

文章编号:2095-7386(2023)06-0057-06

DOI:10.3969/j.issn.2095-7386.2023.06.008

基于卷积神经网络和迁移学习的番茄成熟度检测

吴俊峰¹, 杨柳¹, 崔波¹, 王志诚¹, 徐子龙¹, 宋少云^{1,2}, 张永林^{1,2}

(1. 武汉轻工大学 机械工程学院, 武汉 430023;

2. 湖北省粮油机械工程技术研究中心, 武汉 430023)

摘要:为解决现有番茄成熟度检测模型存在的计算复杂且精度低的问题,以普罗旺斯番茄为样品,利用 HSV 通道数据进行分析,将样本图片划分为绿熟期、变色期和红熟期,建立不同成熟度番茄的数据集。基于迁移学习,对 MobilenetV2、InceptionV3 和 VGG19 三种卷积神经网络模型的超参数进行了优化,建立了基于番茄成熟度的检测模型。进一步分析了 SGD、Adam、Adagrad 优化算法和不同训练轮数的检测效果,发现 MobilenetV2 的表现最好,在优化器为 Adam、训练 100 轮时准确率可达 99.05%,损失值为 0.038 4,满足高精度检测。检测方法能在低损失值下获得更高的检测准确率,可为番茄成熟度检测研究提供参考。

关键词:番茄;迁移学习;卷积神经网络;成熟度分类

中图分类号:TS 212.3

文献标识码:A

Tomato maturity detection based on convolutional neural network and transfer learning

WU Junfeng¹, YANG Liu¹, CUI Bo¹, WANG Zhicheng¹, XU Zilong¹, SONG Shaoyun^{1,2}, ZHANG Yonglin^{1,2}

(School of Mechanical Engineering, Wuhan Polytechnic University, Wuhan, 430023, China)

Abstract: In order to solve the problems of complicated calculation and low accuracy of existing tomato ripener detection models, this paper takes Provence tomatoes as samples and conducts analysis using HSV channel data. The sample images are divided into green ripening, discoloration, and red ripening stages, and a dataset of tomatoes with different maturity levels is established. Based on transfer learning, the hyper parameters of three convolutional neural network models, MobilenetV2, InceptionV3, and VGG19, were optimized, and a detection model based on tomato maturity was established. The detection effects of SGD, Adam and Adagrad optimization algorithms and different training rounds were further analyzed, and it was found that MobilenetV2 had the best performance, with an accuracy of 99.05% and a loss value of 0.038 4 when the optimizer was Adam and 100 training rounds, meeting the high efficiency maturity detection. The detection method can obtain higher detection accuracy under low loss value, and provide reference for the detection of tomato maturity.

Key words: tomato; transfer learning; convolutional neural network; maturity classification

收稿日期:2023-10-21.

作者简介:吴俊峰(2000—),男,硕士研究生,E-mail:2365795728@qq.com.

通信作者:杨柳(1989—),男,副教授,博士,E-mail:yangliuivctry@163.com.

基金项目:湖北省自然科学基金(2022CFB944);湖北省教育厅科研项目(Q20211609);武汉轻工大学校杰出青年科研项目(2020J06).

1 引言

番茄是一种一年生的草本植物,是主要的果蔬作物之一,深受广大消费者喜爱^[1-3]。近年来,随着生活水平的提高,人们对番茄的需求以及品质的要求也越来越高。番茄成熟度检测是保证番茄品质的重要环节,因此在番茄的采摘运输过程中,需要对不同成熟度的番茄进行分类。然而,传统靠人为的对不同成熟度番茄进行分类的方法误差大、效率低,因此开发出误差小且效率高的方法对番茄成熟度进行检测十分重要。

近年来,应用于目标检测和分类的算法层出不穷,特别是机器视觉技术越来越多地被应用在了农业领域^[4-9]。机器视觉技术主要包括了深度学习和迁移学习,伴随着二者的不断发展,出现了各种各样的卷积神经网络算法,有着各自的优势。传统的深度学习需要标注大量数据来对模型进行训练,耗费时间巨大,而迁移学习可以将预训练好的模型移植到特定的领域,训练所需的数据更少,训练模型所花的时间也更少。张莉等^[10]、伍莹芮等^[11]、刘国旭等^[12]学者采用了传统的深度学习方法对番茄成熟度进行了检测,效率较低、训练过程繁琐,而采用迁移学习方法能够克服这些缺点,对数据需求较小、硬件要求较低,学习速度快^[13]。

为了能够获得更高的番茄成熟度检测效率以及准确率,笔者以普罗旺斯番茄为样品,利用 HSV 通道数据进行了分析,将样本图片划分为三个时期,建立了不同成熟度番茄的数据集。基于迁移学习,对 MobilenetV2、InceptionV3 和 VGG19 三种卷积神经网络模型的超参数进行了优化,建立了检测模型,并选择不同的优化器进行不同训练轮数的检测。该方法以期在较低损失值的情况下获得更高的检测准确率,为番茄成熟度检测研究提供参考。

2 材料与方法

2.1 图像采集和预处理

番茄样品是来自中国湖北省襄阳市的普罗旺斯番茄。使用海康工业相机 MV-CS050-10GC-PRO 采集了 800 个不同成熟期的番茄样品照片。根据 GH/T 1193—2021 将番茄样品划分为绿熟期、变色期和红熟期。绿熟期的番茄表面有光泽,从绿色变

成白绿色,此时期可以进行人工催熟、采摘和储存;变色期是从绿熟期到红熟期的过渡阶段,番茄脐部周围开始出现黄色或浅红色晕点,果面红色不到 1/10;红熟期的果面红色高于 1/10。

使用 HSV 颜色空间通道数据将数据集划分为绿熟期、变色期和红熟期。HSV 颜色空间中 H 通道是以角度测量的色调,S 通道是饱和度,V 通道表示颜色的亮度^[14]。与 RGB 相比,HSV 可以更好地反映图像的颜色信息,并具有更显著的特征。如表 1 所示,将采集到的 540 张图像中的 435 张作为训练集,剩下 105 张作为测试集用于训练。

表 1 番茄成熟度数据集/张

Table 1 Tomato maturity dataset

番茄成熟度	训练集	验证集	合计
绿熟期	145	35	180
变色期	145	35	180
红熟期	145	35	180

2.2 设备及模型原理

2.2.1 设备

实验中使用的设备是一个实验平台,采用 Intel i5-8300H 处理器、16 GB RAM 和 NVIDIA GeForce GTX 1060。软件配置环境为 PyCharm2022、Anaconda3、tensorflow2.3 和 Python3.7。

2.2.2 模型选择

Mobilenet 网络是 Google 团队在 2017 年提出的专注移动端或者嵌入式设备中的轻量级 CNN 网络^[15],相比于其他模型,最大的亮点就是采用深度可分离卷积,大大减小了运算量和参数数量。迁移学习时 MobilenetV2 的网络结构参数见表 2。

表 2 迁移学习时 MobilenetV2 的结构参数

Table 2 Migration learning MobilenetV2 structural parameters

类型	输出大小	参数数量
归一化	224×224×3	0
Mbilenetv2 模型	7×7×1 280	2 257 984
全局平均池化	1 280	0
全连接层	3	3 843

InceptionV3 模型在 2014 年由 Google 团队提出^[16],它的优势就是引入了 Inception 结构,添加了两个辅助分类器帮忙训练。迁移学习时的网络结构

如表 3 所示。

表 3 迁移学习时 InceptionV3 的结构参数

Table 3 Migration learning InceptionV3 structural parameters

类型	输出大小	参数数量
归一化	$224 \times 224 \times 3$	0
InceptionV3 模型	$5 \times 5 \times 2\ 048$	21 802 784
全局平均池化	2 048	0
全连接层	3	6 147

VGG 模型于 2014 年由牛津大学著名研究组 Visual Geometry Group 提出^[17],此模型的特点是绝大多数卷积核被替换成了 3×3 ,只有极少数为 1×1 。迁移学习时的网络结构参数如表 4。

表 4 迁移学习时 VGG19 模型的结构参数

Table 4 Migration learning VGG19 structural parameters

类型	输出大小	参数数量
归一化	$224 \times 224 \times 3$	0
VGG19 模型	$7 \times 7 \times 512$	20 024 384
全局平均池化	512	0
全连接层	3	1 539

2.3 优化器的选择

在深度学习中,优化器主要分为基于梯度下降和自适应梯度两大类。在 2010 年前大多数采用基于梯度下降的优化器 SGD,2010 年之后 Adagrad 和 Adam 等自适应梯度的优化器越来越流行。然而,最近的趋势表明,部分研究者已经开始恢复使用 SGD。笔者讨论了 SGD、Adagrad 和 Adam 三种优化器对模型训练损失值和准确率的影响。

SGD 是神经网络训练的基本算法^[18],它能计算网络误差,可在每次批量训练时对误差进行反向传播,并基于一阶梯度信息更新参数。Adagrad 针对 SGD 存在的问题,在 SGD 的基础上引入了二阶动量,能够对不同参数使用不同的学习率进行更新,对于梯度较大的参数,学习率就会变得较小,而对于梯度较小的参数,学习率就会变得较大。Adam 优化器结合了 Momentum 和 RMSProp 算法,引入了 Momentum 的一阶动量和 RMSProp 的二阶动量,可以使收敛速度更快和波动更小,然后在此基础上添加两个校正项,使参数能够自我更新^[19]。

2.4 算法流程

使用迁移学习的方法来训练上述三个模型。迁

移学习的参数设置中 batch size 是指模型训练过程中每次迭代的样本数量,考虑到数据集的大小,将 batch size 选择为 32。学习率是控制模型学习速度的参数,学习率太大容易导致超调现象,太小会导致收敛时间更长,考虑到模型训练时的收敛性和训练时间,将学习率设置为 0.001。epoch 是训练时的次数,考虑到模型的收敛速度,将其设置为 50 轮和 100 轮。设置完参数之后,对数据集进行预处理,并将其归一化为 $224 \times 224 \times 3$ 。在冻结模型主干参数的情况下,对主干模型的输出进行全局平均池化,利用 Softmax 函数将全连接层映射到 (0,1) 之间的分布中,得到番茄图像成熟度预测的类别。算法过程如图 1 所示。

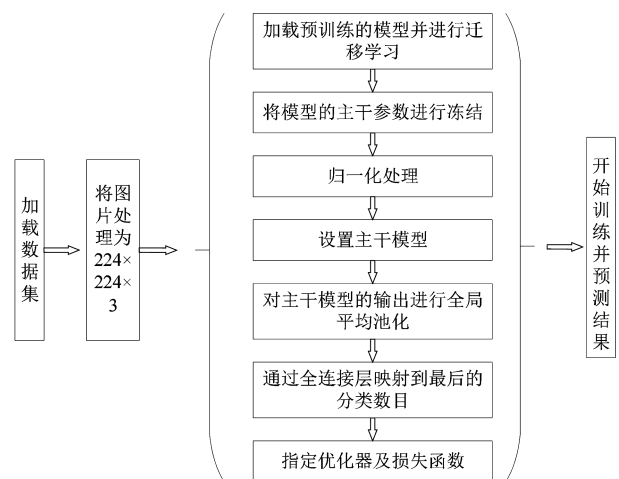


图 1 迁移学习的算法流程

Fig. 1 Flow chart of transfer learning

3 结果与分析

3.1 采用不同优化器的 MobilenetV2 模型

由图 2 可知,使用 SGD 和 Adam 优化器时, MobilenetV2 模型的损失和精度都在 20 个 epoch 内收敛,而使用 Adagrad 优化器,该模型直到大约 50 个 epoch 才接近完成,这表明 SGD 和 Adam 优化器在收敛速度上表现得更好。Adam 优化器的损失最小,从 0.83 逐渐收敛到约 0.03,而 Adagrad 优化器的损失值从 0.9 逐渐收敛到 0.2。SGD 优化器整体上比较平滑,从一开始的 0.5 逐渐收敛到 0.2。就验证集的准确性而言,SGD 优化器和 Adam 优化器在不到 10 个 epoch 内都达到了 99.05% 的准确率。就准确率和损失值而言,Adam 优化器在 MobilenetV2 模型中总能获得更好的结果,Adam 优化器比 SGD 和 Adagrad 总体性能要好。

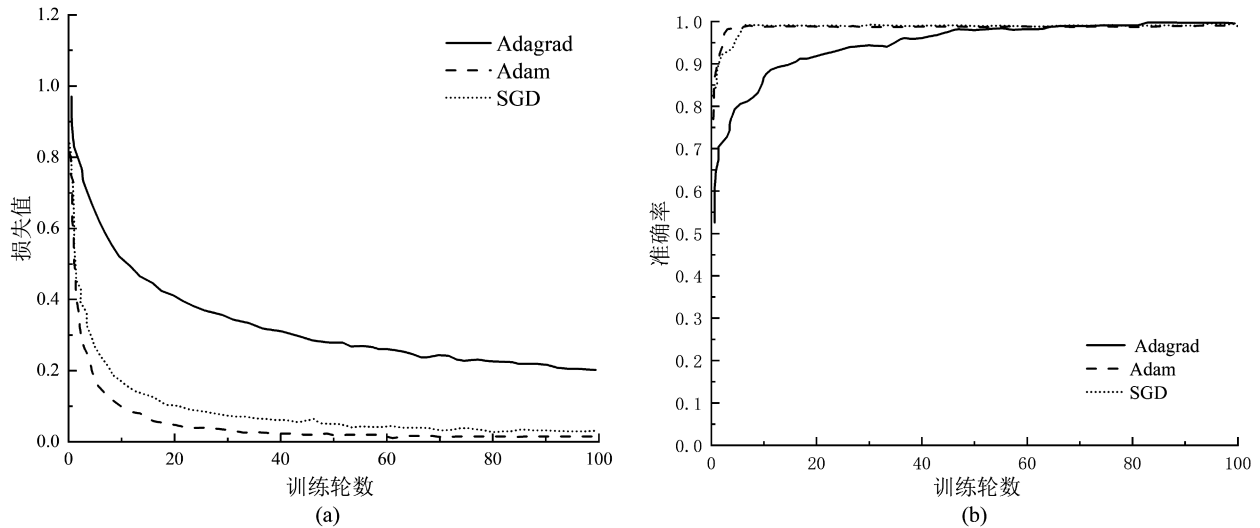


图2 不同优化器下 MobilenetV2 模型训练过程中的损失值和准确率

Fig. 2 Loss and accuracy of MobilenetV2 train process with different optimizers

3.2 采用不同优化器的 InceptionV3 模型

图3说明了 Adam 优化器的损失值收敛得更快,约为 0.05, Adagrad 损失值约为 0.3。开始训练时采用 SGD 优化器的准确率仅为 52%,并且在训练 40 轮之前波动较大,训练 40 轮之后准确率收敛到了 99.05%。使用 Adam 优化器开始训练时准确

率为 75%,然后其准确率快速收敛到 99.05%。使用 Adagrad 优化器训练时初始准确率为 65%,在训练 100 轮时达到 90%。在 InceptionV3 模型中, Adam 优化器总能在准确率和损失值方面取得良好结果。总的来说, Adam 优化器比 SGD 和 Adagrad 的性能要优异。

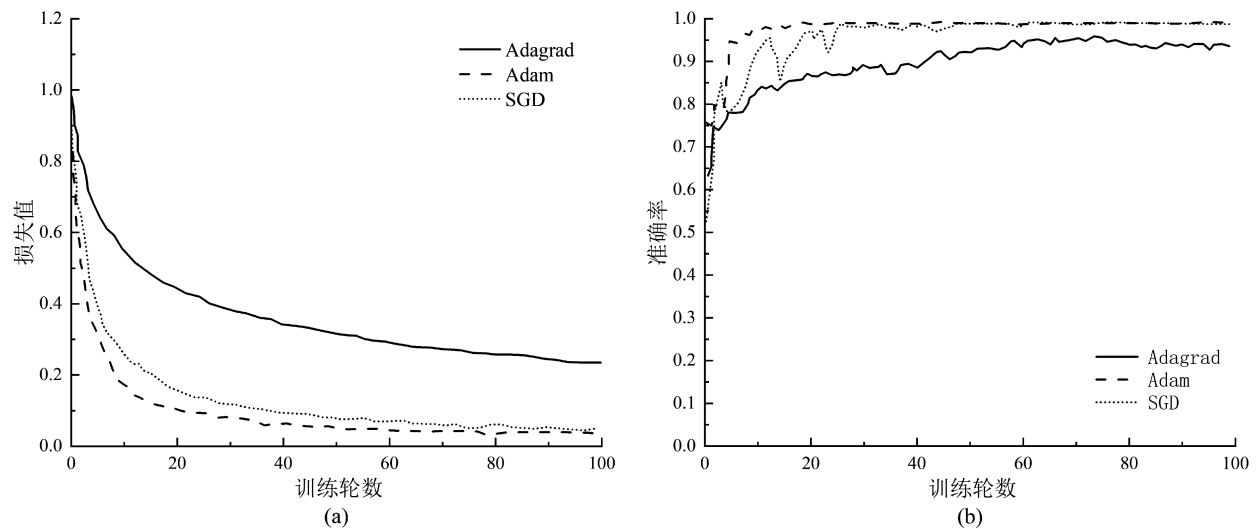


图3 不同优化器下 InceptionV3 模型训练过程中的损失值和准确率

Fig. 3 Loss and accuracy of InceptionV3 train process with different optimizers

3.3 采用不同优化器的 VGG19 模型

由图4可知 VGG 模型训练时采用 Adam 优化器损失值降低得更快,但在训练 100 轮后,损失值仍然保持为 0.3;采用 Adagrad 优化器时损失值从 1.1 缓慢下降到 1.0 左右;采用 SGD 优化器时,训练 100 轮后最终准确率只达到 86%,损失值为 0.4。

使用 Adam 优化器比 SGD 优化器更稳定,在训练 100 轮时达到 92.38% 的准确率。在使用 Adagrad 优化器时,准确率随着训练轮数的增加缓慢提高,在训练 100 轮时达到 78.10%。在 VGG19 模型中, Adam 优化器在准确率和损失值方面也可以获得良好的结果。

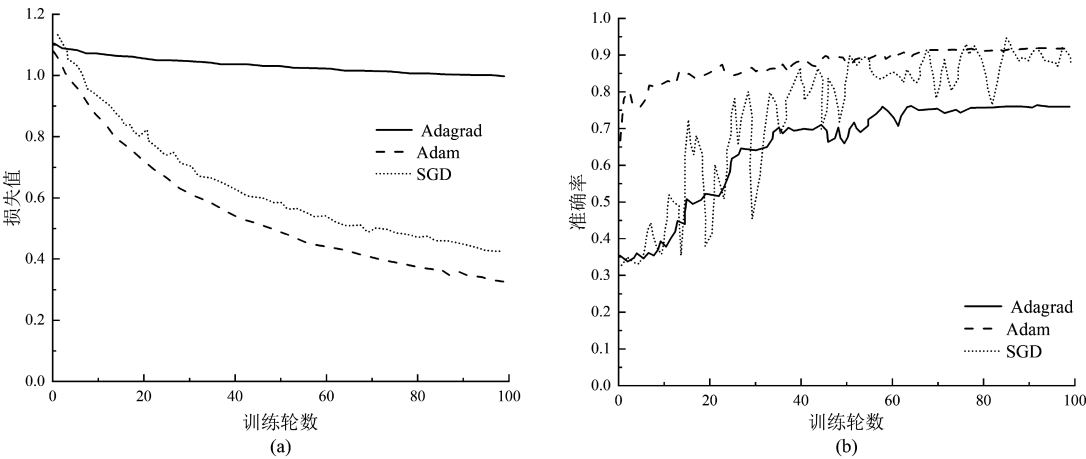


图 4 不同优化器下 VGG19 模型训练过程中的损失值和准确率
Fig 4 Loss and accuracy of VGG19 train process with different optimizers

表 5 各模型不同参数情况下的测试结果
Table 5 Test results for each model with different parameters

模型	优化器	训练轮数	损失值	准确率
MobilenetV2	SGD	50	0.058 4	99.05
		100	0.042 8	99.05
	Adam	50	0.039 1	99.05
		100	0.038 4	99.05
	Adagrad	50	0.238 2	99.05
		100	0.171 4	100
InceptionV3	SGD	50	0.116 0	98.10
		100	0.076 9	99.05
	Adam	50	0.053 9	99.05
		100	0.048 9	99.05
	Adagrad	50	0.395 4	82.86
		100	0.258 8	93.33
VGG19	SGD	50	0.867 9	74.29
		100	0.435 9	86.67
	Adam	50	0.508 0	89.52
		100	0.327 7	92.38
	Adagrad	50	1.030 7	70.48
		100	1.003 1	78.10

从表 5 中可以看出,采用 Adam 优化器的 MobilenetV2 模型在训练 50 轮时达到了 99.05% 的准确率,最小损失值为 0.038 4。SGD 和 Adagrad 的损失值分别为 0.04 和 0.2。InceptionV3 模型在训练 50 轮内,采用 SGD 优化器时实现了收敛,损失值为 0.116 0,准确率为 98.10%;使用 Adam 优化器收敛时损失值为 0.053 9,准确率为 99.05%;使用 Adagrad 优化器时,损失值为 0.395 4,准确率为 82.86%。与其他两个模型相比,VGG19 模型的整体损失值更大,精度更低。在训练 50 轮内,Adagrad 的最大损失值为 1.030 7,最小准确率为 70.48%;当使用 Adam 优化器时,损失值为 0.508 0,准确率为 89.52%;当使用 SGD 优化器时,损失值为 0.867 9,准确率为 74.29%。在训练了 100 轮时,损失值显著降低,准确性显著提高。综上,使用 Adam 优化器的 MobilenetV2 模型总体性能最好。

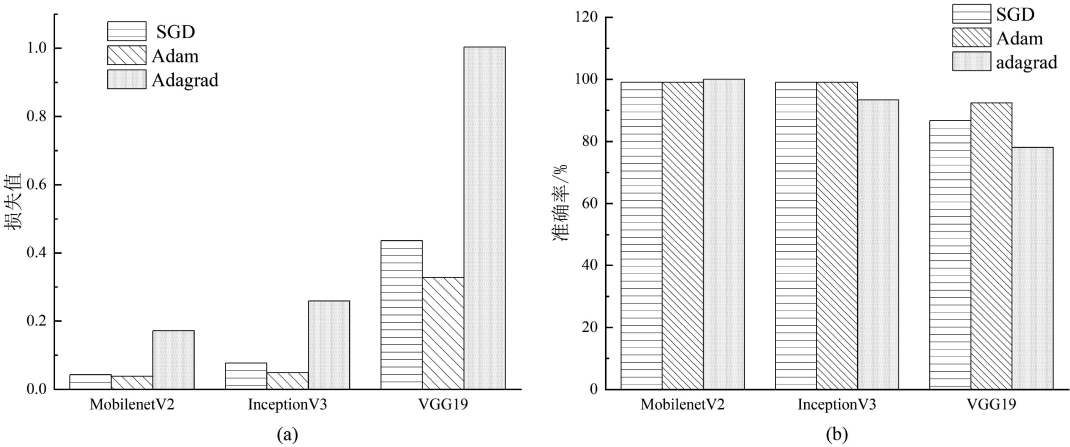


图 5 采用不同优化器训练 100 轮时三个模型的损失值和准确率
Fig. 5 Loss and accuracy of three models with different optimizers for 100 epochs

图 5 说明了 MobilenetV2、InceptionV3 和 VGG19 在分别使用 SGD、Adam 和 Adagrad 优化器训练 100 轮时的损失和准确率。与 InceptionV3 和 VGG19 相比, MobilenetV2 模型在损失和精度方面都表现出色, 三个优化器的最终损失值均低于 0.2, 准确率为 99.05%。InceptionV3 模型的损失值略高于 MobilenetV2, 当优化器为 SGD 和 Adam 时, 其精度相同, 为 99.05%, 但使用 Adagrad 优化器时, 其准确性略低。

4 结论

文中采用了一种基于迁移学习的番茄成熟度检测方法, 比较了三种卷积神经网络, 包括 MobilenetV2、InceptionV3 和 VGG19, 在模型训练过程中, 通过对超参数的调整以及对模型算法的优化, 最终获得了最优模型以及配置。结果表明, MobilenetV2 模型的总体性能优于 InceptionV3 和 VGG19 模型, 可以在最小损失值的情况下获得更高的准确率。Adam 优化器总体上比 SGD 优化器和 Adagrad 优化器表现更好, 在采用 Adam 优化器时模型训练收敛更快、更稳定。当使用 Adam 优化器来对番茄成熟度进行分类时, MobilenetV2 模型在 10 轮训练前已经达到收敛, 准确率为 99.05%, 并且在随后的训练中也非常稳定, 在 10 轮训练之前的损失值也收敛到了 0.03 左右。相比于 Arjenaki 等人提出的一种准确率为 94.5% 的传统机器视觉的番茄分拣系统^[20], 本文中的番茄成熟度检测不需要复杂的预处理, 可以自主学习特征并进行分类。

综上, 文中提出的基于迁移学习的番茄成熟度检测方法的最优配置是 MobilenetV2 模型、batch size 设置为 32、学习率设置为 0.001、优化算法替换为 Adam 优化器、epoch 选择为 100 轮。该识别方法具有较高的识别精度和良好的识别效果, 并且准确率能达到 99.05%。MobilenetV2 是一个轻量级、参数少、运算能力要求低、精度高的模型, 为番茄成熟度鉴定提供了一种有效的方法。随着科学技术的发展, 番茄生产线变得更加智能, 机器视觉将逐渐取代人工在番茄成熟度检测和采摘中的地位, 具有机器视觉系统的机器人将成为采摘成熟番茄的主流。

参考文献:

[1] Alajrami M A, Abu N, Samy S. Type of Tomato Classification Using Deep Learning[J]. International Journal of Academic Pedagogi-

cal Research, 2020, 3 (12), 21-25.

- [2] Salehi B, Sharifi R R, Sharopov F, et al. Beneficial Effects and Potential Risks of Tomato Consumption for Human Health: An Overview[J]. Nutrition, 2019, 62, 201-208.
- [3] Ali M Y, Sina A I, Khandker S S, et al. Nutritional Composition and Bioactive Compounds in Tomatoes and Their Impact on Human Health and Disease: A Review[J]. Foods, 2020, 10(1), 45-77.
- [4] 杨婷婷, 郭志勋, 雷定湘, 等. 机器视觉技术在农业中的应用[J]. 安徽农学通报, 2021, 27 (18): 110-111.
- [5] 马聪, 陈学东. 机器视觉技术在农业领域的应用及发展[J]. 宁夏农林科技, 2021, 62(07): 65-70, 2.
- [6] Da C, Figueroa H E, Fracarolli J A. Computer Vision Based Detection of External Defects on Tomatoes Using Deep Learning[J]. Biosystems Engineering, 2020, 190, 131-144.
- [7] 饶晓慧. 机器视觉的农业图像识别系统分析[J]. 电子世界, 2021, 622(16): 49-50.
- [8] 王玉芯, 戢敏, 周黎明, 等. 机器视觉技术在田间农业机械中的应用[J]. 乡村科技, 2022, 13 (10): 152-154.
- [9] 徐天成, 吴敏, 贺冬仙, 等. 机器视觉在农业工程中的应用[J]. 农业工程, 2021, 11(08): 40-48.
- [10] Zhang L, Jia J, Gui G, et al. Deep Learning Based Improved Classification System for Designing Tomato Harvesting Robot[J]. IEEE Access, 2018, 6, 67940-67950.
- [11] 伍鉴芮, 张志勇, 韩小平, 等. 基于图像处理技术的番茄成熟度检测研究[J]. 农业技术与装备, 2021, 376(04): 50-51, 54.
- [12] Liu G, Mao S, Kim J H. A Mature-Tomato Detection Algorithm Using Machine Learning and Color Analysis[J]. Sensors, 2019, 19(9), 2023-2033.
- [13] Weiss K, Khoshgoftaar T M, Wang D. A Survey of Transfer Learning[J]. Journal of Big Data, 2016, 3(1), 1-40.

(下转第 69 页)

- [15] Ye X, Wang Z, Ma L, et al. Zinc oxide array/polyurethane nanocomposite coating: Fabrication, characterization and corrosion resistance[J]. Surface and Coatings Technology, 2019, 358: 497-504.
- [16] Pholnak C, Sirisathitkul C, Soontaranon S, et al. UV-Vis absorption and small angle X-ray scattering spectra of commercial polyurethane coating filled with zinc oxide[J]. National Academy Science Letters, 2016, 39: 125-128.
- [17] Zhou D, Wang J, Gao Y, et al. Corrosion behavior of tin plate in NaCl solution under different temperature [J]. International Journal of Electrochemical Science, 2017, 12: 192-205.
- [18] Sherif E S M, Alam M A, Alzahrani I S M. Fabrication of different protective coatings and studying their mechanical properties and corrosion behavior in sodium chloride solutions[J]. International Journal of Electrochemical Science, 2015, 10: 373-387.
- [19] 朱清, 刘悠嵘, 江志鹏, 等. 纳米氧化锌掺杂液晶/聚合物膜的阻抗谱特性[J]. 光电工程, 2020, 47(9): 62-70.
- [20] Zhang D, Wang M M, Jiang N, et al. Electrochemical corrosion behavior of Ni-doped ZnO thin film coated on low carbon steel substrate in 3.5% NaCl solution[J]. International Journal of Electrochemical Science, 2020, 15: 4117-4126.

(上接第 62 页)

- [14] 刘寿鑫, 龙伟, 李炎炎, 等. 基于 HSV 色彩空间的低照度图像增强[J]. 计算机工程与设计, 2021, 42(09): 2552-2560.
- [15] Sandler M, Howard A, Zhu M, et al. MobileNetV2: Inverted Residuals and Linear Bottlenecks [C]//In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. NJ: Piscataway, 2018: 4510-4520.
- [16] Szegedy C, Vanhoucke V, Ioffe S, et al. Rethinking the Inception Architecture for Computer Vision[C]//In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. NJ: Piscataway, 2016: 2818-2826.
- [17] Simonyan K, Zisserman A. Very Deep Convolutional Networks for Large-Scale Image Recognition[C]//In Proceedings of the 3rd International Conference on Learning Representations. CA: Berkeley, 2014: 1409-1420.
- [18] Robbins H, Monro S. Stochastic Approximation Method[J]. The Annals of Mathematical Statistics, 1951, 22(3): 400-407.
- [19] Zhang Z. Improved Adam Optimizer for Deep Neural Networks [C]//In 2018 IEEE/ACM 26th International Symposium on Quality of Service. NJ: Piscataway, 2018: 1-2.
- [20] Arjenaki O O, Moghaddam P A, Motlagh A M. Online Tomato Sorting Based on Shape, Maturity, Size, and Surface Defects Using Machine Vision[J]. Turkish Journal of Agriculture and Forestry, 2013, 37(1): 62-68.