

文章编号: 2095-2163(2023)12-0133-06

中图分类号: TP391

文献标志码: A

# 多特征融合的轻量级马铃薯分类研究

郭亚齐<sup>1</sup>, 刘成忠<sup>1</sup>, 韩俊英<sup>1</sup>, 冯全<sup>2</sup>, 张峰<sup>3</sup>, 罗嘉珂<sup>1</sup>

(1 甘肃农业大学 信息科学技术学院, 兰州 730070; 2 甘肃农业大学 机电工程学院, 兰州 730070;

3 甘肃农业大学 农学院/干旱生境作物学国家重点实验室, 兰州 730070)

**摘要:** 针对复杂环境下田间马铃薯品种识别精度不高, 模型体积较大等问题, 提出了一种基于多特征融合的轻量级卷积神经网络 LRnet。LRnet 网络将 ShuffleNetV2 和 MobileNetV2 裁剪后作为双分支网络结构的主干, 分别提取马铃薯叶片和马铃薯块茎的图像特征, 然后通过提出的融合策略将特征进行融合, 从而对马铃薯的 29 个品种进行分类。实验结果表明, 本文提出的 LRnet 相比 ShuffleNetV2、MobileNetV2 和 ResNet 不仅显著提高了马铃薯品种的分类精确度, 而且模型小, 易于端侧部署, 同时也表明了多部位特征融合方法可以显著提高物种的识别精度, 为马铃薯的品种鉴定提供了技术参考, 其融合策略也可为相关研究人员提供一定的研究思路。

**关键词:** 马铃薯分类; 神经网络; 品种鉴定; 多特征融合

## Lightweight potato classification method based on multi-feature fusion

GUO Yaqi<sup>1</sup>, LIU Chengzhong<sup>1</sup>, HAN Junying<sup>1</sup>, FENG Quan<sup>2</sup>, ZHANG Feng<sup>3</sup>, LUO Jiako<sup>1</sup>

(1 College of Information Science and Technology, Gansu Agricultural University, Lanzhou 730070, China;

2 College of Mechanical and Electrical Engineering, Gansu Agricultural University, Lanzhou 730070, China;

3 College of Agronomy, Gansu Agricultural University/State Key Laboratory of Aridland Crop Science, Lanzhou 730070, China)

**Abstract:** Aiming at the problems of low accuracy and large model size of interfield potato varieties in complex environment, a lightweight convolutional neural network (LRnet) based on multi-feature fusion was proposed. LRnet network cuts ShuffleNetV2 and MobileNetV2 as the backbone of the two-branch network structure, extracts the image features of potato leaves and potato tubers respectively, and then fuses the features through the proposed fusion strategy to classify the 29 varieties of potato. Experimental results show that compared with ShuffleNetV2, MobileNetV2 and ResNet using potato leaves for classification, the proposed LRnet not only significantly improves the classification accuracy of potato varieties, but also has a small model and is easy to be deployed end-to-end. This method also shows that the multi-part feature fusion method can significantly improve the accuracy of species identification, which provides technical reference for potato variety identification, and its fusion strategy can also provide certain research ideas for relevant researchers.

**Key words:** classification of potato; neural network; variety identification; multi-feature fusion

## 0 引言

马铃薯是世界上仅次于水稻、小麦、玉米的第四大粮食作物, 凭借其耐旱、耐寒、高产、高营养及口感丰富的优势, 逐渐受到了人们的关注。马铃薯品种的精确判断是马铃薯产业高质量发展的基石, 如何高效、简单、准确的鉴定马铃薯品种是一个重要问题。由于马铃薯品种的不断增多, 通过专家进行品

种鉴定是一项极其耗时且成本高昂的工作。目前, 国内外许多学者将卷积神经网络应用于农业方面取得了较高的检测精度, 但复杂环境下的图像特征提取仍有难度。彭红星等<sup>[1]</sup>提出了一种多尺度融合模块 LMFF, 通过不同大小的卷积核获取图像不同尺度的感受野, 并将其进行融合, 提高了模型对害虫的识别精度。目前许多研究也采用这种方式进行特征融合。不同尺度感受野的融合可以增加特征图的

**基金项目:** 甘肃省高等学校创新基金项目(2021A-056); 甘肃省高等学校产业支撑计划项目(2021CYZC-57); 国家自然科学基金(32160421)。

**作者简介:** 郭亚齐(1994-), 男, 硕士研究生, 主要研究方向: 农业图像分类研究。

**通讯作者:** 刘成忠(1969-), 男, 硕士, 教授, 硕士生导师, 主要研究方向: 智能决策支持系统、农业信息、智能决策支持系统等。Email: liucz@gsau.edu.cn

收稿日期: 2022-12-13

信息量,但获取的特征信息都来自相同的输入图像,而图像在输入时已经进行了一定的放缩,受限于输入图像本身,网络精度与稳定性仍有待提高。徐硕等<sup>[2]</sup>提出了一种融合外观特征和姿态特征的步态识别网络,该网络能够融合外观特征和姿态特征对步态进行识别,实验结果表明步态识别的稳定性和准确性得到较大提高。周晨怡等<sup>[3]</sup>提出一种基于卷积神经网络的多模态生物特征融合模型,将人物的虹膜、手掌静脉和手指静脉的多模态生物特征识别图像特征进行融合,极大程度提高了模型的泛化性和检索精度。受此启发本文提出了一种融合了马铃薯叶片特征和块茎特征的轻量级卷积神经网络 LRnet,该网络能够同时提取马铃薯叶片特征和块茎特征,扩展了马铃薯的特征表示。以自建的马铃薯图像数据库为研究对象,评价 LRnet 网络的有效性。该模型具有参数量少,体积小等优势,方便端侧部署,旨在为农业品种鉴定等方面提供技术支持。

## 1 材料与方法

### 1.1 马铃薯图像数据集






























本文的马铃薯图像数据采集于甘肃省金昌市永登县(北纬 38°25';东经 101°91';大陆性气候,年平均

均气温 5.9 ℃,平均海拔 2 105 m;年平均降雨量 300 mm;年平均日照量:2 659 h,平均无霜期 121 d)。收集的马铃薯品种是甘肃省常食用品种及衍生的试验品种,品种之间图片差异较小。本研究按照 7:3 的比例分别在两块试验田上进行数据采集。采用尼康 COOLPIXB700 数码相机的自动模式,对马铃薯叶片及块茎进行采集时将 ISO 设定为 1 600,最低拍摄快门设定为 1/30 s,图片大小为 2 272×1 704 像素。马铃薯图像数据收集时间为 2021 年 8-9 月,属于自然田间拍摄。在拍摄过程中,摄像机与样本的距离保持在 20±0.1 cm,同时拍摄环境包括各种天气条件和自然光照条件,增加了数据集的健壮性。

选取 29 个马铃薯品种作为本次研究的分类对象,将马铃薯品种按照表 1 进行数字编号。叶片数据共收集了 22 897 张图像,每个品种约 700 张;每个品种大约选择了 100 株;块茎数据集 4 500 张图片,每个品种约 150 张;从顶部、侧面、整体和部分等多个角度对其拍摄图像。图像的命名方式为:马铃薯品种编号+植株编号+视图编号。例如,"1\_1\_1.jpg"代表马铃薯第一个品种的第一株第一视图照片。图 1 中 a-e 分别表示马铃薯叶片和块茎的 5 种视图角。

表 1 本文采用的 29 个马铃薯品种

Table 1 The 29 potato varieties used in this paper

编号	名称	图像	编号	名称	图像	编号	名称	图像
1	陇南 15-8		11	0713-60		21	L1632-5	
2	0726-205		12	201518-53		22	D0903-11	
3	庄 15-11-25		13	凯薯 6 号		23	凯薯 3 号	
4	ly1520-9		14	天 14-4-108		24	0711-67	
5	L1630-6		15	凯薯 5 号		25	L1591-1	
6	D1004-318		16	L14119-1		26	陇薯 6 号(ck)	
7	L16101-11		17	庄 13-5-21		27	14w-5	
8	D0904-11		18	天 14-6-46		28	D10Y1-6	
9	庄 15-1-16		19	天 14-22-28		29	凯薯 4 号	
10	天 14-20-18		20	0726-126				



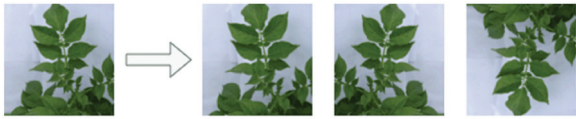
(a) 正视图 (b) 左视图 (c) 右视图 (d) 后视图 (e) 俯视图

图 1 马铃薯数据集不同角度图像

Fig. 1 Potato dataset images from different angles

### 1.2 数据集预处理

数据进行预处理对保持算法的一致性和平滑运行至关重要,在输入图像中添加噪声和模糊可以使深度神经网络学习到更加健壮的特征<sup>[4]</sup>。为了保证模型的稳定性,提高模型的鲁棒性,本研究对数据集进行了增强。图 2 以叶片图片为例,展示了几种数据增强方法。最后将图像放缩到模型设计的输入大小进行网络加载;对于测试集,本研究仅对数据进行放缩,确保能够正常测试模型。



(a) 原始图片 (b) 随机缩放 (c) 水平翻转 (d) 为垂直翻转

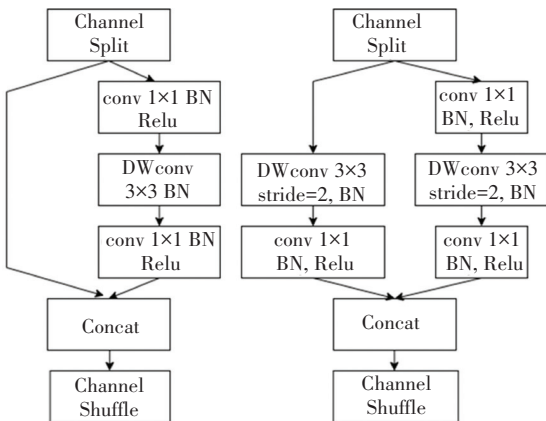
图 2 数据增强策略

Fig. 2 Data augmentation strategy

### 1.3 深度学习模型

#### 1.3.1 叶片特征提取 Lnet 分支

ShuffleNetV1<sup>[5]</sup>是一种轻量级的卷积网络结构,其中使用了深度可分离卷积,降低了参数量和计算量,同时采用 Channel Shuffle 操作,将组间的信息进行交换。如图 3 所示,ShuffleNetV2<sup>[6]</sup>对 ShuffleNetV1 进行了优化,将 ShuffleNetV1 中的 add 操作取消,全部使用 concat<sup>[7]</sup>操作进行特征融合。



(a) 基本单元 (b) 下采样单元

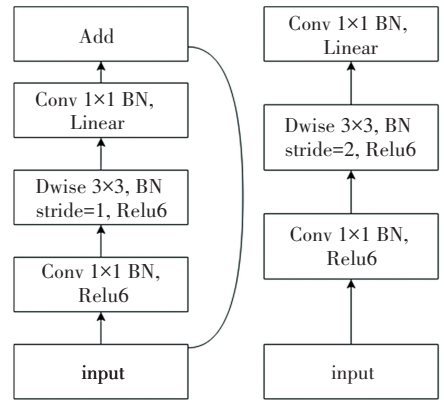
图 3 ShuffleNetV2 的基本结构

Fig. 3 The basic structure of ShuffleNetV2

Lnet 采用 ShuffleNetV2 作为提取马铃薯块茎特征的网络分支,同时对网络进行了量化裁剪,去除了网络最后的全连接层,增加自适应平均池化层,方便后续的特征融合,并将 Relu6 激活函数更换为效果更好的 swish 函数<sup>[8]</sup>以提高模型的精确度。

#### 1.3.2 块茎特征提取 Rnet 分支

MobileNetV1<sup>[9]</sup>是 Google 发布的轻量级神经网络。也使用了深度可分离卷积,节省了参数量和计算量。如图 4 所示,MobileNetV2<sup>[10]</sup>同样使用了深度可分离结构,同时借鉴了 ResNet<sup>[11]</sup>的思想,提出了倒残差结构作为网络基本结构,使用 Add 操作进行特征复用,缓解特征退化问题。



(a) 基本单元

(b) 下采样单元

图 4 MobileNetV2 的基本结构

Fig. 4 The basic structure of MobileNetV2

Rnet 分支网络采用 MobileNetV2 作为提取马铃薯叶片特征的分支,同时在此基础上进行了量化裁剪,并将 Relu6 激活函数更换为效果更好的 swish。同时,为了保证与 Lnet 分支提取的叶片特征融合,去除了网络最后的全连接层,增加自适应平均池化层。

#### 1.3.3 融合策略

研究提出了一种多特征融合架构,从马铃薯的不同部位提取马铃薯特征信息。首先,使用 Lnet 分支提取叶片图像特征,在 Lnet 中使用 Concat 进行多尺度的特征融合充分提取叶片信息;使用 Rnet 分支提取块茎图像特征,在 Rnet 中使用 Add 操作进行特征的复用;两分支提取的特征经过自适应平均池化层后,通过 Concat 进行拼接,然后经过全连层,最后通过 SoftMax<sup>[12]</sup>进行分类。此外,本研究还提出了另外两种融合策略,对比 LRnet 的融合位置,本研究将其命名为早融合和晚融合。

如图 5 所示,叶片和块茎图像分别通过 Lnet 和 Rnet 得到马铃薯叶片特征  $f_1$  和马铃薯块茎特征  $f_2$ ,



经过自适应平均池化后,将  $f_1$  和  $f_2$  级联得到融合特征  $f$ 。采用级联融合的方法将两个输出进行叠加级联操作,只对特征向量的两部分进行合并,不进行其他操作。因此,叠加级联后的特征向量  $f$  包含了两个分支计算出的所有特征信息。这些信息是从两种不同的图像中得到的,使得提取马铃薯特征表示更加全面;接下来对特征  $f$  继续进行全连接操作,然后进行分类。

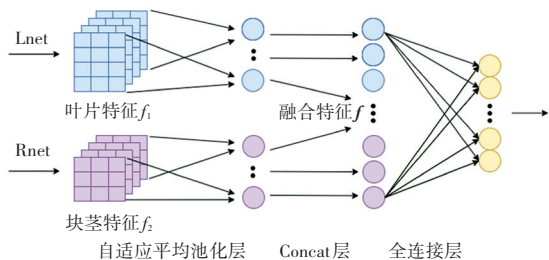


图5 LRnet融合策略

Fig. 5 LRnet fusion strategy

早融合的过程如图6所示。马铃薯叶片和块茎图像分别送入 Lnet 和 Rnet 分支网络进行特征提取,将提取的特征分别经过自适应平均池化层得到  $f_1$  和  $f_2$ 。通过 Add 操作将  $f_1$  和  $f_2$  进行特征融合得到  $f$ ,融合后特征图的大小和维度(即  $f$ ) 在融合后保持不变,然后继续进行卷积操作。这种网络结构很大程度上体现了两部分数据的独立性和相关性。在此过程中,两个模型的特征信息可以通过卷积核进行融合,同时减少了训练参数,提高了训练速度。

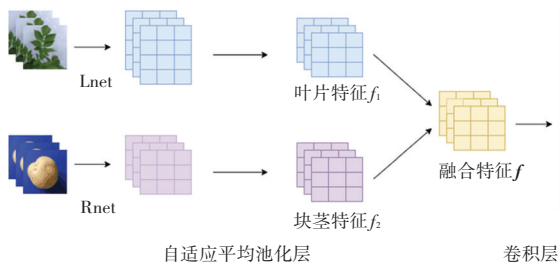


图6 早融合策略

Fig. 6 Early fusion policy

晚融合的过程如图7所示。在此情况下,提出了一种最优权值法来结合单分支的预测概率。在 SoftMax 层之后,分别从 Lnet 和 Rnet 网络输出每个图像的预测概率  $l$  和  $r$ 。最后的融合概率  $y$  是通过将两个分支的输出概率,以一定的权重组合来计算的。融合权值的详细描述如式(1)~式(3)所示。

$$f_l = \frac{t_l}{t_l + t_r} \tag{1}$$

$$f_r = \frac{t_r}{t_l + t_r} \tag{2}$$

$$y_i = f_l \times l_i + f_r \times r_i \tag{3}$$

式中:  $f_l$  和  $f_r$  分别为马铃薯叶片和马铃薯块茎的融合权值,  $t_l$  和  $t_r$  分别为使用马铃薯叶片和马铃薯块茎的单个分支网络的测试精度。在式(3)中,  $l_i$  和  $r_i$  分别表示两个分支输出的叶片分类品种和块茎分类品种的概率。将两种分类结果与融合权重  $f_l$  和  $f_r$  进行融合,即可计算出最终结果。

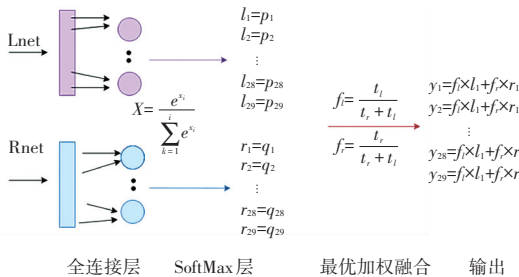


图7 晚融合策略

Fig. 7 Early fusion policy

### 1.3.4 LRnet 网络

本研究提出了一种基于多特征融合的轻量级卷积神经网络 LRnet,该网络将马铃薯叶片和块茎图像特征进行融合,实现了 29 个马铃薯品种的种间精准分类。如图8所示,输入图像由两部分组成,将马铃薯叶片的图像和相同大小的块茎图片同时输入,其块茎与叶片之间一一对应,保证在训练中使用相同的标签。该网络同时提取同株马铃薯的叶片及其块茎的特征,便于叶片与块茎的特征融合。叶片和块茎图片分别输入到双分支网络当中进行图像特征提取,然后经过自适应平均池化层,通过 Concat 层将叶片和块茎的特征进行拼接,经过全连接层得到特征张量,最后通过 SoftMax 得到图像的种类。

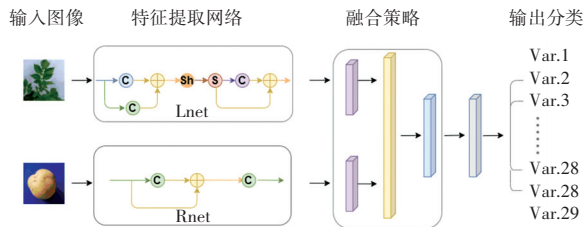


图8 LRnet核心结构

Fig. 8 LRnet core structure

图中, C 表示卷积操作; + 表示特征拼接; Sh 表示通道打乱; S 表示通道拆分。

## 2 实验设计与结果分析

### 2.1 对比实验

为了验证 LRnet 的先进性, 实验过程中使用了相同的训练配置, 并对单叶片数据使用 ShufflenetV2、MobileNetV2 和 ResNet-50 进行训练, 将结果与使用 LRnet 进行训练的结果进行对比。此外, 对于融合策略也进行了对比实验, 验证了 LRnet 的有效性。

### 2.2 模型训练

数据按照 8 : 2 分为训练集和测试集, 其中训练集包含 15 265 幅图像, 测试集包含 3 816 幅图像。研究使用深度学习框架 PaddlePaddle 2.3, 显卡为 Nvidia Tesla V100, 显示内存为 32 GB。采用 PiecewiseDecay<sup>[13]</sup> 用于提供分段设置学习率, 使用 Adam 优化器<sup>[14]</sup> 动态调整每个参数的学习率。加入了 L2 正则化<sup>[15]</sup> 实现 L2 权重衰减, 有助于防止模型对训练数据过拟合; 运用迁移学习<sup>[16]</sup> 方法, 初始化部分权重参数, 加快模型的收敛速度。经过多次调试, 最终模型训练参数见表 2。

表 2 模型的训练配置

Table 2 Training configuration of the model

训练配置	配置参数
Optimizer	Adam
Loss function	CrossEntropyLoss
Batch size	32
learning rate	[0.005, 0.001, 0.000 5, 0.000 1, 0.000 05]
num_workers	4
epochs	100

### 2.3 模型评价

为了在马铃薯分类任务中评价 LRnet 的性能, 本研究采用相同的训练配置, 将训练后的 LRnet 与单一图像模型 ShuffleNetV2、MobileNetV2、ResNet-50 进行比较, 同时也对比了不同的融合策略训练的结果, 比较结果如图 9 所示。

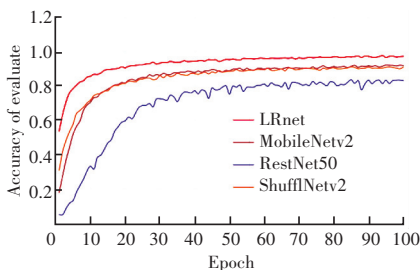


图 9 LRnet 及对比模型在测试集上的精确度

Fig. 9 Accuracy of LRnet and comparison models on the test set

图 9 表示在 100 个 epoch 中每次迭代模型的测试集精确度, 在相同的配置下, 随着 epoch 的增加, LRnet 无论在初始精确度、精确度波动、收敛速度, 还是最终达到的精确度都优于 MobileNetV2、ShuffleNetV2 和 ResNet5 模型, 表明了 LRnet 模型的优越性。表 3 是以上模型的性能对比, 从模型的数量看 MobileNetV2、ShuffleNetV2 和 ResNet50 分别具有 13.5 M、8.75 M、97.69 M, 而 LRnet 只有 7.6 M。本研究将 LRnet 与另外两种融合策略做了对比实验, 其中 LRnet 精确度为 98.34%, 早融合精确度为 96.80%, 晚融合精确度为 96.80%, 虽然略低于 LRnet 网络, 但均高于其他模型, 证明了采用多特征融合的可行性。

表 3 模型数据对比

Table 3 Comparison of model data

模型	参数量/M	精确度/%
LRnet	7.60	98.34
LRnet_early	25.55	96.80
LRnet_late	11.75	96.80
ShuffleNetV2	8.75	91.11
MobileNetV2	13.50	90.64
RestNet50	97.69	84.90

## 3 结束语

针对田间自然环境下的马铃薯品种鉴定, 提出一种多特征融合的马铃薯分类方法。该方法能够同时获取马铃薯叶片图像和块茎图像的特征, 实现对马铃薯种间的高精准分类。以图像数据为基础, 基于 MobileNetV2 和 ShuffleNetV2 模型构建双分支网络 LRnet, 分别对马铃薯叶片及块茎图片进行特征提取, 通过提出的融合策略进行特征融合, 提高了模型的精确度和稳定性。训练数据集包含马铃薯叶片和块茎的图像, 解决了图像特征单一问题, 保证了模型的可靠性。与其他经典网络相比, LRnet 在精确度和稳定性方面表现出明显的优势, 表明通过融合策略对马铃薯叶片和块茎进行特征融合可以提高网络的分类能力。此外, 本研究还提出了另外两种融合网络, 在网络中其具有不同的融合位置。结果表明, LRnet 相对于其他两种方法具有明显的优势, 说明图像特征在网络中的融合位置是影响网络分类性能的一个重要因素。采用测试数据集对模型的泛化能力进行了测试, 最终耦合识别率为 98.34%, 高于单叶片图像模型的识别率, 表明提出耦合模型 LRnet 具有较强的精确度和鲁棒性, 为深度学习模型应用于田间农作物分类提供思路。