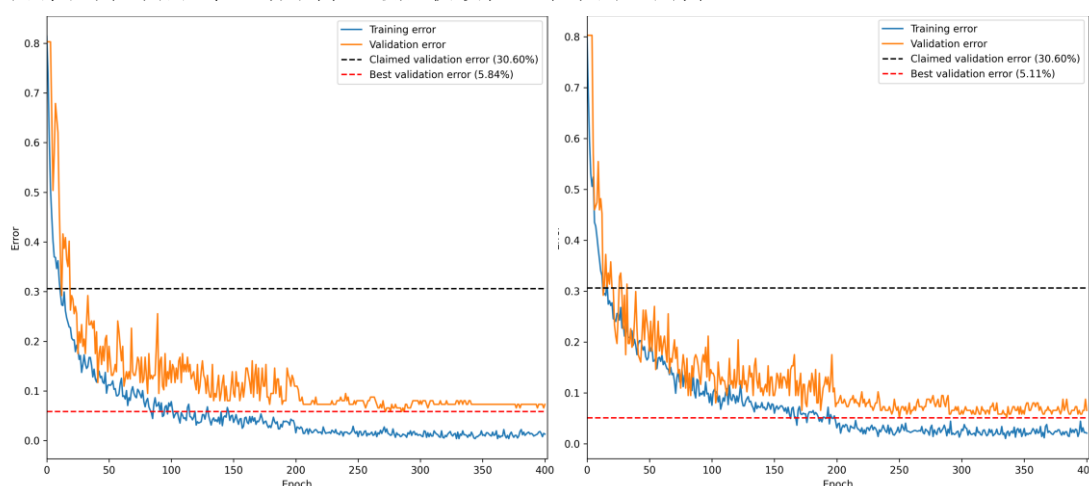
$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN}$$

$$Error_rate = \frac{FP + FN}{TP + TN + FP + FN}$$

145 其中，TP 是正确分类到正例的个数，TN 是正确划分到负例的个数，FP 是错误划分到正例的个数，FN 是错误划分到负例的个数。

3.2 实验结果分析

修改前后的 ShuffleNet V2 检测过程中的性能变化如图 6 所示，随着训练时间的增加，训练和测试错误率逐渐下降，最后收敛在一个小的区间内。



150 图 6 原 ShuffleNet V2 和修改后的 ShuffleNet V2 分类准确率对比

Fig. 6 Comparison of classification accuracy between original ShuffleNet V2 and modified ShuffleNet V2

最终测试的准确率如表 2 所示，最终设计的模型在中草药分类测试集上有更好的效果，分类效果达到 94.89%。

155 表 2 测试集实验结果对比

Tab. 2 Comparison of experimental results on test dataset

方法	Accuracy	Error rate
原 ShuffleNet V2	94.16%	5.84%
修改后的 ShuffleNet V2	94.89%	5.11%

4 总结

160 本文首先对中草药图像分类在医学、生活中的意义进行了阐述，之后对中草药图像分类时的难点进行了分析。在轻量级分类网络 ShuffleNet V2 的基础上，使用数据增强扩展数据集，使用 ACBlock 替换 ShuffleNet V2 基本单元卷积以提升图像特征提取能力。最终，本文

设计的模型在中草药数据集上进行测试，达到较高的准确率。但本文也仍然存在一些缺陷，比如用于实验的中草药种类偏少、实际图像数量少，后续将扩展数据量进行实验，在实际使用中发挥更大价值。

[参考文献] (References)

- 165 [1] 仝小林,朱向东,赵林华,李青伟,雷焯,宋斌,邵建柱,王强,李修洋,郑玉娇,樊瑗晗.加强我国新发突发传染病中医药应急防控体系建设的战略思考[J].中国科学院院刊,2020,35(09):1087-1095
- [2] Ma N, Zhang X, Zheng H T, et al. Shufflenet v2: Practical guidelines for efficient cnn architecture design[C]. Proceedings of the European conference on computer vision (ECCV). 2018: 116-131.
- 170 [3] Titoriya A, Sachdeva S. Breast cancer histopathology image classification using AlexNet[C]. 2019 4th International Conference on Information Systems and Computer Networks (ISCON). IEEE, 2019: 708-712.
- [4] Kaur T, Gandhi T K. Automated brain image classification based on VGG-16 and transfer learning[C]. 2019 International Conference on Information Technology (ICIT). IEEE, 2019: 94-98.
- [5] 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. 2016: 770-778.
- 175 [6] Howard A G, Zhu M, Chen B, et al. Mobilenets: Efficient convolutional neural networks for mobile vision applications[J]. arXiv preprint arXiv:1704.04861. 2017.
- [7] 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. 2018: 6848-6856.
- 180 [8] Sandler M, Howard A, Zhu M, et al. Mobilenetv2: Inverted residuals and linear bottlenecks[C]. Proceedings of the IEEE conference on computer vision and pattern recognition. 2018: 4510-4520.
- [9] Ding X, Guo Y, Ding G, et al. Acnet: Strengthening the kernel skeletons for powerful cnn via asymmetric convolution blocks[C]. Proceedings of the IEEE/CVF international conference on computer vision. 2019: 1911-1920.