

基于改进 YOLO v4 的玉米种子外观品质检测方法

范晓飞 王林柏 刘景艳 周玉宏 张 君 索雪松
(河北农业大学机电工程学院, 保定 071001)

摘要: 针对玉米种子在外观品质检测中需要快速识别与定位的需求,提出了一种基于改进 YOLO v4 的目标检测模型,同时结合四通道(RGB + NIR)多光谱图像,对玉米种子外观品质进行了识别与分类。为了减少改进后模型的参数量,本文将主干特征提取网络替换为轻量级网络 MobileNet V1。为了进一步提升模型的性能,通过试验研究了空间金字塔池化(Spatial pyramid pooling, SPP)结构在不同位置上对模型性能的影响,最终选取改进 YOLO v4 - MobileNet V1 模型对玉米种子外观品质进行检测。试验结果表明,模型的综合评价指标平均 F1 值和 mAP 达到 93.09% 和 98.02%,平均每检测 1 幅图像耗时 1.85 s,平均每检测 1 粒玉米种子耗时 0.088 s,模型参数量压缩为原始模型的 20%。四通道多光谱图像的光谱波段可扩展到可见光范围之外,并能够提取出更具有代表性的特征信息,并且改进后的模型具有鲁棒性强、实时性好、轻量化的优点,为实现种子的高通量质量检测和优选分级提供了参考。

关键词: 玉米种子; 外观品质; 多光谱图像; YOLO v4; MobileNet V1

中图分类号: TP391.4 文献标识码: A 文章编号: 1000-1298(2022)07-0226-08 OSID: 

Corn Seed Appearance Quality Estimation Based on Improved YOLO v4

FAN Xiaofei WANG Linbai LIU Jingyan ZHOU Yuhong ZHANG Jun SUO Xuesong
(College of Mechanical and Electrical Engineering, Hebei Agricultural University, Baoding 071001, China)

Abstract: Aim to identify and position corn seed, an object detection model based on improved YOLO v4 was proposed. This model combined with multi-spectral images with four channels (RGB + NIR), the appearance quality of corn seeds was identified and classified. In order to reduce the number of parameters in the model, the trunk feature extraction network was replaced with the lightweight network MobileNet V1. To improve the performance of the model, the effect of spatial pyramid pooling (SPP) structure on the model performance was studied. Finally, the improved YOLO v4 - MobileNet V1 model was selected to detect the appearance quality of corn seeds. The experimental results showed that the comprehensive evaluation indexes F1 and mAP of the model reached 93.09% and 98.02%, respectively. The average detection time of each image was 1.85 s, and the average detection time of each corn seed was 0.088 s. And the number of model parameters was compressed to 20% of the original model. The spectral band of four channel multi-spectral image can be extended beyond the visible range. Image can extract more representative feature information. The improved model had the advantages of strong robustness, good real-time performance and lightweight. It can provide a reference for high throughput quality detection and optimal classification of seeds.

Key words: corn seeds; appearance quality; multi-spectral images; YOLO v4; MobileNet V1

0 引言

玉米是世界上重要的粮食作物,广泛种植于世界各地,其产量和贸易量在近几年不断攀升^[1],流

通过程中,外观品质是影响玉米种子价格的重要因素之一^[2]。玉米种子在储存和运输过程中很容易受到破损和发霉问题的影响,而表型缺陷是种子品质评价的重要指标。当前种子品质检测以人工为

收稿日期: 2021-08-02 修回日期: 2021-10-25
基金项目: 国家自然科学基金面上项目(32072572)、河北省重点研发计划项目(20327403D)、河北省高层次人才项目(E2019100006)和河北农业大学人才引进研究项目(YJ201847)
作者简介: 范晓飞(1979—),男,教授,博士,主要从事农业人工智能研究,E-mail: leopardxf@163.com

主,效率低且具有主观性,因此将较为先进的目标检测技术应用到种子筛选当中具有重要意义。

许多学者对种子品质检测开展了相关研究,将机器学习^[3]技术应用到种子品质检测当中并且取得了一定的研究成果。刘双喜等^[4]使用了一种基于最远优先遍历的 DBSCAN 玉米种子纯度识别的机器学习算法,该方法对玉米种子纯度识别正确率达 93.3%。闫小梅等^[5]提出一种以玉米种子冠部与侧面颜色作为特征向量进行纯度识别的方法,其玉米种子纯度的准确率为 96.32%,但需要对种子的冠部和侧部进行图像获取,过程较繁琐。王玉亮等^[6]提出了一种基于多对象有效特征提取和主成分分析优化神经网络的玉米种子品种识别方法,通过提取形态特征和颜色特征来进行品种识别,种子平均识别率为 98.47%。

随着深度学习的发展,研究者将卷积神经网络模型应用于农作物种子质量检测当中,取得较好效果^[7-9]。近年来兴起的目标检测技术被广泛应用于农业当中,为玉米种子品质检测提供了参考^[10-13]。

在典型的深度学习目标检测算法中,一类是基于区域候选的目标检测,代表性算法有 R-CNN^[14]、Fast R-CNN^[15]、Faster R-CNN^[16]等,另一类是基于回归的目标检测,利用端到端的思想,将图像归一化后直接输入一个卷积神经网络中,回归预测出目标物体的类别和位置信息。代表性算法有 YOLO (You only look once)^[17]系列、SSD (Single shot multibox detector)^[18]系列等。其中区域候选的目标检测虽然在准确率上存在一定的优势,但是速度较慢,而 YOLO 系列算法有着高准确率和高检测速度的优点,适合于种子品质的检测。

在图像采集方面,传统的 RGB 图像通常在可见光范围内进行获取,其波段为 400 ~ 650 nm,多光谱相机采集的图像是由多个单通道的灰度图像组成,其波段扩展到可见光范围之外,每幅灰度图像都具备自身的光谱响应特性^[19]。而某些物质的特性区别会在特定的波段内显现出来,其波段较宽的多光谱相机可以提取出更具有代表性的特征信息,因此本文用多光谱相机采集的四通道 (RGB + NIR) 图像建立玉米种子质量图像数据库,用于玉米种子外观品质检测,同时对 YOLO v4 结构进行改进。

1 检测方法

1.1 数据采集

以不同外观品质的玉米种子作为研究对象,选取的玉米品种为郑单 985、科育 186、济玉 517。采用丹麦 JAI 公司的 4 通道多光谱面阵相机 (AD -

130GE 型)进行图像数据采集,相机采用棱镜分光技术结合双 CCD 采集图像,可以同时获取可见光图像和近红外图像 (图 1),其中可见光区域的波段为 400 ~ 650 nm,近红外区域的波段为 760 ~ 1 000 nm,中心波长中 R 通道为 610 nm、G 通道为 550 nm、B 通道为 450 nm、NIR 通道为 790 nm。采用 KOWA 公司生产的 LM25NC3 型镜头,焦距为 25 mm,该镜头拥有极高的色彩还原性、低失真性。图像分辨率为 1 296 像素 × 964 像素,每个颜色通道的像素深度为 8 位,并采用白色 LED 环形光源、近红外光源和增强图像对比度的白色背光源。相机高度固定为 35 cm,为了避免外界光源的影响,图像采集在封闭暗箱中进行。图像采集平台如图 2 所示。

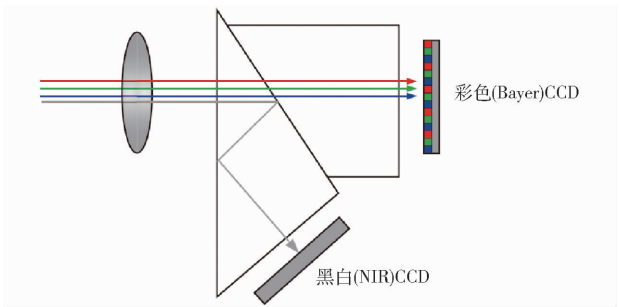


图 1 棱镜分光技术示意图

Fig. 1 Prism spectroscopic technology

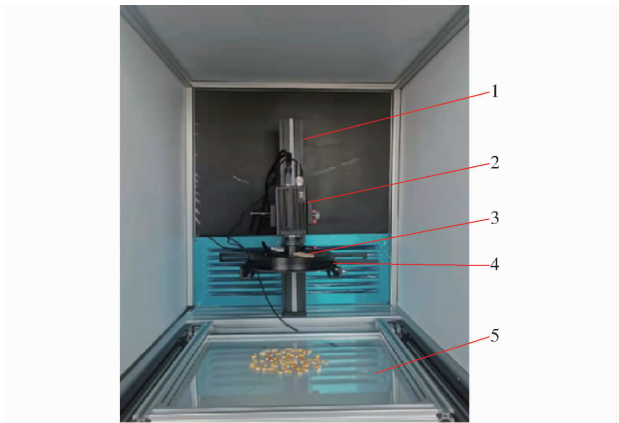


图 2 图像采集平台

Fig. 2 Image acquisition platform

1. 支架 2. 相机 3. 环形白光光源 4. 环形近红外光源 5. 白色背光源

本文采集原始玉米种子数据 200 组,每组图像中包含 RGB 图像和 NIR 图像,如图 3 所示。将色泽鲜亮、饱满的籽粒作为良好籽粒,将带有破损、虫蚀、发霉的籽粒作为带有缺陷籽粒,经统计外观良好的籽粒有 1 130 粒,外观带有缺陷的籽粒有 1 056 粒。由于 NIR 图像在波段 760 ~ 1 000 nm 的近红外光谱范围内采集,因此 NIR 图像包含人眼视觉之外的图像信息,本试验通过增加近红外波段下的图像,为模型提供更多的特征信息。

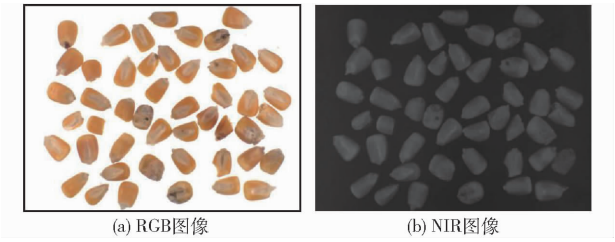


图 3 原始图像
Fig. 3 Original images

本文对 RGB、NIR 图像使用相同的数据增强方式(亮度调整、旋转、加入高斯噪声等操作)来增加图像数量,提升训练后模型的性能,最终获取 650 组图像,其中外观良好的籽粒有 5 082 粒,外观带有缺陷的籽粒有 4 613 粒。训练模型采用 PASCAL VOC 数据集格式,用 Labellmg 标注工具对每幅图像的每个玉米种子进行手工标注矩形框,将外观良好的玉米种子标签设为 good,外观带有缺陷的玉米种子标签设为 bad,选取其中 10 组图像作为验证集,其中外观良好的种子有 110 粒,外观带有缺陷的种子有 100 粒,用于模型最终的性能评估,将剩余的 640 组图像按照比例 4:1 划分为训练集和验证集,用于模型训练。

1.2 YOLO v4 目标检测算法

BOCHKOVSKIY 等在 YOLO v3 的基础上,提出了 YOLO v4 模型,整个网络结构分为主干特征提取网络 CSPDarkNet53^[20],在 DarkNet53 的基础上进一步增加了底层信息的融合,增强信息的提取能力;加强特征提取网络 FPN (Feature pyramid networks) + PAN(Path aggregation network)^[21-22]和空间金字塔池化(Spatial pyramid pooling)^[23];最后为预测网络,利用获得的特征进行预测。其中空间金字塔池化位于主干网络和颈部网络的结合处,如图 4 所示,其将输入特征图分别通过最大池化的方式变为不同尺寸的特征图,然后将不同尺寸(5×15、9×9、13×13)的特征图与原特征图进行连接操作,作为新的特征图,采用这种方式能够更好地增加卷积核的感受野。在 YOLO v4 的颈部网络部分使用特征金字塔网络(FPN)和路径聚合网络(PAN),如图 5 所示,FPN 将高层的特征图通过上采样的操作与底层的特征图进行连接,增加了特征图的信息量,PAN 通过下采样连接底层特征和高层特征,缩短了各层之间的融合路径,提升了网络特征的提取能力。

1.3 MobileNet V1 网络模型

为了使模型达到更好的性能,通常将模型网络设计得更深更复杂,这样会导致模型参数数量和计算量的增加,从而对硬件的要求会有所提高。而 MobileNet V1^[24]作为具有代表性的小模型,其使用

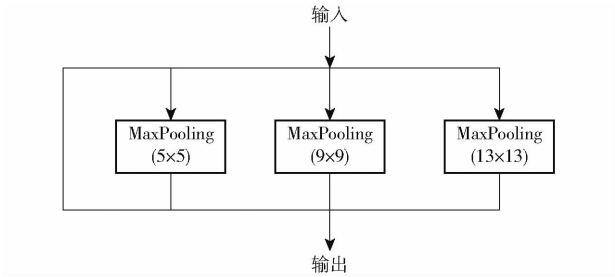


图 4 SPP 网络结构图
Fig. 4 SPP network structure diagram

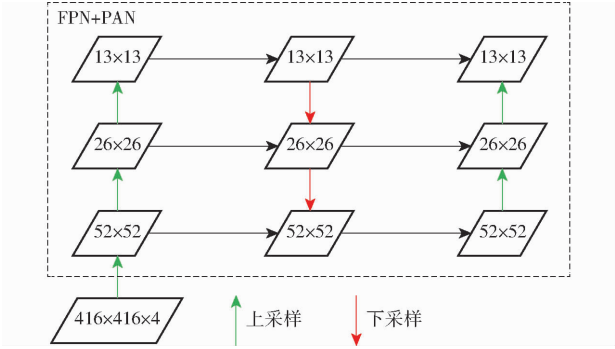


图 5 FPN 和 PAN 网络结构图
Fig. 5 FPN and PAN network structure diagram

深度可分离卷积(图 6)构建的轻量级模型可以达到较好的检测效果。深度可分离卷积包含逐深度卷积 (Depthwise convolution) 和逐点卷积 (Pointwise convolution) 两个过程。逐深度卷积将每一个通道与对应的每一个卷积核进行卷积操作,逐点卷积采用 1×1×M 的卷积核,M 为输入特征图的通道数。通过 1×1 卷积操作,实现特征图在深度上的加权融合。利用 3×3 的深度可分离卷积与传统的卷积方法相比,模型参数量降低到原来的 1/10,同时引入宽度因子 α 在每层对网络的输入输出通道数进行缩减,引入分辨率因子 β 用于控制输入和内部层表示,即用分辨率因子控制输入的分辨率。MobileNet V1 通过深度可分离卷积结构实现了在性能没有明显损失的前提下使模型参数量和计算量降低。

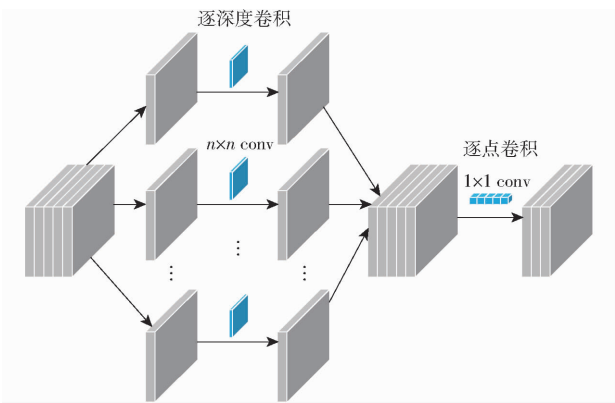


图 6 深度可分离卷积结构图
Fig. 6 Depth-wise separable convolution structure diagram

1.4 改进 YOLO v4 目标检测模型

将多光谱相机采集的 4 通道图像作为输入,将其尺寸设为 416×416 。在 YOLO v4 中运用 CSPDarkNet53 作为主干特征提取网络,整个网络包含 72 个卷积层,虽然原始的 YOLO v4 模型性能优异,考虑到使用过深的网络会增加模型的参数数量和检测时间,同时本文中目标的背景为白色背光板,并不需要深层的神经网络,因此将 MobileNet V1 模型作为主干特征提取网络,从而减少 YOLO v4 网络的参数量,提升计算效率。改进后的 YOLO v4 – MobileNet V1 模型的网络结构如图 7

所示。选择 MobileNet V1 模型的 P5、P11、P13 作为 3 个不同的特征层,输出尺寸为 52×52 、 26×26 、 13×13 。在主干特征提取网络的尾端 P13 处保留了原来的 SPP 结构,同时为了更好地提取每一个特征层的全局信息,在 P5、P11 处分别增加了相应的 SPP2、SPP1 (图 8),将每个特征层进行最大池化,并与原来特征层进行融合,从而提升了不同尺寸特征层的全局信息。然后将 3 个特征层输入 FPN 和 PAN 网络中,实现高层特征信息与低层特征信息的融合,进而输入到预测网络中进行预测。

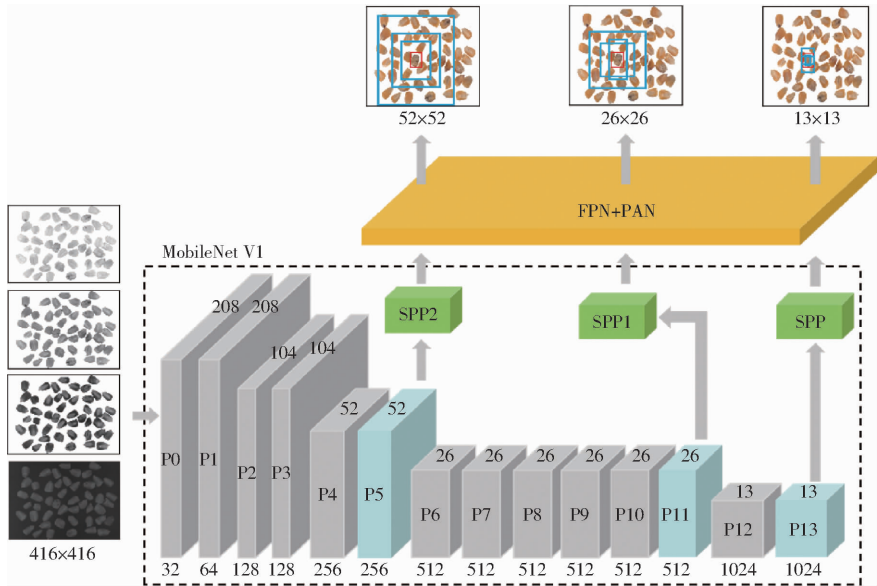


图 7 YOLO v4 – MobileNet V1 结构图

Fig. 7 YOLO v4 – MobileNet V1 structure diagram

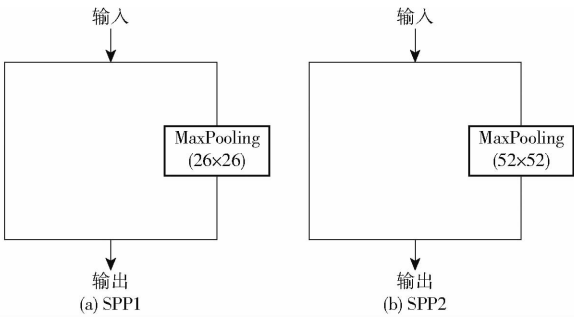


图 8 SPP1 和 SPP2 网络结构

Fig. 8 SPP1 and SPP2 network diagram

YOLO v4 – MobileNet V1 模型的损失函数部分与 YOLO v4 模型相同,由分类损失函数、回归损失函数和置信度损失函数 3 部分组成。在回归损失函数中,传统方法使用非极大值抑制 (Non-maximum suppression, NMS) 算法将预测框按置信度排序,将得分最高的框与剩余的框作交并比 (Intersection over union, IoU) 计算,从而找到最合适的预测框。YOLO v4 模型使用损失函数完全交并比 (Complete intersection over union, CIoU) [25] 代替传统方法来优

化回归损失,CIoU 将目标与预测框之间的距离、重叠率、尺寸以及惩罚项都考虑进去,使得目标框回归变得更加稳定。

1.5 模型的评估方法

采用精确率 (Precision)、召回率 (Recall)、F1 值、平均精度 (AP) 和平均精度均值 (mAP) 作为评价指标。

2 试验与结果分析

2.1 模型训练

试验条件为 Windows 10、64 位操作系统, CUDA 版本为 10.0, 采用基于 Python 编程语言的 Tensorflow 与 Keras 深度学习框架。计算机配置为 GeForce GTX 2070 显卡, 8 GB 显存, Intel (R) Core (TM) i5 – 9400F 处理器, 主频 2.90 GHz。

模型输入图像尺寸为 $416 \times 416 \times 4$, 在训练过程中使用随机旋转、随机缩放、随机拉伸、随机色域变换的数据增强方法, 由于输入图像为四通道图像, 其色域变换中首先将三通道图像进行变换, 然后再

加上 NIR 图像,从而实现四通道的色域变换。试验采用迁移学习的方式进行训练,将模型的主干特征提取网络使用 VOC 数据集的预训练权重进行初始化参数设置,模型共训练 150 个迭代周期(Epoch),使用 Adam 作为优化算法。前 50 次将学习率设为 1×10^{-3} ,批量大小设为 16,将主干特征提取网络进行冻结训练,从而加快收敛速度,避免预训练权值被破坏,后 100 次将学习率设为 1×10^{-4} ,批量大小设为 8,对主干特征提取网络进行解冻,以更小的初始学习率对整个模型进一步训练,从而加快整个网络的训练时间。其训练损失值曲线如图 9 所示。

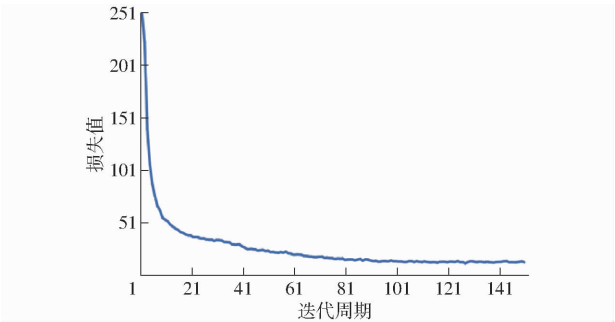


图 9 损失值变化曲线
Fig. 9 Change curve of loss value

为了选出整体训练性能足够高的模型,本文以 mAP 最高的权值文件作为最终模型,模型进行了 150 次迭代,每 3 次迭代输出 1 个模型,得到 50 个模型,在这 50 个模型中找出 1 个 mAP 最高的模型(图 10)。当 mAP 趋于稳定状态时,最大值为 98.02%,此时的模型即为本文选用的模型。

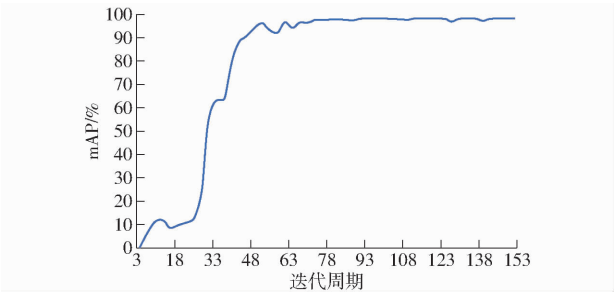


图 10 mAP 随迭代周期的变化曲线
Fig. 10 mAP value change curve with epoch

2.2 RGB 图像和 RGB + NIR 图像的目标检测模型对比

首先对 RGB 三通道玉米种子图像进行试验,将 MobileNet V1、MobileNet V2^[26] 以及 MobileNet V3^[27] 的主干网络用于 YOLO v4 网络的特征提取,同时对比以 CSPDarkNet53 作为主干网络的 YOLO v4 模型。将训练完成后的模型对玉米种子图像的验证集进行验证,其结果如表 1 所示。YOLO v4 – MobileNet V1 模型 mAP 最高,为 96.33%,性能优于 YOLO v4 – MobileNet V2 和 YOLO v4 – MobileNet V3 模型,相较于 YOLO v4 – CSPDarkNet53 模型平均 F1 值降低 0.25 个百分点,但 mAP 提高 12.70 个百分点,因此本文选取 MobileNet V1 替换 CSPDarkNet53 作为 YOLO v4 的主干网络。然后本文在四通道(RGB + NIR)的玉米种子图像下进行试验,其结果如表 2 所示,从表中可以看出 4 种模型的平均 F1 值相较于 RGB 图像都有所增加。试验说明四通道(RGB + NIR)玉米种子图像经过卷积神经网络处理可以提取出优于 RGB 图像的特征信息,因为 RGB 图像是在波段 400 ~ 650 nm 的可见光区域获取,而 NIR 图像是在波段 760 ~ 1 000 nm 的近红外区域获取,然而不同质量的玉米种子会存在不同特性的物质,其在特定的光谱范围下的表现会有所不同,四通道图像中每幅灰度图像都具备自身的光谱响应特性。因此多光谱图像为模型提供更具有代表性的特征信息,从而达到提升模型性能的目的。

2.3 改进 YOLO v4 – MobileNet V1 模型与不同模型的对比

为了验证改进 YOLO v4 – MobileNet V1 模型的性能,对改进型 YOLO v4 – MobileNet V1 结构中的 SPP1、SPP2 进行了改变,将 SPP1、SPP2 分别改为图 11 所示结构,由于原始网络的 SPP 中最大池化层为 13×13 ,与最后一个特征层相同,因此将其他 SPP 结构的最大池化层尺寸设为与提取特征图的尺寸相同,其他 2 个的最大池化层按相同比例减少,将

表 1 RGB 图像下模型的检测结果

Tab.1 Detection results of model under RGB image

模型	类别	准确率/%	召回率/%	F1 值/%	平均 F1 值/%	AP/%	mAP/%
YOLO v4 – CSPDarkNet53	good	95.79	82.73	88.78	90.78	80.14	83.63
	bad	97.47	88.51	92.77		87.12	
YOLO v4 – MobileNet V1	good	88.33	96.36	92.17	90.53	97.04	96.33
	bad	96.00	82.76	88.89		95.61	
YOLO v4 – MobileNet V2	good	86.40	98.18	91.91	89.98	97.20	96.07
	bad	97.22	80.46	88.05		94.95	
YOLO v4 – MobileNet V3	good	87.70	97.27	92.24	89.79	97.23	95.90
	bad	97.18	79.31	87.34		94.57	

表 2 RGB + NIR 图像下模型的检测结果

Tab. 2 Detection results of model under RGB + NIR image

模型	类别	准确率/%	召回率/%	F1 值/%	平均 F1 值/%	AP/%	mAP/%
YOLO v4 – CSPDarkNet53	good	96. 56	83. 65	89. 64	91. 49	85. 16	86. 70
	bad	97. 38	89. 62	93. 34		88. 23	
YOLO v4 – MobileNet V1	good	91. 15	93. 64	92. 38	91. 10	96. 96	96. 50
	bad	93. 75	86. 21	89. 82		96. 03	
YOLO v4 – MobileNet V2	good	90. 27	92. 73	91. 48	90. 26	96. 37	96. 24
	bad	94. 81	83. 91	89. 03		96. 10	
YOLO v4 – MobileNet V3	good	88. 24	95. 45	91. 70	89. 95	96. 87	95. 24
	bad	95. 95	81. 61	88. 20		95. 60	

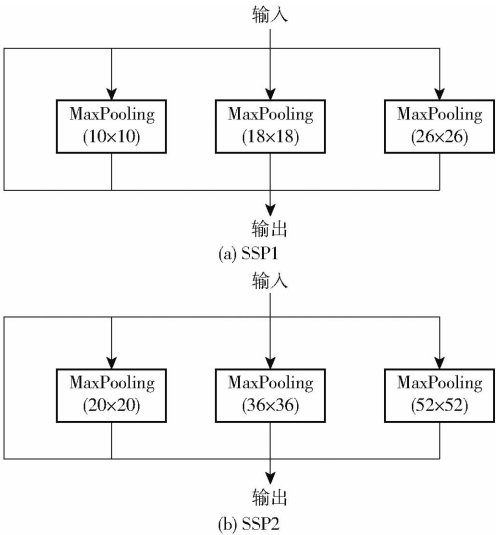


图 11 SPP 网络结构

Fig. 11 SPP network structure

此模型命名为 YOLO v4 – MobileNet V1 – 3SPP。将

训练后改进 YOLO v4 – MobileNet V1 模型、YOLO v4 – MobileNet V1 – 3SPP 模型、YOLO v4 – MobileNet V1 模型在相同验证集下进行验证试验,其结果如表 3 所示, YOLO v4 – MobileNet V1 – 3SPP 模型的平均 F1 值和 mAP 均高于 YOLO v4 – MobileNet V1 模型,而改进 YOLO v4 – MobileNet V1 模型的平均 F1 值高于 YOLO v4 – MobileNet V1 – 3SPP 模型 1. 13 个百分点,高于 YOLO v4 – MobileNet V1 模型 1. 99 个百分点,其 mAP 高于 YOLO v4 – MobileNet V1 – 3SPP 模型 1. 16 个百分点,高于 YOLO v4 – MobileNet V1 模型 1. 50 个百分点。试验说明增加过多最大池化层的模型效果并没有优于只有特征图尺寸相同的最大池化层模型,因此增加 SPP 结构可以使其通过最大池化层提取全局信息来增加特征图的感受野,从而分离出最显著的上下文特征,提升模型的特征提取能力。

表 3 不同 SPP 结构下模型的检测结果

Tab. 3 Test results of models with different SPP structures

模型	类别	准确率/%	召回率/%	F1 值/%	平均 F1 值/%	AP/%	mAP/%
YOLO v4 – MobileNet V1	good	91. 15	93. 64	92. 38	91. 10	96. 96	96. 50
	bad	93. 75	86. 21	89. 82		96. 03	
YOLO v4 – MobileNet V1 – 3SPP	good	90. 00	98. 18	93. 91	91. 96	97. 75	96. 86
	bad	98. 63	82. 76	90. 00		95. 96	
改进 YOLO v4 – MobileNet V1	good	91. 53	98. 18	94. 71	93. 09	98. 31	98. 02
	bad	97. 40	86. 21	91. 46		97. 73	

2. 4 四通道 (RGB + NIR) 图像下不同模型的对比

为了证明改进 YOLO v4 – MobileNet V1 模型的性能,本文使用四通道 (RGB + NIR) 验证集图像对原始 YOLO v4 模型、YOLO v4 – MobileNet V1 模型、YOLO v4 – MobileNet V1 – 3SPP 模型、改进 YOLO v4 – MobileNet V1 模型进行了验证,对比分析了每种模型对玉米种子图像的检测结果、模型的准确性、总参数量和检测单幅图像的耗时。在检测结果中选取较稀疏排列的玉米种子和较密集排列的玉米种子(图 12),从图中可以看出,改进 YOLO v4 – MobileNet V1 模型可以准确地定位每粒种子图像各

1 幅,并且可以准确地识别出单粒玉米种子品质的好坏,并在不同种子密度条件下,模型都能达到较好的识别效果,图像中种子稀疏与密集的摆放形式并不影响模型的性能。而 YOLO v4 – MobileNet V1 – 3SPP 模型在较密集排列的玉米种子图像中有 3 个籽粒没有检测出,说明 SPP 结构中较多的最大池化层并没有达到较好的效果。模型的性能如表 4 所示,其中改进 YOLO v4 – MobileNet V1 模型的平均 F1 值和 mAP 最高,说明模型的准确性最高,性能最强。由于 MobileNet V1 轻量级网络代替了原始模型中的 CSPDarkNet53 网络,因此模型的总参数量变为

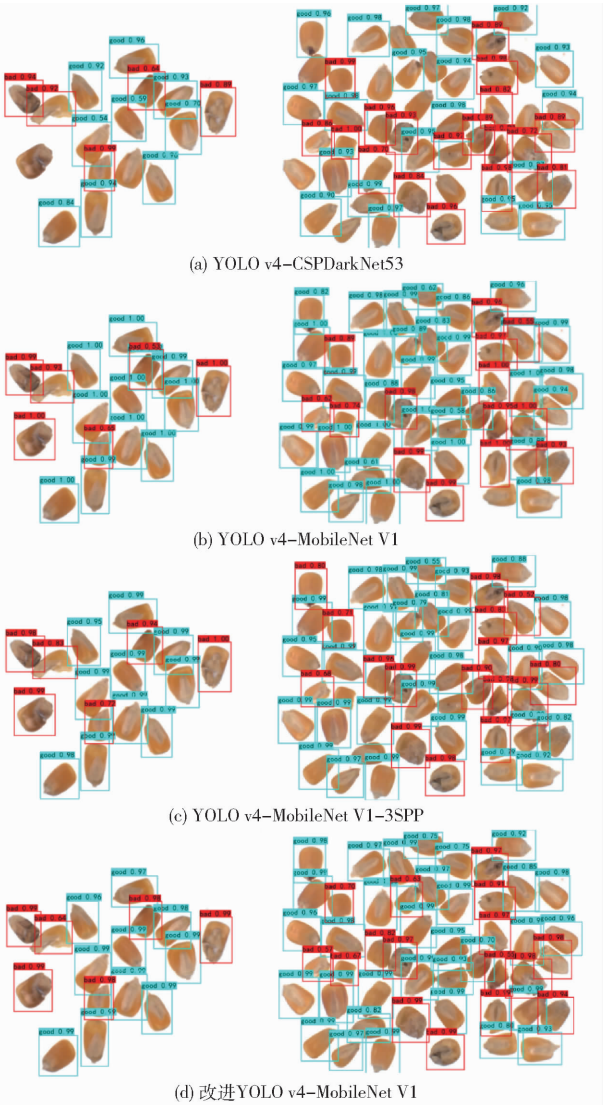


图 12 不同模型的玉米种子检测结果

Fig. 12 Test results of corn seeds by different models

原来的 20%,在模型检测单幅图像的时间上,由于网络层数的减少,时间缩短了约 0.5 s。最终改进

表 4 模型的性能对比

Tab. 4 Performance comparison of models

模型	平均 F1 值/%	mAP/%	参数数量	检测时间/s
YOLO v4 - CSPDarkNet53	91.49	86.70	6.401×10^7	2.00
YOLO v4 - MobileNet V1	91.10	96.50	1.233×10^7	1.87
YOLO v4 - MobileNet V1 - 3SPP	91.96	96.86	1.300×10^7	1.84
改进 YOLO v4 - MobileNet V1	93.09	98.02	1.284×10^7	1.85

YOLO v4 - MobileNet V1 在检测 10 组 RGB + NIR 图像中共消耗 18.5 s,平均每检测 1 粒玉米种子耗时 0.088 s,达到实时检测效果。

3 结论

(1)对比三通道 (RGB) 图像和四通道 (RGB + NIR) 图像的模型训练结果表明,四通道图像的平均 F1 值和 mAP 较高,近红外图像包含可见光之外的特征信息,而不同品质的种子特性会在不同的光谱波段范围内显现出来,其四通道图像所包含的信息量高于三通道图像,因此将四通道图像作为输入。

(2)为了使模型的计算量和参数数量减少,将 YOLO v4 主干网络 (CSPDarkNet53) 替换为轻量级的 MobileNet V1,同时对比了其他 MobileNet 模型,结果表明,MobileNet V1 的性能最优,同时模型的总参数量变为原始模型的 20%,本研究在确保精度的前提下,大幅度减少了模型的总参数量。

(3)改进 YOLO v4 - MobileNet V1 目标检测模型中由于增加了 2 个 SPP 结构,融合了特征图的全局特征,更好地增加了卷积核的感受野,并提取出上下位的重要信息,因此模型的平均 F1 值和 mAP 分别达到 93.09% 和 98.02%,平均每检测 1 幅图像耗时 1.85 s,平均每检测 1 粒玉米种子耗时 0.088 s。

参 考 文 献

[1] 李春宏. 玉米种子市场价格分析[J]. 种子科技, 2004(4):3-4.
LI Chunhong. Market price analysis of corn seed[J]. Seed Science & Technology, 2004(4):3-4. (in Chinese)

[2] 付恭华. 中国粮食生产的多维成本研究[D]. 北京:中国农业大学, 2014.
FU Gonghua. A study on multidimensional cost of grain production in China[D]. Beijing: China Agricultural University, 2014. (in Chinese)

[3] TU K L, LI L J, YANG L M, et al. Selection for high quality pepper seeds by machine vision and classifiers[J]. Journal of Integrative Agriculture, 2018, 17(9):1999-2006.

[4] 刘双喜,王盼,张春庆,等. 基于优化 DBSCAN 算法的玉米种子纯度识别[J]. 农业机械学报, 2012, 43(4):188-192.
LIU Shuangxi, WANG Pan, ZHANG Chunqing, et al. Maize purity identification based on improved DBSCAN algorithm[J]. Transactions of the Chinese Society for Agricultural Machinery, 2012, 43(4):188-192. (in Chinese)

[5] 闫小梅,刘双喜,张春庆,等. 基于颜色特征的玉米种子纯度识别[J]. 农业工程学报, 2010, 26(13):46-50.
YAN Xiaomei, LIU Shuangxi, ZHANG Chunqing, et al. Purity identification of maize seed based on color characteristics[J]. Transactions of the CSAE, 2010, 26(13):46-50. (in Chinese)

[6] 王玉亮,刘贤喜,苏庆堂,等. 多对象特征提取和优化神经网络的玉米种子品种识别[J]. 农业工程学报, 2010, 26(6):199-204, 389.
WANG Yuliang, LIU Xianxi, SU Qingtang, et al. Maize seeds varieties identification based on multi-object feature extraction

- and optimized neural network[J]. Transactions of the CSAE, 2010, 26(6): 199–204, 389. (in Chinese)
- [7] 侯俊铭, 姚恩超, 朱红杰. 基于卷积神经网络的蓖麻种子损伤分类研究[J]. 农业机械学报, 2020, 51(增刊1): 440–449. HOU Junming, YAO Enchao, ZHU Hongjie. Classification of castor seed damage based on convolutional neural network[J]. Transactions of the Chinese Society for Agricultural Machinery, 2020, 51(Supp.1): 440–449. (in Chinese)
- [8] 祝诗平, 卓佳鑫, 黄华, 等. 基于 CNN 的小麦籽粒完整性图像检测系统[J]. 农业机械学报, 2020, 51(5): 36–42. ZHU Shiping, ZHUO Jiaxin, HUANG Hua, et al. Wheat grain integrity image detection system based on CNN[J]. Transactions of the Chinese Society for Agricultural Machinery, 2020, 51(5): 36–42. (in Chinese)
- [9] MICHA K, PRZEMYSAW G, PIOTR M S. Varietal classification of barley by convolutional neural networks[J]. Biosystems Engineering, 2019, 184(2): 155–165.
- [10] 张宏鸣, 付振宇, 韩文霆, 等. 基于改进 YOLO 的玉米幼苗株数获取方法[J]. 农业机械学报, 2021, 52(4): 221–229. ZHANG Hongming, FU Zhenyu, HAN Wenting, et al. Detection method of maize seedlings number based on improved YOLO[J]. Transactions of the Chinese Society for Agricultural Machinery, 2021, 52(4): 221–229. (in Chinese)
- [11] 张晴晖, 孔德肖, 李俊菽, 等. 基于逆运动学降维求解与 YOLO v4 的果实采摘系统研究[J]. 农业机械学报, 2021, 52(9): 15–23. ZHANG Qinghui, KONG Dexiao, LI Junqiu, et al. Design of fruit picking system based on inverse kinematics dimension reduction and YOLO v4[J]. Transactions of the Chinese Society for Agricultural Machinery, 2021, 52(9): 15–23. (in Chinese)
- [12] RUI S A, FG A, ZZ A, et al. Improved multi-classes kiwifruit detection in orchard to avoid collisions during robotic picking[J]. Computers and Electronics in Agriculture, 2021, 182: 106052.
- [13] WU L, MA J, ZHAO Y, et al. Apple detection in complex scene using the improved YOLOv4 model[J]. Agronomy, 2021(11): 476.
- [14] GIRSHICK R, DONAHUE J, DARREL T, et al. Rich feature hierarchies for accurate object detection and semantic segmentation[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2014: 580–587.
- [15] GIRSHICK R. Fast R-CNN[C]//Proceedings of the IEEE International Conference on Computer Vision, 2015: 1440–1448.
- [16] REN S, HE K, GIRSHICK R, et al. Faster R-CNN: towards real-time object detection with region proposal networks[J]. IEEE Transactions on Pattern Analysis & Machine Intelligence, 2017, 39(6): 1137–1149.
- [17] REDMON J, DIVVALA S, GIRSHICK R, et al. You only look once: unified, real-time object detection[C]//Computer Vision & Pattern Recognition, IEEE, 2016.
- [18] LIU W, ANGUELOV D, ERHAN D, et al. SSD: single shot multiBox detector[C]//European Conference on Computer Vision. Springer International Publishing, 2016.
- [19] 郭文锦, 王红武, 陈绍江, 等. 基于近红外光谱的商品玉米品种快速鉴别方法[J]. 光谱学与光谱分析, 2010, 30(5): 1248–1251. WU Wenjin, WANG Hongwu, CHEN Shaojiang, et al. Fast discrimination of commercial corn varieties based on near infrared spectra[J]. Spectroscopy and Spectral Analysis, 2010, 30(5): 1248–1251. (in Chinese)
- [20] WANG C Y, LIAO H, YEH I H, et al. CSPNet: a new backbone that can enhance learning capability of CNN[C]//2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops. IEEE, 2020.
- [21] LIN T Y, DOLLAR P, GIRSHICK R, et al. Feature pyramid networks for object detection[C]//2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). IEEE Computer Society, 2017.
- [22] LIU S, QI L, QIN H, et al. Path aggregation network for instance segmentation[C]//CVPR, 2018: 8759–8768.
- [23] HE K, ZHANG X, REN S, et al. Spatial pyramid pooling in deep convolutional networks for visual recognition[J]. IEEE Transactions on Pattern Analysis & Machine Intelligence, 2014, 37(9): 1904–1916.
- [24] HOWARD A G, ZHU M, CHEN B, et al. Mobilenets: efficient convolutional neural networks for mobile vision applications[J]. arXiv preprint, 2017: 1704.04861.
- [25] ZHENG Z, WANG P, LIU W, et al. Distance-IoU loss: faster and better learning for bounding box regression[C]//AAAI Conference on Artificial Intelligence, 2020.
- [26] SANDLER M, HOWARD A, ZHU M, et al. Mobilenetv2: inverted residuals and linear bottlenecks[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2018: 4510–4520.
- [27] HOWARD A, SANDLER M, CHU G, et al. Searching for Mobilenetv3[C]//Proceedings of the IEEE/CVF International Conference on Computer Vision, 2019: 1314–1324.