

## 改进 LeNet-5 模型在大米分选算法上的应用

甘骥榕<sup>1</sup>, 苏芳<sup>1</sup>, 练坤玉<sup>1</sup>, 徐道际<sup>2</sup>, 董玉德<sup>1\*</sup>

(1. 合肥工业大学数字化设计与制造省级重点实验室, 合肥 230009; 2. 安徽宏实光机电高科有限公司, 合肥 230051)

**摘要:** 针对大米特征提取与分类的算法难以确定的问题, 改进 LeNet-5 卷积神经网络模型并研究其在大米分选问题上的表现。首先对大米原始图像进行预处理、提取单粒大米的图像建立大米图像样本库, 然后对原始的 LeNet-5 模型提出改进方案并进行实验, 最后将改进模型与若干传统分类方法、3 个轻量化卷积神经网络模型进行对比。改进 LeNet-5 模型大米形选准确率为 97.2%, 色选准确率 90.6%, 处理速度约 5 300 粒·s<sup>-1</sup>, 分类效果与分类速度显著优于其他对比方法。实验结果证明, 改进的 LeNet-5 模型可以高效分选碎米与垩白米, 且能有效减少实际分选前准备工作的工作量。

**关键词:** 机器视觉; 机器学习; 卷积神经网络; 大米分选

中图分类号: TP391

文献标识码: A

文章编号: 1672-352X (2019)03-0549-05

### Rice classification algorithm research based on modified LeNet-5 model

GAN Qirong<sup>1</sup>, SU Fang<sup>1</sup>, LIAN Kunyu<sup>1</sup>, XU Daoji<sup>2</sup>, DONG Yude<sup>1</sup>

(1. Provincial Key Laboratory of Digital Design and Manufacture, Hefei University of Technology, Hefei 230009;

2. Anhui Hongshi Optoelectronic High-tech Co.Ltd., Hefei 230051)

**Abstract:** Aiming at the problem that the algorithm of rice feature extraction and classification is difficult to determine, this paper improves the LeNet-5 convolutional neural network model and studies its performance on rice sorting. In this paper, the rice original image was preprocessed and the image of single grain rice was extracted to establish a rice image sample database. Then the original LeNet-5 model was improved and tested. Finally, the improved model and several traditional classification methods, 3 lightweight convolutional neural network models were compared. The improved LeNet-5 model has a rice shape selection accuracy of 97.2%, a color selection accuracy of 90.6%, and a processing speed of about 5300 particles·s<sup>-1</sup>. The experimental results show that the improved LeNet-5 model can efficiently sort broken rice and chalky rice, and can effectively reduce the workload of preparation before actual sorting.

**Key words:** machine vision; machine learning; convolution neural network; rice sorting

大米是人类主食之一, 是人类生存的必需品, 保障大米的品质具有重要的意义, 而大米分选是生产中品控必不可少的重要一环。

大米分选最初是以人工进行的。人工处理存在主观性, 人工分选的成本、效率问题也会随着工作时间而愈加明显。随着计算机技术的进步, 机器视觉技术被引入大米分选领域。相较于人类视觉, 机器视觉的分选标准设定更加准确, 色彩、形状的信息采集结果更加稳定, 封闭工作空间、控制环境光

照更加方便, 稳定工作时间也大大增长, 具有极大的优势。

传统的图像分类方法中特征提取与特征分类两部分结合不紧密, 随着需要分类的农产品种类日益丰富、标准日益严格, 所用算法的选择与训练难度也越来越高。此时, 采用卷积神经网络进行图像分类成为一个更优解。

卷积神经网络在农业领域已有应用, 如李敬<sup>[1]</sup>使用 6 层卷积神经网络取代专业的技术人员和技术

收稿日期: 2018-10-15

基金项目: 合肥市“基于物联网的精准农业光电分选设备产业化关键技术研究”(ZR201711290011) 资助。

作者简介: 甘骥榕, 硕士研究生。E-mail: ganqirong@foxmail.com

\* 通信作者: 董玉德, 博士, 教授。E-mail: yddong@hfut.edu.cn

设备对烟草病害进行识别分类，丁文宽<sup>[2]</sup>针对成熟辣椒的特征量较少和复杂的外界环境的特点使用卷积神经网络对辣椒进行识别与定位，袁培森等<sup>[3]</sup>使用端到端的卷积神经网络技术直接作用于菊花的原始图像数据，针对菊花进行了高效识别。

1998年LéCun<sup>[4]</sup>提出了LeNet-5模型，该模型在手写数字识别上取得了良好效果，标志着卷积神经网络真正面世。随后，多种层数更深、参数越多的卷积神经网络模型被提出，在运算量显著增加的同时有效提升了准确度。随后一些经典的轻量化神经网络如MobileNet<sup>[5]</sup>、Xception<sup>[6]</sup>和ShuffleNet<sup>[7]</sup>等面世，在准确率、模型体积和运算速度等之间进行了取舍，可以在手机等便携式设备上使用，促进了卷积神经网络的应用。

综合考虑到大米分选设备的性能限制与大米分选的准确度要求，本研究采用结构相对简单、运算量较低的LeNet-5模型来进行大米的分选，并对其改进以获得更高的准确度。通过研究，拟验证大米色选机采用卷积神经网络进行大米图像分选的有效性与可行性，为其他农产品分选提供指导。

### 1 卷积神经网络

#### 1.1 卷积神经网络简介

卷积神经网络是一种多层次堆叠的模型，它输入的是诸如图像、音频等原始的数据，然后经过卷积、降采样、非线性激励等操作，对原始数据进行特征提取与抽象，从低层提取基础特征如轮廓、拐角等，从高层提取更抽象、更整体的语义特征。以大米图片为输入，对本研究所使用的神经网络模型中间各层运算所得特征图进行提取，结果（图1）可以看出，底层提取到了大米的边缘、形状、色块和斑点等基础特征，中层更加关注边缘的主要部分、拐角以及大米整体形状，最高层所得到的特征图已经没有明显的局部特征，关注的是大米颗粒的整体。

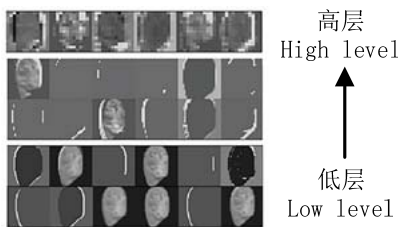


图1 中间层可视化

Figure 1 Visualization of intermediate layers

#### 1.2 卷积神经网络的训练

卷积神经网络的训练即是以目标函数为标准，对参数进行调整而减小误差。如图2所示，首先从

样本库中取一组样本图像，基于当前参数对样本进行预测，根据选用的目标函数来计算误差。然后根据误差，使用反向传播算法更新各层的参数，使更新后的模型做出的预测误差更小。随着参数的不断更新，最终误差将缩小到一个稳定值，此时模型训练完成。

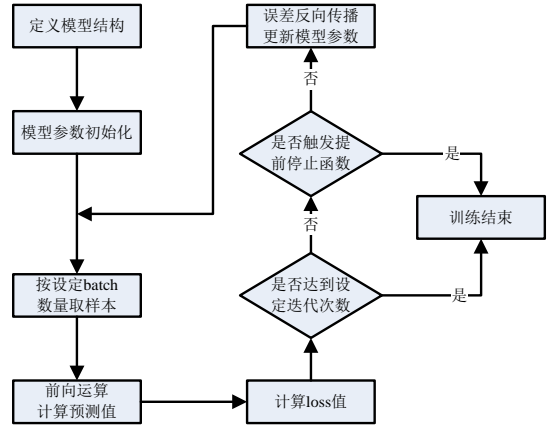


图2 卷积神经网络训练流程图

Figure 2 Training flowchart of convolutional neural network

### 2 LeNet-5 模型改进

原始的LeNet-5模型针对的是10类手写数字的识别，近年来卷积神经网络的相关理念有了巨大的进步，因此对LeNet-5进行改进以获得更好的性能，并在后文中针对输入图像尺寸、激励函数、优化算法等改进项目进行测试。

LeNet-5模型的输出层原本由10个欧式径向基函数组成，它们与手写的10个数字一一对应，最后共输出10个分类，在大米识别的问题中，形选与色选分别将大米分为2类，因此最终输出更改为由Softmax函数输出2类。

输入图片的尺寸与模型的计算量直接相关，在保证高正确率的前提下，应该优先选择较小的尺寸以减少运算负担，缩短运算耗时。LeNet-5模型输入的图片尺寸为32×32像素，本研究算法实验中选取了32×32像素、48×48像素、64×64像素等3个不同的输入尺寸来研究其影响，对各层的尺寸参数进行调整以适应新的输入。

激励函数是人工神经网络中重要的一环，激励函数能引入非线性因素，选用合适的激励函数能有效提高神经网络的性能表现。LeNet-5模型使用的是Sigmoid函数，Sigmoid函数两侧梯度趋近于0，容易造成梯度消失的现象，减缓收敛速度。梯度消失在网络层数多的时候尤其明显，是加深网络结构的主要障碍之一。算法测试中，将LeNet-5原有的

Sigmoid 函数更换为 Tanh、Relu 和 Elu 等激励函数来进行测试。

VGG16 中使用 3 个  $3 \times 3$  像素卷积核来代替  $7 \times 7$  像素卷积核, 使用了 2 个  $3 \times 3$  像素卷积核来代替  $5 \times 5$  像素卷积核。如图 3 所示, 这种替换方式在保证感受野不变的同时提升了网络的深度, 增加了非线性, 避免过拟合。原 LeNet-5 模型两个卷积层都采用了  $5 \times 5$  像素卷积核, 改进的 LeNet-5 模型中分别用 2 层  $3 \times 3$  像素卷积核进行替换, 并在卷积后增加激励函数 (图 3)。

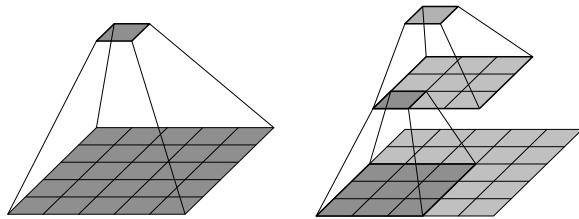


图 3 采用更小的卷积核

Figure 3 Apply smaller convolution kernels

优化算法 (Optimizer) 是影响卷积神经网络效果的一个重要因素, 优化算法经历了 SGD、SGDM、NAG、AdaGrad、AdaDelta、Adam 等算法的发展历程。原 LeNet-5 模型使用的是随机梯度下降法 (Stochastic gradient descent, SGD), 实验中使用 Adam 方法进行替换。

Adam 方法由 Kingma 等<sup>[8]</sup>于 2014 年提出, 该方法利用梯度的一阶矩估计和二阶矩估计对待训练参数的学习率进行动态调节, 并将每一次的迭代学习率限定在一个固定范围内, 使参数变化更缓和,

加快训练速度。Adam 方法实施起来很简单, 计算效率高, 存储需求小, 对梯度的对角线重新定标不变, 并且非常适合于数据和参数数量较大的问题, 该方法也适用于目标不稳定的问题和噪音复杂、梯度稀疏的问题。

模型的下采样方式也进行了改进。平均下采样的将滤波器范围内的值求和并平均, 得到输出, 而最大下采样将范围内元素最大的值作为输出, 最大下采样方法还能引入一定的不变性, 结合往年研究经验, 最终将下采样方法更改为最大下采样。

### 3 分类测试

针对提出的改进方案, 以准确率为主要指标, 首先采用原始 SGD 算法、固定学习率  $1 \times 10^{-6}$  针对各项改进内容进行测试, 然后在改进模型上修改优化算法为 Adam 算法再次测试, 对改进效果进行验证。测试中所得结果均为 10 次测试的平均结果。

仅修改输入图像尺寸所得的模型训练结果很不稳定, 受随机初始化的影响较大。测试反复进行了多次, 大部分情况下 loss 减小速度很慢甚至停滞不变, 难以收敛到较小的值, 此时准确率在 50% 上下波动或由于模型内发生梯度爆炸而变为 Nan, 少数情况下可以正常训练到较高准确率。研究这一部分结果, 发现大的输入尺寸能有效地提升模型分类效果, 因此在后续测试中均采用分辨率为  $64 \times 64$  的样本图片进行测试。

在上述修改基础上, 对模型所用激励函数的测试结果如表 1 所示。

表 1 采用不同激励函数的分类测试结果

Table 1 Result of classification test using different activation function

序号 No.	激励函数 Activation function	形选准确率/% Accuracy(shape)	色选准确率/% Accuracy(chalky)
1	Sigmoid	无稳定结果	无稳定结果
2	Tanh	93.10	无稳定结果
3	Relu	97.20	80.20
4	Elu	97.20	78.20

注: Tanh 函数的形选测试结果为训练得到结果且训练轮数达到 2 000 时的模型准确率。

Note: Test result about Tanh in shape classification is the accuracy of model which have got a result and training epoch equals to 2 000.

表 2 修改模型结构后的色选测试结果

Table 2 Result of chalky classification test after modifying the model's structure

激励函数 Activation function	优化算法 Optimization methods	准确率/% Accuracy
Relu	SGD	80.2
Elu	SGD	78.2
Relu	Adam	90.7
Elu	Adam	90.4

测试结果表明: 使用 Sigmoid 函数时模型无法训练出可用模型; tanh 函数在形选问题上能得到可用模型, 但是训练速度非常缓慢, 在色选问题上无法训练出可用模型; 采用 Relu、Elu 函数的模型能训练得到可用模型。经过比较, 最终采用 Relu、Elu 函数的改进方案。

在测试过程中, 形选的多次测试结果均在 97% 至 98.5% 之间波动, 考虑到样本人工标记时无法避

免的主观性问题，可以认为形选问题准确率已经受限于样本库人工标记本身的准确度，而色选问题所得准确率还有提升的空间，因此后续步骤基于色选问题进行测试。

在使用小卷积核加深网络、增加激励函数、修改优化算法与下采样方式后再进行测试，结果如表 2 所示。

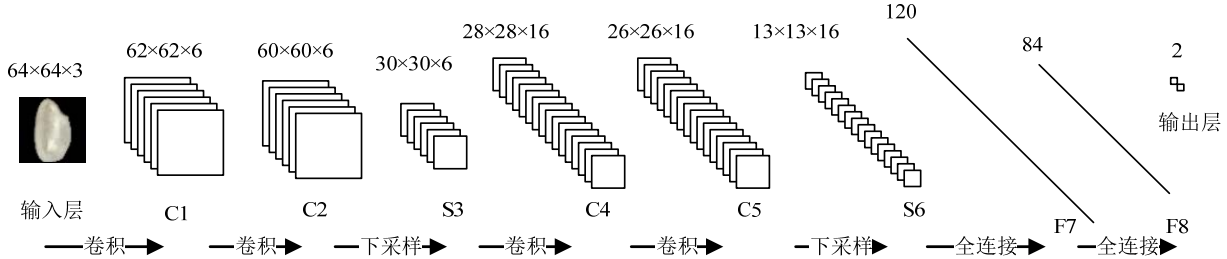


图 4 改进模型示意图

Figure 4 Schematic of improved model

### 4 算法性能

#### 4.1 评价指标

实验中的大米形选与色选是二分类问题，即大米每次都被分成正类 (Positive) 或负类 (Negative)。对一个机器学习的二分类问题来说，预测结果会出现 4 种情况：真正 (True Positive, TP)、假正 (False Positive, FP)、假负 (False Negative, FN)、真负 (True Negative, TN)。其中真正与真负分别指正确分类的正、负样本，假正、假负分别指被错误分类为正的负样本、被错误分类为负的正样本。

结合大米色选机的实际生产情况，选择以下几个评价指标：

可见修改后准确率有显著提高。与此同时，Adam 优化算法还大大缩短了训练所需要的时间。

改进前，原 LeNet-5 模型在形选问题上表现不稳定，在色选问题上完全无法收敛，改进后，在色选与形选问题上均能获得稳定的结果，多次实验所得结果波动很小，改进是有效果的。最终，修改后的模型结构如图 4 所示。

准确率 (Accuracy)，即正确分类的样本结果数与总样本数之比，是最常见的评价指标，计算方式如式 (1)。

$$ACC = \frac{TP+TN}{TP+TN+FP+FN} \tag{1}$$

色选精度，也称为选出率 (SEL)，指从大米中选出的不合格大米占不合格大米总量的百分数。可见色选精度与真负率是同一个概念的。在实际生产中，单次色选精度如果过低，则不合格物料残余过多，为了确保最终大米的质量，需要多次筛选。计算方式如式 (2)。

$$SEL = \frac{TN}{TN+FP} \tag{2}$$

表 3 分类测试表现

Table 3 Performance of classification

分类类型 Classification Types	分类方法 Classification Methods	准确率/% Accurate	色选精度/% Sorting Accuracy	带出比 Carryover Ratio
形选 Shape	原始模型	93.60	93.10	11.50
	改进模型	98.00	97.20	81.00
色选 Chalky	原始模型	分选失败	分选失败	分选失败
	改进模型	90.70	90.60	9.85

注：原始 LeNet-5 模型训练给出的结果不稳定，形选取估计值，色选分选失败。

Note: The results of the original LeNet-5 model are unstable, shape classification use estimated value, color classification fails.

带出比 (CAR)：色选机抛弃的大米中真正不合格大米数量与错判的正常大米数量之比。如果带出比过小，说明带出的正常大米多，会造成原料的浪费，影响成本，这个指标在实际生产中通常根据实际需要进行调整<sup>[9]</sup>。计算方式如式 (3)。

$$CAR = \frac{TN}{FN} \tag{3}$$

#### 4.2 大米分类测试

大米分类测试结果如表 3 所示。

在实验电脑上对样本库中包含单粒大米的

64×64 像素样本图像进行分选测试, 改进 LeNet-5 模型分选千粒大米平均耗时 188 ms, 折合 1 s 约处理 5 300 粒。改进模型虽然相较原模型增加了运算量, 处理耗时随之增加, 但是仍然保持了足够快的速度, 因此准确度与速度的交换是值得的。

## 5 结论

卷积神经网络将特征提取与分类器的优化整合在一起, 自动提取特征、自动优化参数, 简化了训练的步骤, 降低了对图像预处理等步骤的要求。卷积神经网络还能充分利用大量样本数据, 且能利用新样本进一步训练, 更适合色选机的工作模式。

相较于原始的 LeNet-5 模型, 改进后的模型采用更大的图像输入尺寸来引入更多信息、采用双层 3×3 像素卷积层替换单层 5×5 像素卷积层来加深网络、使用 Relu 激励函数减轻梯度消失的现象、使用 Adam 优化算法增强训练效率, 最终有效提升了模型的分选效果。同时, 改进的 LeNet-5 模型在保证准确度的同时仍然有足够快的处理速度, 为后续模型的改动以及色选机其他处理步骤留下更充足的时间。

## 参考文献:

- [1] 李敬. 基于卷积神经网络的烟草病害自动识别研究[D]. 泰安: 山东农业大学, 2016.
- [2] 丁文宽. 基于卷积神经网络和机器视觉的辣椒检测与识别[D]. 天津: 天津理工大学, 2017.
- [3] 袁培森, 黎薇, 任守纲, 等. 基于卷积神经网络的菊花花型和品种识别[J]. 农业工程学报, 2018, 34(5): 152-158.
- [4] LÉCUN Y, BOTTOU L, BENGIO Y, et al. Gradient-based learning applied to document recognition[J]. Proceedings of the IEEE, 1998, 86(11):2278-2324.
- [5] HOWARD A G, ZHU M, CHEN B, et al. MobileNets: Efficient convolutional neural networks for mobile vision applications[J]. arXiv:1704.04861, 2017.
- [6] CHOLLET F. Xception: Deep learning with depthwise separable convolutions[C]//Proceedings of the IEEE conference on computer vision and pattern recognition. Honolulu, 2017.
- [7] ZHANG X Y, ZHOU X N, LIN M X, et al. ShuffleNet: an extremely efficient convolutional neural network for mobile devices[C]//2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Salt Lake City, 2018.
- [8] KINGMA D P, BA J. Adam: A method for stochastic optimization[C]//San Diego: International Conference on Learning Representations. 2015.
- [9] 李素梅. 如何正确认识白米色选机[J]. 粮食与饲料工业, 2002(7): 12-14.