

DOI: 10.13652/j.spjx.1003.5788.2022.90149

# 基于深度学习卷积神经网络的花生籽粒完整性检测

Peanut kernel integrity detection based on deep learning convolution neural network

张军锋<sup>1</sup>尚展垒<sup>2</sup>ZHANG Jun-feng<sup>1</sup> SHANG Zhan-lei<sup>2</sup>

(1. 河南水利与环境职业学院, 河南 郑州 450011; 2. 郑州轻工业大学, 河南 郑州 450002)

(1. Henan Vocational College of Water Conservancy and Environment, Zhengzhou, Henan 450008, China; 2. Zhengzhou University of Light Industry, Zhengzhou, Henan 450002, China)

**摘要:**目的: 准确区分完整花生、果仁破损花生和表皮破损花生。方法: 提出了一种基于深度学习卷积神经网络(CNN)的花生籽粒完整性检测方案。搭建了花生籽粒色选系统, 建立了花生籽粒图像库; 利用改进的密度峰值聚类(DPC)算法对 CNN 卷积核进行自适应压缩, 有效平衡网络深度和运算效率; 采用改进的麻雀搜索算法对 CNN 超参数配置和网络结构进行优化, 得到适用于花生籽粒完整性检测的 CNN 模型。结果: 相比于 DL-CNN、CO-Net 等检测方法, 该方案识别准确率提高了 5.41%~13.92%, 花生籽粒单幅图像检测时间缩短了约 16.9%。结论: 该方法可有效提高花生籽粒完整性检测的准确率和实时性。

**关键词:** 花生籽粒; 图像处理; 卷积神经网络; 密度峰值聚类; 麻雀搜索算法; 准确率

**Abstract: Objective:** To accurately distinguish intact peanut, nut damaged peanut and epidermis damaged peanut. **Methods:** A peanut seed integrity detection scheme based on deep learning convolution neural network (CNN) was proposed. The peanut seed color selection system was established and a peanut seed image database was also established; The improved density peak clustering (DPC) algorithm was used to adaptively compress the CNN convolution kernel to effectively balance the network depth and operation efficiency; The improved sparrow search algorithm was used to optimize the CNN super parameter configuration and network structure, and the CNN model suitable for peanut grain integrity detection was obtained. **Results:** Compared with other de-

tection methods, this scheme improved the recognition accuracy by about 5.41%~13.92%, and the detection time of single image of peanut grain was shortened by about 16.9%. **Conclusion:** This method effectively improves the accuracy and real-time of peanut grain integrity detection.

**Keywords:** peanut seeds; image processing; convolutional neural network; density peak clustering; sparrow search algorithm; accuracy

花生籽粒图像快速准确识别是花生籽粒完整性分选的关键<sup>[1]</sup>。近年来, 随着深度学习理论的快速发展, 卷积神经网络(CNN)作为一种具有代表性的深度学习技术逐渐被应用于图像识别领域<sup>[2-5]</sup>。与传统视觉图像检测技术相比, CNN 蕴含着大量的图像特征信息, 极大地提高了图像检测效率<sup>[6]</sup>。

学者们围绕 CNN 食品完整性检测开展了一系列研究并取得了一定成果。赵志衡等<sup>[1]</sup>提出了一种基于改进 CNN 的花生籽粒完整性识别算法, 该算法对花生分类的准确率可以达到 98.18%, 但是并没有考虑真实应用环境对检测精确度的影响。张忠志等<sup>[7]</sup>采用双分支卷积神经网络对红枣缺陷进行检测, 具有较好的分类识别效果, 但是该方法依赖大规模训练数据集, 而且模型参数配置过程耗时较长且得到的参数未必最优。张瑞青等<sup>[8]</sup>采用迁移学习卷积神经网络对花生荚果等级进行识别, 分类识别准确率达到 95.43%, 但该模型采用随机梯度下降技术进行模型参数优化, 效率较低。谢为俊等<sup>[9]</sup>在 AlexNet 网络基础上搭建适用于油茶籽完整性检测的 CNN 模型, 并验证了该模型的有效性。因此, 运用 CNN 技术进行食品目标识别检测具有广阔的应用前景。

研究拟结合花生籽粒完整性图像检测特点, 提出深度学习卷积神经网络花生籽粒完整性检测方法, 设计改进的

**基金项目:** 河南省高等学校青年骨干教师培养计划项目(编号: 2018GGJS281)

**作者简介:** 张军锋(1982—), 男, 河南水利与环境职业学院副教授, 硕士。Email: zhangzj5566@126.com

**收稿日期:** 2022-02-18

密度峰值聚类算法和麻雀搜索算法,并对神经网络模型进行优化,最后通过对比试验来验证该方法的有效性,旨在为花生籽粒完整性检测研究提供理论依据。

## 1 花生籽粒检测系统和卷积神经网络

基于彩色色选设备研究花生籽粒完整性图像检测方法,以准确区分完整花生、果仁破损花生和表皮破损花生为目的,搭建如图 1(a)所示的花生籽粒筛选系统,利用工业线阵彩色 CCD 相机实时获取花生籽粒图像信息,图像信息经边缘检测、图像分割以及滤波去噪处理后,采用提出的检测算法进行筛选分类,最后利用空气喷枪将判定为果仁破损花生、表皮破损花生进行剔除得到完整花生籽粒。

### 1.1 花生籽粒图像库

建立花生籽粒图像数据库,数据库中包括完整花生、果仁破损花生和表皮破损花生 3 类,如图 1(b)所示。对采集得到的花生籽粒图像进行人工标记,为增加训练样本数量和多样性,采用镜像<sup>[10]</sup>、平移<sup>[11-12]</sup>、旋转<sup>[13-14]</sup>方法对图像进行有效数据增广,最终得到 3 600 幅花生籽粒图片(每个分类各 1 200 幅),统一设置图片格式为 224 像素×224 像素正方形图像。模型训练时,从图像数据库中随机选取 80%(每类 840 幅,共 2 520 幅)样本为训练集,剩余 20%为测试集。

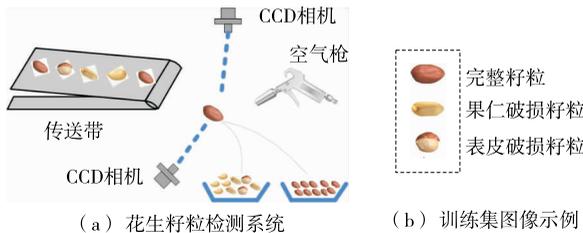


Figure 1 Schematic diagram of peanut grain detection system and image database

### 1.2 卷积神经网络

卷积层、池化层、全连接层是 CNN 的主要组成部分,其中,卷积层和池化层交替出现,实现对数据的特征学习;全连接层采用全连接的方式与最后一个池化层进行连接,实现对提取特征的分类。

1.2.1 卷积层 对于大小为  $M \times N$  的输入图像,用矩阵  $x$  描述,设卷积核为  $m \times n$  的矩阵  $w$ ,偏置为  $b$ ,卷积层通过多个卷积核  $w$  和激活函数  $f(\cdot)$ ,实现对输入数据的卷积运算和构建特征映射关系,得到一系列特征图  $h$ (如图 2 所示)。卷积计算表达式为:

$$h = f(x \otimes w + b), \quad (1)$$

式中:

$h$ ——输入图像经卷积计算后的特征图;

$f(\cdot)$ ——激活函数;

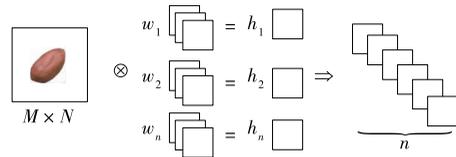


Figure 2 Schematic diagram of convolution kernel feature extraction process

$x$ ——卷积核  $M \times N$  的矩阵;

$w$ ——卷积核  $m \times n$  的矩阵;

$b$ ——偏置项。

1.2.2 池化层 池化层主要用来压缩卷积层输出特征,保留主要特征,池化层不改变卷积层输出的特征图数量。通过式(2)和式(3)分别定义最大值池化  $y_{\max}$  和平均值池化  $y_{\text{ave}}$  的数学表达式:

$$y_{\max} = \max(x_1, \dots, x_l), \quad (2)$$

$$y_{\text{ave}} = \frac{x_1 + \dots + x_l}{l}, \quad (3)$$

式中:

$x_1$ ——池化区域第 1 个元素像素值;

$x_l$ ——池化区域第  $l$  个元素像素值;

$l$ ——步长。

1.2.3 全连接层 全连接层将特征映射到标记空间,包含隐含层和输出层,对于全连接隐含第  $k$  层,设定输入值为  $x^{k-1}$ ,权值向量为  $w^k$ ,偏置为  $c^k$ ,利用激活函数  $f(\cdot)$  得到  $k$  层输出  $z^k$ :

$$z^k = f(w^k x^{k-1} + c^k). \quad (4)$$

当  $k=1$  时,  $x^{k-1}$  为最后一层池化层铺展后值。全连接输出层通常采用 Softmax 函数<sup>[15]</sup>进行概率分布处理以便于计算损失函数。

1.2.4 卷积神经网络存在的主要问题 主要问题有:① 多个卷积核得到的图像特征冗余度缺乏定量分析,影响了网络运算效率;② 网络结构和参数设置对分类结果影响较大,传统网络参数设定方式很容易导致网络陷入局部最优;③ 扩展网络深度能够提升网络性能,但是也增加了网络参数规模和计算复杂度。

## 2 改进卷积神经网络

为提升卷积神经网络性能,结合花生籽粒图像相对简单的特点对传统 AlexNet 模型进行改造:设计 2 层卷积层、2 层池化层,采用 2 层全连接层替代原有 3 层全连接层;将批归一化(BN)<sup>[16]</sup>置于卷积层激活函数后,加快网络训练速度;利用改进的密度峰值聚类(IDPC)算法对卷积核特征进行自适应压缩,采用改进的麻雀搜索算法对网络超参数结构进行优化配置。图 3 为改进卷积神经网络示意图。

### 2.1 密度峰值聚类压缩卷积核提取特征

DPC(密度峰值聚类算法)参数相对简单、鲁棒性强,

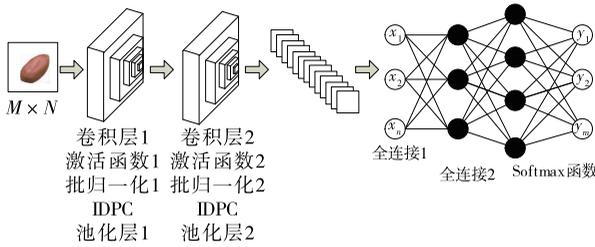


图 3 改进卷积神经网络结构示意图

Figure 3 Structure diagram of improved convolutional neural network

对任意形状数据均有良好的适用性<sup>[17-18]</sup>。对于第  $k$  个卷积层,采用  $n$  个卷积核  $\{w_1, \dots, w_n\}$  对图层进行卷积操作,得到特征图  $\{h_1, \dots, h_n\}$ ,利用 DPC 算法对  $\{h_1, \dots, h_n\}$  进行聚类分析,得到  $C$  个聚类,选取距离每个聚类中心  $(v_1, \dots, v_c)$  最近的特征图为最佳特征图  $\{h'_1, \dots, h'_c\}$ ,  $\{h'_1, \dots, h'_c\}$  对应的卷积核即为最佳卷积核  $\{w'_1, \dots, w'_c\}$ ,从而简化了网络参数。DPC 认为数据点  $h_i$  能够成为聚类中心的前提是  $h_i$  的局部密度  $\rho_i$  要高于周围其他点,且具有较大最近点距离  $\delta_i$ 。

$$\rho_i = \sum_j \chi(d_{ij} - d_c), \text{ 其中 } \chi(x) = \begin{cases} 1, x < 0 \\ 0, \text{else} \end{cases}, \quad (5)$$

$$\delta_i = \min_{i: \rho_j > \rho_i} d_{ij}, \quad (6)$$

$$\gamma_i = \rho_i \delta_i, \quad (7)$$

式中:

- $\rho_i$ ——特征图的  $h_i$  的局部密度;
- $\rho_j$ ——特征图的  $h_j$  的局部密度;
- $d_{ij}$ ——数据点  $h_i$  与  $h_j$  的距离;
- $\delta_i$ ——数据点  $h_i$  的最近点距离;
- $d_c$ ——截断距离;
- $\gamma_i$ ——聚类中心判定参数。

确定每个数据点的  $\rho_i$ 、 $\delta_i$  后, DPC 按照  $\gamma_i$  大小,选取前  $C$  个点为聚类中心,其余点划分到距离最近聚类中心所在分类。

研究<sup>[19]</sup>表明, DPC 对大规模数据集的聚类结果是稳健的,但对小规模数据集较为敏感,即当出现某个点分配错误时,可能会使得更多的点分配错误。为此,对数据点分配过程进行改进,引入核距离度量  $D_{ij}$  替代  $d_{ij}$ ,其数学表达式见式(8),并定义  $h_i$  和  $h_j$  的相似性  $\zeta_{ij}$ ,其数学表达式见式(9)。

$$D_{ij} = \|\Phi_r(X_j) - \Phi_r(X_i)\|^2, \quad (8)$$

$$\zeta_{ij} = 1/(1 + D_{ij}), \quad (9)$$

式中:

$\Phi$ ——核函数。

从式(9)可以看出,  $h_i$  与  $h_j$  距离越近,它们的相似性越高,因此属于同一聚类的可能性越大。定义数据点  $h_i$  属于第  $c$  ( $c = 1, \dots, C$ ) 个聚类的概率  $p_{i,c}$ ,其数学表达式为:

$$p_{i,c} = \frac{\sum_{j \in KNN_{i,y_i=c}} \lambda_{ij} \times \zeta_{ij}}{\sum_{t \in KNN_j} \zeta_{ij}}, \text{ 其中 } \lambda_{ij} = \quad (10)$$

式中:

$y_i$ —— $h_i$  的分类标号,  $y_i = c$  表示  $h_i$  属于第  $c$  个聚类;

$KNN_j$ —— $h_i$  的  $k$  近邻集合<sup>[20]</sup>。

确定  $h_i$  的所有概率  $\{p_{i,1}, \dots, p_{i,c}\}$  后,将其划分到概率最大的聚类中。

采用改进的 DPC (IDPC) 算法对卷积核进行提取,使得网络参数规模由  $n$  降到  $C$ ,最大限度地降低了卷积核冗余度,简化了网络参数。图 4 为改进密度峰值聚类压缩卷积核提取特征示意图。

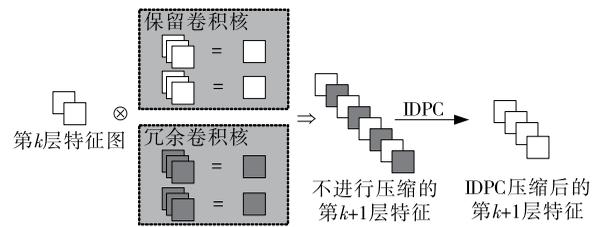


图 4 改进密度峰值聚类压缩卷积核提取特征示意图

Figure 4 Feature extraction diagram of IDPC compressed convolution kernel

## 2.2 网络超参数、网络结构优化配置

设计的具有 2 层卷积层、2 层池化层、2 层全连接层的改进卷积神经网络涉及的网络参数有训练批量大小  $b_1$ , 卷积核数目  $k_1, k_2$ , 矩阵大小  $m_1 \times n_1, m_2 \times n_2$ , 选取  $3 \times 3, 5 \times 5, 7 \times 7$  为备选卷积核大小; 激活函数类型, 选取  $f_1$  为 sigmoid 函数<sup>[10]</sup>,  $f_2$  为 Tanh 函数<sup>[11]</sup>,  $f_3$  为 ReLU 函数<sup>[11]</sup>,  $f_4$  为 Maxout 函数<sup>[12]</sup> 为备选激活函数; 池化操作类型, 选取  $P_1$  最大值池化  $y_{\max}$ 、 $P_2$  平均值池化  $y_{\text{ave}}$  为备选池化操作; 全连接层隐含神经元个数  $q_1, q_2$ , 神经元中心  $(c_1^1, \dots, c_{q_1}^1), (c_1^2, \dots, c_{q_2}^2)$ 。选取改进的麻雀搜索算法 (SSA)<sup>[20]</sup> 对 CNN 超参数配置和网络结构进行优化, 得到适用于花生籽粒完整性检测的 CNN 模型。

2.2.1 改进 SSA 更新进化策略 以训练误差作为 SSA 目标函数  $F(\text{SSA})$ , 麻雀个体  $X$  编码定义为:

$$X = (b_1, k_1, k_2, M_1, M_2, f^1, f^2, A_1, A_2, q_1, q_2) (c_1^1, \dots, c_{q_1}^1) (c_1^2, \dots, c_{q_2}^2), \quad (11)$$

式中:

$M_1, M_2$ ——2 层卷积层卷积核数大小类型, 且  $M_1 (M_2) = 1, 2, 3$  分别表示大小为  $3 \times 3, 5 \times 5, 7 \times 7$ ;

$f^1, f^2$ ——激活函数类型, 且  $f^1 (f^2) = 1, 2, 3, 4$  分别表示激活函数为  $f_1, f_2, f_3, f_4$ ;

$A_1, A_2$ ——2 层池化层操作类型, 且  $A_1 (A_2) = 1, 2$  分别表示  $P_1$  最大值池化、 $P_2$  平均值池化。

从  $X$  编码定义可以看出, 既具有连续编码位  $(c_1^1, \dots,$

$c_{q_1}^1$ ),  $(c_1^2, \dots, c_{q_2}^2)$ , 又具有整数编码位  $b_1, M_1, M_2, f^1, f^2, A_1, A_2$ , 如果仍采用传统更新进化方式, 会产生大量不符合要求的解, 为此分别重新定义 SSA 更新进化策略:

(1) 对于连续编码位  $(c_1^1, \dots, c_{q_1}^1), (c_1^2, \dots, c_{q_2}^2)$ , 定义子编码  $X'_1 = [(c_1^1, \dots, c_{q_1}^1), (c_1^2, \dots, c_{q_2}^2)]$ , 对于个体  $X_a$ , 引入加速进化机制和深度搜索机制:

$$X'_{a,1} = X'_{a,1} + \text{rand}(-1, 1)(X'_{\text{best},1} - X'_{a,1}), \quad (12)$$

$$X'_{a,1} = X'_{a,1} + \text{rand}(-1, 1)(X'_{i,1} - X'_{j,1}) + \text{rand}(0, 1)(X'_{\text{best},1} - X'_{a,1}), \quad (13)$$

式中:

$X'_{\text{best},1}$ ——种群最优解  $X_{\text{best}}$  对应的连续编码位,  $X_a \neq X_i \neq X_j$ .

麻雀种群内一部分个体采用加速进化机制进行更新, 直接向最优个体进行学习, 提升了算法收敛效率; 一

$$X_a = (X'_{a,1}, X'_{a,2}, X'_{a,3}), \quad (15)$$

$$X'_{a,1} = \begin{cases} X'_{a,1} + \text{rand}(-1, 1)(X'_{\text{best},1} - X'_{a,1}) \\ X'_{a,1} = X'_{a,1} + \text{rand}(-1, 1)(X'_{i,1} - X'_{j,1}) + \text{rand}(0, 1)(X'_{\text{best},1} - X'_{a,1}) \end{cases}, \quad (16)$$

$$X'_{a,2} = \begin{cases} X'_{a,2} + [\text{rand}(-1, 1)(X'_{\text{best},2} - X'_{a,2})] \\ X'_{a,2} = X'_1 + [\text{rand}(-1, 1)(X'_{i,2} - X'_{j,2}) + \text{rand}(0, 1)(X'_{\text{best},2} - X'_{a,2})] \end{cases}, \quad (17)$$

$$X'_{a,3} = \begin{cases} \alpha(X'_{\text{best},3} \rightarrow X'_{a,3}) \\ \alpha\{X'_{\text{best},3} \rightarrow \alpha[X'_{j,3} \rightarrow \alpha(X'_{i,3} \rightarrow X'_{a,3})]\} \end{cases} \quad (18)$$

2.2.2 改进卷积神经网络实现 利用 IDPC 算法提取每层最佳卷积核, 采用 ISSA 优化网络超参数和网络结构, 最终得到适用于花生籽粒完整性检测的 CNN 模型。其实现过程如图 5 所示。

### 3 仿真试验

#### 3.1 优化网络参数

使用图像识别领域基准 MNIST 数据集、CIFAR-10 数据集和花生籽粒图像库训练集进行模型训练, 其中 MNIST 数据集包括 60 000 张单通道灰度图像训练集、10 000 张测试集, CIFAR-10 数据集包括 60 000 张 3 通道

部分个体采用深度搜索机制, 向更多个体进行学习, 扩展了样本搜索空间, 提高了算法收敛精度。

(2) 对于整数编码位  $k_1, k_2, b_1, q_1, q_2$ , 定义子编码  $X'_2 = (b_1, k_1, k_2, q_1, q_2)$ , 由于  $X'_2$  编码位在连续整数范围内变化, 可以等效为连续编码, 即对于个体  $X_a$ , 有:

$$X'_{a,2} = X'_{a,2} + [\text{rand}(-1, 1)(X'_{\text{best},2} - X'_{a,2})], \quad (14)$$

其中,  $[\cdot]$  表示向下取整数。

(3) 对于整数编码位  $M_1, M_2, f^1, f^2, A_1, A_2$ , 定义子编码  $X'_3 = (M_1, M_2, f^1, f^2, A_1, A_2)$ , 对于  $X'_{a,3}$ , 定义其向个体  $X'_{j,3}$  学习更新方式为随机选取  $X'_{j,3}$  内  $\alpha$  个不同编码为替代  $X'_{a,3}$  ( $1 \leq \alpha \leq 6$ ), 并描述为  $X'_{a,3} = \alpha(X'_{j,3} \rightarrow X'_{a,3})$ 。联合式(12)~式(14)可得  $X_a$  更新进化公式为:

彩色图像训练集、10 000 张测试集。采用深度学习卷积神经网络(DL-CNN)、只利用 ISSA 优化网络结构的卷积神经网络(ISSA-CNN)对 3 种数据集的训练集进行模型训练。

由表 1 和图 6 可知, DL-CNN、ISSA-CNN 能够给出优化后的网络参数。对于 MNIST、CIFAR-10 以及花生图像测试数据集, DL-CNN 给出的激活函数优化结果为  $f_2, f_1, f_3$  (第一层)、 $f_3, f_1, f_2$  (第二层); 对于每层卷积核数目, DL-CNN 优化后的参数明显小于 ISSA-CNN, 这表明, IDPC 算法的引入能够降低卷积核数目规模, 简化网络结构。

利用花生籽粒图像库训练集验证归一化方法和全连接层数分别对模型性能的影响, 其中归一化方法选取 BN、LRN<sup>[15]</sup>, 全连接层数分别设置为 1 层、2 层, 其他网络参数选取 DL-CNN 优化后的参数, 结果见表 2。由表 2 可知, 归一化方法 LRN 的引入对模型准确率影响较小, 但是大幅增加了网络参数训练时间; 而 BN 的引入能够大幅度提升模型准确率, 而且模型训练时间小于 LRN; 2 层全连接层准确率高于 1 层全连接层, 但是两者的模型训练时间相近。综上, 采用 BN 归一化方法和 2 层全连接层使模型具有更高的准确率和训练时间。

#### 3.2 对比试验结果

为进一步验证 DL-CNN 性能, 选取标准 CNN、ISSA-CNN 和文献[9]提出的 CO-Net, 对 MNIST 数据集、CIFAR-10 数据集和花生籽粒图像库中的训练集进行对比试验。

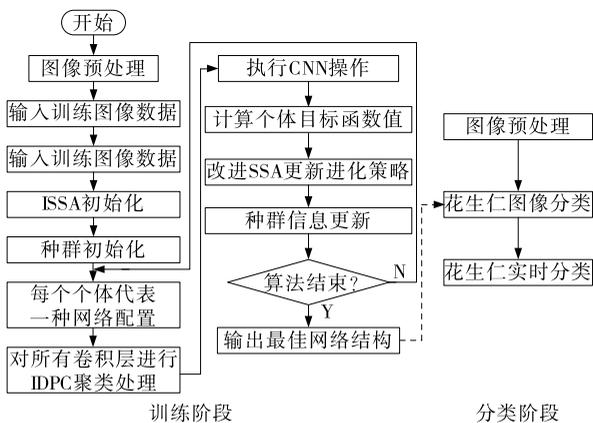


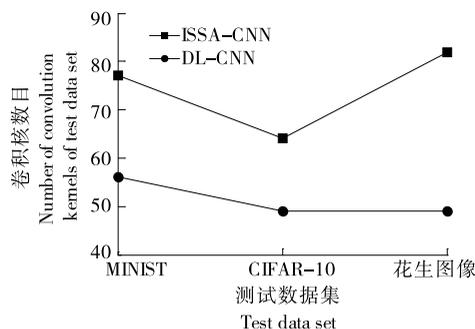
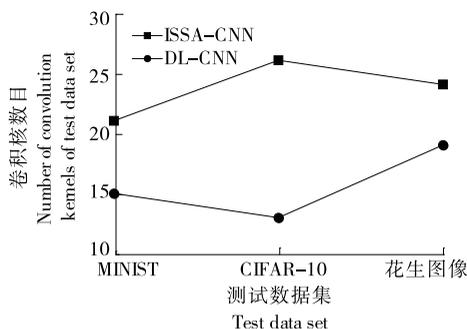
图 5 改进卷积神经网络实现流程图

Figure 5 Implementation flow chart of improved convolutional neural network

表 1 参数优化范围及 DL-CNN 算法优化结果

Table 1 parameter optimization range and optimization results of DL-CNN algorithm

参数	参数优化范围	CNN 参数	ISSA-CNN 优化后参数			DL-CNN 优化后参数		
			MNIST	CIFAR-10	花生图像	MNIST	CIFAR-10	花生图像
$b_1$	[10,100]	10	12	13	17	11	14	16
$k_1$	[4,90]	10	21	26	24	15	13	19
$M_1$	$3 \times 3, 5 \times 5, 7 \times 7$	$5 \times 5$	$5 \times 5$	$3 \times 3$	$5 \times 5$	$3 \times 3$	$5 \times 5$	$3 \times 3$
$f^1$	$f_1, f_2, f_3, f_4$	$f_1$	$f_1$	$f_1$	$f_3$	$f_2$	$f_1$	$f_3$
$A_1$	$y_{max}, y_{ave}$	$y_{ave}$	$y_{max}$	$y_{ave}$	$y_{ave}$	$y_{ave}$	$y_{ave}$	$y_{ave}$
$k_2$	[4,90]	16	77	64	82	56	49	49
$M_2$	$3 \times 3, 5 \times 5, 7 \times 7$	$5 \times 5$	$3 \times 3$	$3 \times 3$	$5 \times 5$	$3 \times 3$	$3 \times 3$	$3 \times 3$
$f^2$	$f_1, f_2, f_3, f_4$	$f_1$	$f_1$	$f_3$	$f_4$	$f_3$	$f_1$	$f_2$
$A_2$	$y_{max}, y_{ave}$	$y_{ave}$	$y_{ave}$	$y_{ave}$	$y_{max}$	$y_{ave}$	$y_{max}$	$y_{ave}$
$q_1$	[4,200]	120	140	182	193	117	169	197
$q_2$	[4,200]	24	42	58	73	112	78	134



(a) 第1层卷积核优化对比

(b) 第2层卷积核优化对比

图 6 卷积核数目优化对比结果

Figure 6 Comparison results of convolution kernel number optimization

表 2 归一化方法、全连接层对模型性能影响结果

Table 2 Effects of normalization method and full connection layer on model performance

全连接层层数	归一化方法	网络训练时间/h	准确率/%
1	不采用	0.52	83.32
	BN	0.63	90.19
	LRN	0.94	84.41
2	不采用	0.53	88.14
	BN	0.64	97.32
	LRN	1.06	87.94

由图 7 可知,对于 MNIST 数据集,DL-CNN 能够在迭代 250 次时找到全局最优解,要快于其他 3 种算法,而且收敛精度(分类准确率)达到了 97.63%,高于其他 3 种算法。对于 CIFAR-10 数据集,DL-CNN 收敛精度与 CO-Net 相当,均超过 97%,明显高于其他 2 种算法,而且 DL-CNN 仅需迭代 200 次左右就能够找到全局最优解,收敛速度更快。对于花生籽粒图像库训练集,DL-CNN 无论

是在收敛速度上还是收敛精度上都优于其他 3 种算法,且其他 3 种算法的收敛精度低于 94%,表明这 4 种算法中 DL-CNN 陷入局部最优的可能性最小。由表 3 可知,DL-CNN 无论是对于训练集还是测试集,分类准确率均高于其他 3 种算法,分类准确率几乎保持在 97%以上,对于花生籽粒图像识别准确率提高了约 5.41%~13.92%。而 CNN 准确率最低,仅为 86.03%,未达到实际应用要求。综上,DL-CNN 对单幅图像的运算速度与其他 3 种算法相当,能够以 14.1 ms 左右的速度实现对单幅图像的检测识别。

3.3 真实环境试验结果

对构建的花生籽粒筛选系统进行真实环境试验仿真。CCD 相机得到线阵图像经过处理后得到单粒花生图像,耗时约 2~3 s,然后采用 DL-CNN 进行图像筛选,最后利用空气喷枪将判定为破损花生进行剔除,最终实现花生籽粒完整性检测。对一批花生籽粒进行筛选,得到图像检测准确率为 97.63%,实际筛选出完整籽粒正确率为 91.14%,表明在真实环境下 DL-CNN 模型仍然能够以 97.63%的准确率实现对花生图像的高精度识别。考虑空

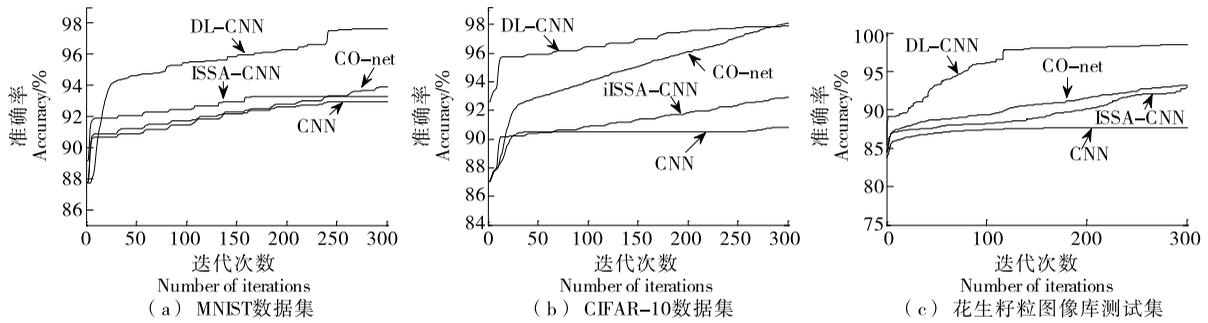


图 7 4 种模型准确率收敛曲线

Figure 7 Convergence curves of accuracy of four models

表 3 4 种模型测试集试验结果

Table 3 Experimental results of four model test sets

模型	准确率/%			平均一幅图测试时间/ms		
	MNIST	CIFAR-10	花生籽粒图像	MNIST	CIFAR-10	花生籽粒图像
DL-CNN	97.63	97.42	98.17	14.1	15.2	14.7
CO-Net	93.54	97.02	93.61	18.6	16.2	17.8
ISSA-CNN	91.87	91.32	92.17	15.7	16.1	17.7
CNN	91.05	90.11	86.03	8.4	7.7	9.1

气喷枪精度、传输带稳定性等因素,整个花生籽粒完整性筛选准确率仍然能够达到了 91.14%。因此,该系统能够很好地实现对花生籽粒的筛选,具有一定推广应用价值。

#### 4 结论

对花生籽粒筛选问题进行研究,提出了基于 DL-CNN 的花生籽粒完整性检测方案。通过引入 IDPC 算法对卷积核进行自适应压缩处理和 ISSA 对网络参数配置和网络结构进行优化,有效提升了识别准确率,而且基于 DL-CNN 的花生籽粒完整性筛选系统能够以 91.14% 的准确率实现真实环境下的花生完整性检测,证明了该方案的有效性和可行性。下一步将重点研究提升筛选系统的稳定性。

#### 参考文献

[1] 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 Zhi-heng, SONG Huan, ZHU Jiang-bo, et al. Identification algorithm and application of peanut kernel integrity based on convolution neural network[J]. Transactions of the Chinese Society of Agricultural Engineering, 2018, 34(21): 195-201.

[2] HINTON G E, OSINDERO S, TEH Y W. A fast learning algorithm for deep belief nets[J]. Neural Computation, 2014, 18(7): 1 527-1 554.

[3] HAN Dong-mei, LIU Qi-gang, FAN Wei-guo. A new image classification method using CNN transfer learning and web data augmentation[J]. Expert Systems with Applications, 2018, 95: 43-56.

[4] GENG Lei, XU Wen-long, ZHANG Fang, et al. Dried jujube classification based on double branch deep fusion convolution neural net-

work[J]. Food Science and Technology Research, 2018, 24(6): 1 007-1 015.

[5] 蓝金辉, 王迪, 申小盼. 卷积神经网络在视觉图像检测的研究进展[J]. 仪器仪表学报, 2020, 41(4): 167-182.  
 LAN Jin-hui, WANG Di, SHEN Xiao-pan. Research progress on visual image detection based on convolutional neural network[J]. Chinese Journal of Scientific Instrument, 2020, 41(4): 167-182.

[6] 马涌. 基于机器视觉的颗粒状农作物色选系统研究[D]. 哈尔滨: 哈尔滨工业大学, 2016.  
 MA Yong. Research on granular plant color selection system based on machine vision[D]. Harbin: Harbin Institute of Technology, 2016.

[7] 张忠志, 薛欢庆, 范广玲. 基于改进卷积神经网络的红枣缺陷识别[J]. 食品与机械, 2021, 37(8): 158-162, 192.  
 ZHANG Zhong-zhi, XUE Huan-qing, FAN Guang-ling. Research on jujube defect recognition method based on improved convolution neural network[J]. Food & Machinery, 2021, 37(8): 158-162, 192.

[8] 张瑞青, 李张威, 郝建军, 等. 基于迁移学习的卷积神经网络花生荚果等级图像识别[J]. 农业工程学报, 2020, 36(23): 171-180.  
 ZHANG Rui-qing, LI Zhang-wei, HAO Jian-jun, et al. Image recognition of peanut pod grades based on transfer learning with convolutional neural network[J]. Transactions of the Chinese Society of Agricultural Engineering, 2020, 36(23): 171-180.

[9] 谢为俊, 丁冶春, 王凤贺, 等. 基于卷积神经网络油茶籽完整性识别方法[J]. 农业机械学报, 2020, 51(7): 13-21.  
 XIE Wei-jun, DING Ye-chun, WANG Feng-he, et al. Integrity recognition of camellia oleifera seeds based on convolutional neural network[J]. Transactions of the Chinese Society for Agricultural Machinery, 2020, 51(7): 13-21.

(下转第 36 页)

- GAO Fu-chao, CUI Chang-ri, WEI Yue, et al. Determination of 6 furfural compounds contents in royal jelly by high performance liquid chromatography[J]. *Journal of Food Safety and Quality*, 2014, 5(11): 3 603-3 609.
- [13] 刘卫义, 沙均响, 刘建芝, 等. 一种测定酿酒辅料稻壳中糠醛含量的方法[J]. *酿酒*, 2019, 46(4): 110-111.  
LIU Wei-yi, SHA Jun-xiang, LIU Jian-zhi, et al. A method for determination of furfural in rice husk[J]. *Liquor Making*, 2019, 46(4): 110-111.
- [14] TEIXIDÓ E, MOYANO E, SANTOS F J, et al. Liquid chromatography multi-stage mass spectrometry for the analysis of 5-hydroxymethylfurfural in foods[J]. *Journal of Chromatography A*, 2008, 1 185(1): 102-108.
- [15] 张东飞, 邹金梅, 张朝晖, 等. 液相色谱—串联质谱法测定茶叶中的 11 种农药残留[J]. *食品安全质量检测学报*, 2018, 9(2): 294-298.  
ZHANG Dong-fei, ZOU Jin-mei, ZHANG Zhao-hui, et al. Determination of 11 pesticide residues in tea by liquid chromatography tandem mass spectrometry[J]. *Journal of Food Safety and Quality Inspection*, 2018, 9(2): 294-298.
- [16] 丁立平, 王丹红, 蔡春平. 液液萃取—气相色谱法直接测定罐装茶饮料中 7 种痕量卤乙酸[J]. *分析实验室*, 2021, 40(4): 449-453.  
WANG Li-ping, WANG Dan-hong, CAI Chun-ping. Direct determination of 7 kinds of haloacetic acids in canned tea drinks by liquid-liquid extraction coupled with gas chromatography[J]. *Chinese Journal of Analysis Laboratory*, 2021, 40(4): 449-453.
- [17] BARKER S A. Matrix solid-phase dispersion[J]. *Journal of Chromatography A*, 1998, 885(1/2): 115-127.
- [18] 国家市场监督管理总局. 食品安全国家标准 植物源性食品中 208 种农药及其代谢物残留量的测定 气相色谱—质谱联用法: GB 23200.113—2018[S]. 北京: 中国标准出版社, 2018: 2-3.  
State Administration for Market Regulation. National food safety standard—Determination of 208 pesticides and metabolites residues in foods of plant origin—Gas chromatography-tandem mass spectrometry method: GB 23200.113—2018 [S]. Beijing: China Standards Press, 2018: 2-3.
- [19] 朱书强. 食品安全分析中的样品前处理方法研究[D]. 兰州: 兰州大学, 2018: 10-12.  
ZHU Shu-qiang. Sample pretreatment methods in food safety analysis[D]. Lanzhou: Lanzhou University, 2018: 10-12.
- [20] 叶学敏. 新型 QuEChERS 方法在果蔬农残分析中的应用研究的应用研究[D]. 杭州: 浙江工业大学, 2020.  
YE Xue-min. Applications of novel QuEChERS protocols for analysis of pesticide residues in fruits and vegetables[D]. Hangzhou: Zhejiang University of Technology, 2020.
- [21] 全国认证认可标准化技术委员会. 实验室质量控制规范 食品理化检测: GB/T 27404—2008[S]. 北京: 中国标准出版社, 2009: 31.  
National Standardization Technical Committee on Certification and Accreditation. Criterion on quality control of laboratories—Chemical testing of food: GB/T 27404—2008[S]. Beijing: China Standards Press, 2009: 31.
- 
- (上接第 29 页)
- [10] 黄凯奇, 任伟强, 谭铁牛. 图像物体分类与检测算法综述[J]. *计算机学报*, 2014, 37(6): 1 225-1 240.  
HUANG Kai-qi, REN Wei-qiang, TAN Tie-niu. A review on image object classification and detection [J]. *Chinese Journal of Computers*, 2014, 37(6): 1 225-1 240.
- [11] 常亮, 邓小明, 周明全. 图像理解中的卷积神经网络[J]. *自动化学报*, 2016, 9(42): 1 302-1 303.  
CHANG Liang, DENG Xiao-ming, ZHOU Ming-quan. Convolutional neural networks in image understanding[J]. *Acta Automatica Sinica*, 2016, 9(42): 1 302-1 303.
- [12] 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 Bin, ZHONG Jin-ying. Review of convolution neural network[J]. *Journal of University of South China (Science and Technology)*, 2016, 30(3): 66-72.
- [13] 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 Hong-tao, ZHANG Qin-chuan. Applications of deep convolutional neural network in computer vision[J]. *Journal of Data Acquisition and Processing*, 2016, 31(1): 1-17.
- [14] 王永利, 曹江涛, 姬晓飞, 等. 基于卷积神经网络的 PCB 缺陷检测与识别算法[J]. *电子测量与仪器学报*, 2019, 33(8): 78-84.  
WANG Yong-li, CAO Jiang-tao, JI Xiao-fei, et al. PCB defect detection and recognition algorithm based on convolutional neural network[J]. *Chinese Journal of Scientific Instrument*, 2019, 33(8): 78-84.
- [15] CHEN Y S, LIN Z H, ZHAO X, et al. Deep learning-based classification of hyperspectral data[J]. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 2014, 7(6): 2 094-2 107.
- [16] ROSEBROCK A. Deep learning for computer vision with python-starter Bundle[M]. Baltimore: Py Image Search, 2017: 189-190.
- [17] RODRIGUEZ A, LAIO A. Clustering by fast search and find of density peaks[J]. *Science*, 2014, 344(6 191): 1 492-1 496.
- [18] CHEN Y, TANG S, PEI S, et al. DHeat: A density heat-based algorithm for clustering with effective radius[J]. *IEEE Transactions on Systems Man & Cybernetics Systems*, 2018, 48(4): 649-660.
- [19] CHEN Ye-wang, TANG Sheng-yu, ZHOU Li-da, et al. Decentralized clustering by finding loose and distributed density cores[J]. *Information Sciences: An International Journal*, 2018, 433/434: 510-526.
- [20] SEYEDALI M, AMIR H G, SEYEDEH Z M, et al. Salp Swarm Algorithm: A bio-inspired optimizer for engineering design problems[J]. *Advances in Engineering Software*, 2017, 114: 163-191.