

基于协调注意力的花生荚果品质分级

Classification of peanut quality based on coordinated attention

王春龙

蒋仲铭

鲍安红

WANG Chun-long JIANG Zhong-ming BAO An-hong

(西南大学工程技术学院, 重庆 400716)

(College of Engineering and Technology, Southwest University, Chongqing 400716, China)

摘要:目的:解决花生荚果品质分级过程中模型参数占用内存大、识别精度低、识别速度慢的问题。方法:提出一种基于深度学习和图像处理的花生荚果品质分级方法,在SqueezeNet模型的基础上,通过引入的协调注意力模块(Coordinate Attention)将得到的特征图分别编码成一对方向感知和位置敏感的注意力图,加强获取特征图中感兴趣区域信息的能力;采用梯度集中(Gradient Centralization)的策略改进优化算法;优化最末的fire层及卷积层的参数。提出优化模型CG-SqueezeNet,并应用于花生荚果品质分级。结果:与经典模型试验对比,CG-SqueezeNet模型在实际花生荚果图像数据库上的检测准确率为97.83%,参数内存仅为2.52 MB。结论:该方法适合部署在移动终端等嵌入式资源受限设备上,有助于实现对花生荚果品质的实时准确识别。

关键词:深度学习;品质分级;机器视觉;协调注意力;梯度集中;花生

Abstract: Objective: This study focuses on solving the problems of large memory consumption, low recognition accuracy and slow recognition speed in the classification process of peanut quality.

Methods: A method for classification of peanut quality based on deep learning and image processing was proposed. The Coordinate Attention module was firstly introduced to encode the obtained feature graph into a pair of direction-aware and position-sensitive attention graph, which improved the ability to obtain the information of the region of interest of the graph. Then, Gradient Centralization was used to improve the optimizer. By modifying the parameters of the last fire layer and the convolution layer. An

improved model, CG-SqueezeNet, was applied to peanut pod quality grading. **Results:** The classical convolutional network models VGG16, AlexNet, DenseNet121, ResNet50, Squeezenet were improved, and five different base classifier models were trained by transfer learning. By comparing with the classic model, it was found that the CG-SqueezeNet model could better learn the features of the region of interest in the image. The detection accuracy of the actual peanut pod image database was 97.83%, and the parameter memory was only 2.52 MB. **Conclusion:** The method is suitable for deployment on embedded resource-limited devices such as mobile terminals, which helps to realize real-time and accurate identification of peanut pod quality. **Keywords:** deep learning; quality grading; computer vision; Coordinate Attention; Gradient Centralization; peanut

花生筛选是花生销售前的重要工序,不仅可以提高商品花生的品质等级,实现质优价优,还可以实现花生产品的多渠道利用,减少花生产品浪费。人工筛选存在劳动强度大、效率低下等问题。而机械式^[1]的筛选装备根据花生荚果的大小特征进行分级,分级过程中很容易损坏花生荚果,分级的效果差。随着科技的发展,机器视觉技术被广泛应用于食品质量检测领域^[2]。如将机器视觉应用在对花生仁的外观品质检测上、利用掩模及亮度校正算法实现脐橙表面缺陷分割^[3]等。然而,传统的图像识别过程需要操作者利用经验手动提取特征,并且多次择优特征,操作过程繁琐且易受主观因素影响,难以得到准确稳定的缺陷特征。

基于卷积神经网络(convolutional neural network, CNN)的图像识别技术,没有过往复杂人工提取图像特征的过程,直接实现了自动化提取特征^[4-6]。随着迁移学习技术的发展,在大型数据集上训练学习得到的模型权重值和特征可以容易地迁移到小型数据集上,减少了模型的计算量,使得深度学习模型也适用于小数据集。薛

基金项目:国家自然科学基金项目(编号:4111900075);重庆市自然科学基金项目(编号:4312000227);重庆市研究生科研创新项目(编号:CYS211117)

作者简介:王春龙,男,西南大学在读硕士研究生。

通信作者:鲍安红(1969—),女,西南大学教授,博士。

E-mail: bah69@swu.edu.cn

收稿日期:2022-02-27 **改回日期:**2022-06-25

勇等^[7]利用 GoogLeNet 模型对苹果的外观进行检测,优化后识别准确率为 91.91%。杨志锐等^[8]采用改进的 AlexNet 模型识别红枣缺陷,模型的平均识别精度为 94.79%。赵志衡等^[9]提出一种利用卷积神经网络识别花生籽粒完整性的算法,测试样本的识别精度为 98.18%。何进荣等^[10]使用模型融合的方式优化改进了苹果的分级模型,分级精度为 97.84%。

目前深度学习识别模型在鲜花生荚果品质分级方面的准确率和稳定性有待提升。为提升智能检测方法的准确率和稳定性,提出一种基于 SqueezeNet 基础的天府花生分级模型。首先将协调注意力嵌入 SqueezeNet 模型中;然后通过超参数调优及正则化技术,减少模型参数大小,使用改进的梯度优化算法提高模型识别准确率,通过对 5 种经典模型的对比试验,进一步验证该模型的稳定性,旨在为鲜花生荚果智能筛选机器的嵌入式部署提供理论依据。

1 材料与试验

1.1 材料与软硬件设备

天府 11 号花生样本于 2021 年 8 月分 3 批次采购于重庆市市区超市及周边农场。

计算机参数为 Intel Core i7-11700F CPU,主频 2.40 GHz,16 GB 内存,搭载为 NVIDIA GeForce GTX 3060 显卡,12 GB 显存。开发平台为 Ubuntu 20.04 系统,采用 Pytorch 深度学习框架,IDE 为 Jupyter Notebook,配置 Python 3.8,利用 OpenCV 进行图像预处理,拍摄设备为思蓝绿 CCD 工业相机、手动对焦镜头(200 万像素)。

1.2 构建花生荚果平衡数据集

根据食用花生标准及市场需求将花生荚果划分为 4 个等级:正常果、破损果、发霉果、发芽果。采集的原始数据中正常果图像 6 000 幅,破损果图像 1 632 幅,发霉果 1 230 幅,发芽果 1 520 幅,利用数据集平衡技术平衡数据集。取出最大值 6 000 作为基准,通过保留原图、旋转、水平镜像和垂直镜像的方法将其他荚果种类的图像也扩增到 6 000 幅,正常果、破损果、发霉果、发芽果的图像数比例为 1:1:1:1。至此,平衡后的数据集样本总数为 24 000 幅,每幅图像像素大小为 640×480 ,格式为 .jpg。平衡后的数据集训练的模型通常有更好的泛化能力和识别准确度^[11-12]。数据集中 60% 作为训练集,20% 作为验证集,剩余 20% 作为测试集评估模型。花生荚果的具体数据集样本个数划分如表 1 所示。

2 模型及方法

2.1 单个分类模型

VGG16、AlexNet、ResNet50、DenseNet121、SqueezeNet 是目前农业检测领域广泛使用的 5 种经典模

表 1 试验数据集划分

Table 1 Partition of experiment data

级别	训练集/幅	验证集/幅	测试集/幅
正常果	3 600	1 200	1 200
破损果	3 600	1 200	1 200
发芽果	3 600	1 200	1 200
发霉果	3 600	1 200	1 200
总计	14 400	4 800	4 800

型,VGG16 主要通过大幅增加卷积神经网络深度提高模型的性能;AlexNet 首次在 CNN 中采用 ReLU 激活函数,通过加入 dropout 层减少训练过程中的过拟合;ResNet50 引入残差模块,消除了梯度消失和网络退化的问题;DenseNet121 构建每层到其他层的连接,具有一定的正则化效果,缓解了模型过拟合并减少了特征冗余;SqueezeNet 网络参数内存小。

上述 5 种模型在结构上差异较大,但都由卷积层、密集连接层、分类器组成。常用的激活函数包括 Sigmoid、tanh、ReLU 等。试验采用 ReLU(Rectified Linear Unit,线性整流函数)激活函数,该函数为线性函数,网络的收敛速度快^[13],且不存在梯度消失和梯度饱和的问题^[14],其数学表达式为:

$$f(x) = \max(0, x), \quad (1)$$

经 softmax 函数输出一个 4 行 1 列的向量,4 为类别数。softmax 函数的数学表达式为:

$$y_j = \frac{e^j}{\sum_i e^i}, \quad (2)$$

式中:

y_j ——图像为第 j 个类别的概率;

e^j ——图像为第 j 个类别的上层函数值;

i ——共 i 个类别数。

2.2 协调注意力模块

协调注意模块^[15](CoordAtt)可以看作是一个计算单元,旨在提高网络模型的学习表达能力,计算单元具有相同的输入输出尺寸,模型的结构如图 1 所示。

设输入的特征图像为 X ,分别使用卷积核 $(H, 1)$ 、 $(1, W)$ 沿着图像的两个坐标进行卷积池化,卷积操作的数学表达式为:

$$Z_c^h(h) = \frac{1}{W} \sum_{0 \leq i < W} x_c(h, i), \quad (3)$$

$$Z_c^h(w) = \frac{1}{H} \sum_{0 \leq i < H} x_c(j, w), \quad (4)$$

式中:

(W, H) ——模块输入特征图尺寸;

(w, h) ——输出特征图尺寸;

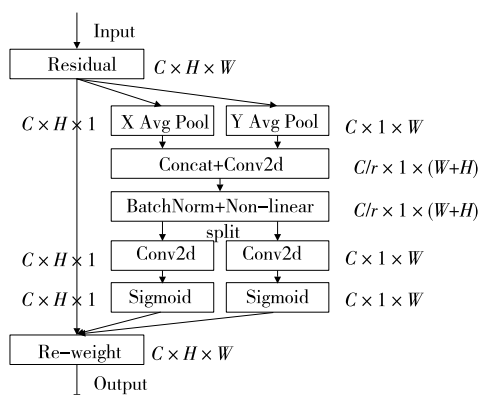


图 1 协调注意力模块

Figure 1 Coordinate attention block

c ——输出通道数。

通过式(3)和式(4)两种卷积池化,协调注意力模块可以在一个空间方向上捕捉依赖关系,并在另一个空间方向上保持精确的位置信息,这有助于网络更准确地定位感兴趣的区域。接着将两个方向上的聚合特征连接起来,并输入 1×1 的卷积核转化式 F_1 ,操作的数学表达式如式(5)所示。

$$f = \delta(F_1([Z^h, Z^w])), \quad (5)$$

式中:

δ ——非线性激活函数;

F_1 ——卷积核转化公式;

f ——将空间信息在水平方向和垂直方向上编码的中间特征图。

然后使用 $\text{split}()$ 方法把中间特征图划分为 f^h 和 f^w 两个独立张量,进一步使用 1×1 的卷积转化数学式 F_h 和 F_w ,得到与输入特征图像 X 相同尺寸的输出, σ 为 Sigmoid 激活函数,操作过程的数学表达式为:

$$g^h = \sigma(F_h(f^h)), \quad (6)$$

$$g^w = \sigma(F_w(f^w)). \quad (7)$$

由式(8)得到协调注意力模块计算的结果。

$$y_c(i, j) = x_c(i, j) \times g_c^h(i) \times g_c^w(j). \quad (8)$$

2.3 梯度集中

为了加快训练速度,提高泛化性能,对于深度学习神经网络,目前从激活、权重和梯度 3 个角度已经提出了多种优化技术。梯度集中^[16]通过引入新的优化函数,对模型反向传播的梯度进行优化,使权空间和输出特征空间

都正则化,从而提高训练速度和泛化能力。优化函数如式(9)所示,自变量为反向传播的梯度。

$$\Phi(\nabla_{w_i} \mathcal{L}) = \nabla_{w_i} \mathcal{L} - \frac{1}{M} \sum_{j=1}^M \nabla_{w_i} \mathcal{L}, \quad (9)$$

式中:

$\nabla_{w_i} \mathcal{L}$ ——反向传播的梯度;

M ——卷积核的个数。

2.4 协调注意力 SqueezeNet 模型(CG-SqueezeNet)

SqueezeNet 模型中的 fire 模块包括两层。通常定义一个 fire 模块为 $\text{fire}(A, B, C_1, C_2)$, 其中 A, B 分别代表 squeeze 层的输入、输出通道数, C_1, C_2 分别代表 expand 层中 $1 \times 1, 3 \times 3$ 卷积核的输出通道数目,结构如图 2 所示。

从经典模型出发,构建一种协调注意力 SqueezeNet 模型(Coordinate Attention SqueezeNet, CG-SqueezeNet),其基本架构如图 3 所示,包括 8 个 Fire 层,2 个卷积层、1 个协调注意力层、4 个池化层。与 SqueezeNet 模型相比,CG-SqueezeNet 模型的改进有:① 在 fire10 之前引入 CoordAtt 协调注意力模块,模型可以更快更准确地学习感兴趣区域的特征。② 将 fire10 模块中的 expand $1 \times 1, 3 \times 3$ 卷积层的输出通道数设置为 128, conv11 层的输入输出通道数分别设置成 256, 4。由于通道数的减少,网络参数量降低。③ 使用梯度优化算法,加速训练过程,提高模型泛化性。

3 结果与分析

3.1 超参数选取

基于试验的硬件水平,批大小设置为 32,所有图片经过一次学习称作一次迭代(epoch),取批尺寸为 2 的次方能够充分利用显卡的并行计算能力,若设置为较大的数

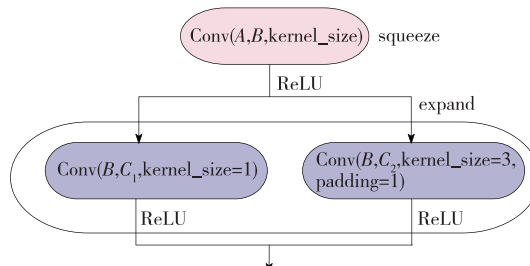


图 2 fire 模块结构

Figure 2 Fire modular configuration

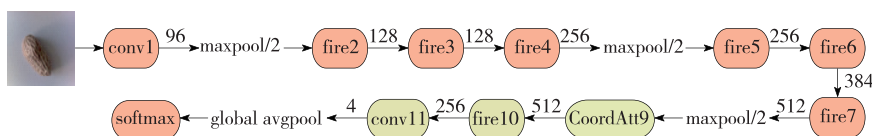


图 3 协调注意力 SqueezeNet 模型的网路架构

Figure 3 Network architecture of the Coordinate Attention SqueezeNet (CG-SqueezeNet) model

值如 128,会导致模型收敛到最优解的时间过长;若设置为 8 或 16 等较小数值,又容易产生震荡波动。训练模型时采用了带动量的随机梯度下降优化算法(SGD),momentum=0.9。采用了随机丢弃(dropout)技术,其参数 p 为 0.5,将初始学习率设置为 0.001,每 20 次迭代(epoch)后学习率减小为原来的 10%。由于采用了变学习率,初始学习率设置的很大,因而模型收敛速度很快,每个模型都只迭代了 50 次。

3.2 试验结果

表 2 记录了不同网络的训练集和测试集的准确率、模型参数内存大小。对于平衡后的数据集,CG-SqueezeNet 网络的分类效果最好,在测试集上准确率达到到了 97.83%;AlexNet 在测试集准确率为 93.17%,在训练网络中准确率最低。从参数内存需求的角度来看,VGG16 的参数内存需求为 537 MB;内存需求最小的是 CG-SqueezeNet,为 2.52 MB。因此如果想得到一个分类速率性能高的网络,宜选择 CG-SqueezeNet 模型,其训练过程如图 4 所示。

经测试集测试,4 800 个鲜花生荚果样本中,检测正确样本数为 4 688,总体准确率为 97.84%。1 200 个正常果中 12 个被误判为发霉果;1 200 个破损果样本全部判断正确;1 200 个发霉果中 80 个被误判为正常果,5 个误判为发芽果。1 200 个发芽样本全部判断正确。综上,改进的集成网络结构能够较好地识别鲜花生荚果的缺陷特征。

3.3 数据可视化

类激活图(class activation map,CAM)^[17]指针针对指定的输入图像生成的热力图,代表图像的各个位置对该类别的重要程度。图 5 是合格花生荚果、破损果、发霉

表 2 花生荚果分类结果

Table 2 Classification results of peanut pod

网络模型	训练集	测试集	参数内存/
	准确率/%	准确率/%	MB
VGG 16	95.61	94.73	537
AlexNet	96.08	93.17	228
ResNet50	98.01	95.54	94.4
DenseNet121	97.99	95.38	28.5
SqueezeNet	96.08	94.17	2.97
CG-SqueezeNet	98.28	97.83	2.52

果、发芽果的特征可视化示意图。特征图中红色表示神经网络判定花生荚果类别的高权重区域,蓝色表示网络分类低权重的区域,色彩越红越表示这块区域对分类结果的影响越大。从图 5 可以看出,网络在分类为正常果时高红亮区域较为均匀地集中在荚果本身,分类为破损果的重要部位在完整与破损衔接处,分类为发霉图像时会找到大面积黑亮区域及褐斑的位置,分类为发芽果时网络的关注点在于检测目标物的牙尖区域。

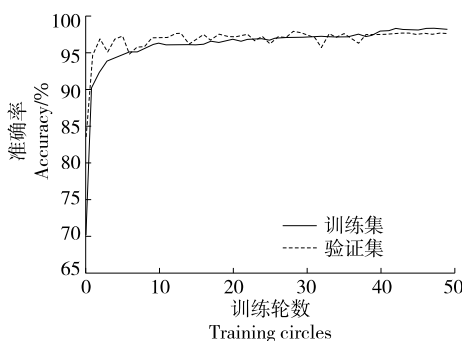


图 4 训练结果

Figure 4 Training results

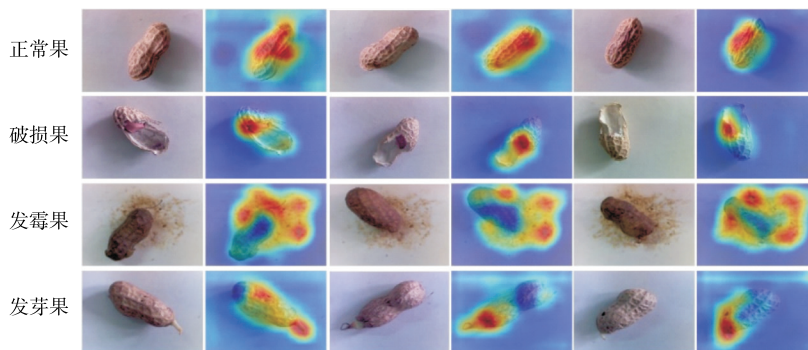


图 5 Grad-CAM 可视化测试集花生荚果特征区域

Figure 5 Visualization of test set peanut pod's feature region using Grad-CAM

4 结论

研究提出了一种嵌入协调注意力策略,该策略在学习特征的过程中更加注重位置信息,增强模型对图像感

兴趣区域特征的学习。利用梯度集中方法,改进优化算法,在测试集上对比不同经典模型,结果表明 CG-SqueezeNet 模型的预测结果会更加稳定、准确率达到 97.83%。与传统的迁移模型 VGG16、AlexNet、

ResNet50、DenseNet121、SqueezeNet 相比, CG-SqueezeNet 模型的内存仅为 2.52 MB,适用于小型嵌入式设备部署。后期将搭建花生的在线检测系统,探究模型在实用性方面的不足。

参考文献

- [1] 吕尚武, 尚书旗, 王伟伟, 等. 花生除杂(清选)分级机的设计与研究[J]. 农机化研究, 2019, 41(9): 71-75.
LU S W, SHANG S Q, WANG D W, et al. Design and research of peanut cleaning and sorting machine[J]. Journal of Agricultural Mechanization Research, 2019, 41(9): 71-75.
- [2] 王翔宇, 李鑫星, 张领先, 等. 农业主要病害检测与预警技术研究进展分析[J]. 农业机械学报, 2016, 47(9): 266-277.
WANG X Y, LI X X, ZHANG L X, et al. Research progress analysis of mainly agricultural diseases detection and early warning technologies[J]. Transactions of the Chinese Society for Agricultural Machinery, 2016, 47(9): 266-277.
- [3] 张明, 李鹏, 邓烈, 等. 基于掩模及亮度校正算法的脐橙表面缺陷分割[J]. 中国农业科学, 2019, 52(2): 327-338.
ZHANG M, LI P, DENG L, et al. Segmentation of navel orange surface defects based on mask and brightness correction algorithm[J]. Scientia Agricultura Sinica, 2019, 52(2): 327-338.
- [4] 杨斌, 钟金英. 卷积神经网络的研究进展综述[J]. 南华大学学报(自然科学版), 2016, 30(3): 66-72.
YANG B, ZHONG J Y. Review of convolution neural network[J]. Journal of University of South China (Science and Technology), 2016, 30(3): 66-72.
- [5] 卢宏涛, 张秦川. 深度卷积神经网络在计算机视觉中的应用研究综述[J]. 数据采集与处理, 2016, 31(1): 1-17.
LU H T, ZHANG Q C. Applications of deep convolutional neural network in computer vision[J]. Journal of Data Acquisition and Processing, 2016, 31(1): 1-17.
- [6] 陈超, 齐峰. 卷积神经网络的发展及其在计算机视觉领域中的应用综述[J]. 计算机科学, 2019, 46(3): 63-73.
CHEN C, QI F. Review on Development of convolutional neural network and its application in computer vision [J]. Computer Science. 2019, 46(3): 63-73.
- [7] 薛勇, 王立扬, 张瑜, 等. 基于 GoogLeNet 深度迁移学习的苹果缺陷检测方法[J]. 农业机械学报, 2020, 51(7): 30-35.
XUE Y, WANG L Y, ZHANG Y, et al. Defect detection method of apples based on GoogLeNet deep transfer learning[J]. Transactions of the Chinese Society for Agricultural Machinery, 2020, 51(7): 30-35.
- [8] 杨志锐, 郑宏, 郭中原, 等. 基于网中网卷积神经网络的红枣缺陷检测[J]. 食品与机械, 2020, 36(2): 140-145.
YANG Z R, ZHENG H, GUO Z Y, et al. Detection of jujubes based on the neural network with network convolution[J]. Food & Machinery, 2020, 36(2): 140-145.
- [9] 赵志衡, 宋欢, 朱江波, 等. 基于卷积神经网络的花生籽粒完整性识别算法及应用[J]. 农业工程学报, 2018, 34(21): 195-201.
ZHAO Z H, SONG H, ZHU J B, et al. Identification algorithm and application of peanut kernel integrity based on convolution neural network[J]. Transactions of the CSAE, 2018, 34(21): 195-201.
- [10] 何进荣, 石延新, 刘斌, 等. 基于深度学习的富士苹果外部品质分级方法[J]. 农业机械学报, 2021, 52(7): 379-385.
HE J R, SHI Y X, LIU B, et al. External quality grading method of Fuji apple based on deep learning[J]. Transactions of the Chinese Society for Agricultural Machinery, 2021, 52(7): 379-385.
- [11] 王铁伟, 赵瑶, 孙宇馨, 等. 基于数据平衡深度学习的不同成熟度冬枣识别[J]. 农业机械学报, 2020, 51(S1): 457-463, 492.
WANG T W, ZHAO Y, SUN Y X, et al. Identification of Jujube with different maturity based on data balance deep learning[J]. Transactions of the Chinese Society for Agricultural Machinery, 2020, 51(S1): 457-463, 492.
- [12] 高霁月, 倪建功, 杨昊岩, 等. 基于数据平衡和深度学习的开心果品质视觉检测方法[J]. 农业机械学报, 2021, 52(7): 367-372.
GAO J Y, NI J G, YANG H Y, et al. Visual detection of pistachios based on data balance and deep learning[J]. Transactions of the Chinese Society for Agricultural Machinery, 2021, 52(7): 367-372.
- [13] WANG S H, XIE S P, CHEN X Q, et al. Alcoholism identification based on an AlexNet transfer learning model[J]. Frontiers in Psychiatry, 2019, 10: 1-13.
- [14] 周飞燕, 金林鹏, 董军. 卷积神经网络研究综述[J]. 计算机学报, 2017, 40(6): 1 229-1 251.
ZHOU F Y, JIN L P, DONG J. Review of convolutional neural network[J]. Chinese Journal of Computer, 2017, 40(6): 1 229-1 251.
- [15] HOU Q, ZHOU D, FENG J. Coordinate attention for efficient mobile network design[C]// Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2021: 13 713-13 722.
- [16] YONG H, HUANG J, HUA X, et al. Gradient centralization: A new optimization technique for deep neural networks[C]// European Conference on Computer Vision. [s.l.]: Springer, Cham, 2020: 635-652.
- [17] SELVARAJU R R, COGSWELL M, DAS A, et al. Grad-CAM: Visual explanations from deep networks via gradient-based localization[J]. IEEE International Conference on Computer Vision (ICCV), 2017: 618-626.