

全召茂,陈学海,马志艳,等.融合图像增强和迁移学习的YOLOv8n夜间苹果检测方法[J].华中农业大学学报,2024,43(5):1-9.
DOI:10.13300/j.cnki.hnlkxb.2024.05.001

融合图像增强和迁移学习的YOLOv8n 夜间苹果检测方法

全召茂¹,陈学海^{1,2},马志艳^{1,2},杨光友^{1,2},张灿¹

1.湖北工业大学农机工程研究设计院,武汉 430068; 2.湖北省农机装备智能化工程技术研究中心,武汉 430068

摘要 针对夜间场景下苹果识别率低、实时性差的问题,提出一种融合图像增强和迁移学习的YOLOv8n夜间苹果检测方法。首先,在YOLOv8n前端嵌入Zero-DCE模块增强夜间图像,更清晰地呈现苹果的轮廓和细节,降低夜间苹果图像的识别难度;其次,使用SPD-Conv进行下采样,增强模型细粒度特征的提取能力;在此基础上,针对夜间苹果数据集样本量少的问题,采用迁移学习训练策略,选取含有苹果类别的MS COCO数据集作为源域数据集,对于夜间场景下的目标域数据集,利用Zero-DCE增加其与日间苹果图像的相似度并在源域模型上微调目标域模型。基于上述方法,在夜间苹果图像数据集上进行了试验,结果显示,所提方法的模型精确率 P 为97.0%、召回率 R 为93.4%、平均精度均值 $mAP@0.5:0.95$ 为74.6%,较YOLOv8n原始模型分别提升2.3、1.9和4.3个百分点,同时该模型的推理速度为22 帧/s,可以满足实时性要求。消融试验显示,图像增强与迁移学习结合使用的效果超过两者单独使用时的效果之和。研究表明,改进后的模型在处理重叠、遮挡、绿果和光线过暗等复杂情形时都比原始模型表现更优,具有良好的鲁棒性。

关键词 苹果; 夜间检测; 图像增强; 迁移学习; YOLOv8

中图分类号 TP391.4 **文献标识码** A **文章编号** 1000-2421(2024)05-0001-09

中国是全球最大的苹果生产国,2021年产量达到4 597.34万t,占据全球总产量的54.7%^[1-2]。在我国苹果产业快速发展和人力资源短缺的背景下,果农对智能农业装备的需求不断增长^[3-4]。视觉系统对苹果采摘机器人具有重要意义^[4],同时在苹果的产量估算^[5]、果实分拣^[6]和表型分析^[7]等任务中也起着关键作用。目前,大部分采摘机器人与监控设备都是白天作业,如能利用夜间进行连续作业,将提高全天工作效率,有力地推动苹果产业的发展。

夜间环境下光线弱,采集的图像存在暗淡、模糊和阴影等问题^[8],同时枝叶枝干遮挡、果实密集重叠等情况又进一步增加了识别难度,准确识别夜间苹果是有挑战性的任务。近年来,一些研究者基于图像处理的方法对夜间果实检测进行了探索。Font等^[9]利用人工光源采集成熟葡萄的RGB图像,并通过检测葡萄表面的球面反射峰值进行葡萄计数,该方法的平均计数误差为14%;赵德安等^[10]将R-G色

差分割法用于苹果采摘机器人夜间检测,该方法仅在良好照明条件下表现较佳,在不考虑果实的粘连和遮挡等因素的前提下,正确识别率为83.7%;戴家裕^[11]使用明场漫射正面照明作为夜间补光方式,解决了苹果表面阴影和反光的问题,并通过直方图均衡化和gamma校正进一步对图像增强,该方法能够同时克服夜间环境和同色系苹果的识别难点,对绿苹果的识别成功率为87%。随着深度学习技术的发展,以YOLO^[12]为代表的目标检测算法广泛应用于夜间果实检测。熊俊涛等^[13]借鉴DenseNet^[14]的思想改进YOLOv3,增强了模型在夜间复杂环境中检测成熟柑橘的稳健性,精确率达到97.67%,但该方法对暗光图像识别效果欠佳,且容易将密集多目标识别为一个目标。孙宝霞等^[15]设计了多光源结合的视觉系统,并基于YOLOv4对夜间自然环境下成熟柑橘进行识别与表征缺陷检测,准确率为95.3%,该方法的有效性依赖于光源系统的设计。何斌等^[16]利用

收稿日期:2023-12-01

基金项目:湖北省科技创新人才计划项目(2023DJC088);湖北省农机装备补短板核心技术应用攻关项目(HBSNYT202208)

全召茂,E-mail:tzhaomao@163.com

通信作者:陈学海,E-mail:cxh@hbut.edu.cn

YOLOv5检测夜间环境中的番茄,但由于番茄的绿果在夜间环境下的颜色容易与叶片、秆茎等混淆,且重叠果实的边界区别较为困难,导致番茄的绿果和多果识别率低于红果。

综上所述,传统图像处理方法在夜间果实检测任务中存在精度低、速度慢和通用性差等局限性,难以满足实际作业要求,基于深度学习的方法具备更优的性能,但也有不足之处,许多研究者沿用白天检测环境的思路,忽视了夜间环境的特殊性,导致面对重叠、遮挡、绿果和光线过暗等情形时检测效果欠佳,同时部分研究还过度依赖于外部辅助光源系统弥补图像质量的不足,此外也没有利用迁移学习策略把日间数据纳入到模型训练中。针对上述问题,本研究提出了一种融合图像增强与迁移学习的YOLOv8n的夜间苹果检测方法,以期实现对夜间苹果的准确实时检测,从而提升苹果采摘机器人及相关机器视觉设备的全天工作效率。

1 材料与方法

1.1 数据集

本研究将MS COCO的训练集^[17]作为源域数据,其中包含80个类别的118 287幅图像,苹果类别的实例个数有5 776个。目标域数据来自2019年现代果园环境下苹果图像数据集(<https://www.agridata.cn/data.html#/datadetail?id=289878>),该数据集由华盛顿州立大学农业自动化和机器人实验室创

建,用于产量估算和机器人收割,其中含有夜间环境下人工照明的苹果图像273幅。将拍摄过于模糊、遮挡十分严重等容易对神经网络模型训练时产生误导的苹果图像进行裁剪,以提升数据质量,确保模型可以学到代表性的特征。经过处理后的数据集共1 092幅,包含无遮挡、枝叶枝干遮挡、果实重叠、单目标和多目标果实等多种情景。依据数据标注原则,使用LabelImg将图像中苹果实例的最小外接矩形标注为真实框,得到包含苹果实例的中心坐标和宽高信息的xml文件,并保存为PASCAL VOC格式,标注结果示例如图1所示。

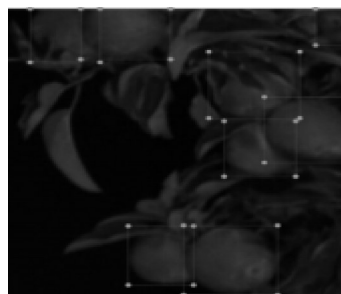
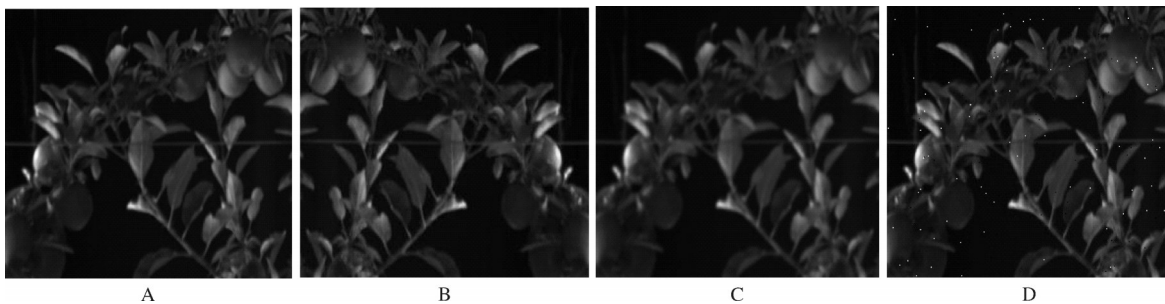


图1 数据标注结果示例

Fig.1 Example of data annotation results

将标注后的数据集按7:3划分,得到764幅图像的训练集和328幅图像的测试集。然后,在训练集上采用翻转、适度的模糊和噪声等数据增强技术,以提升模型的适应能力,增强结果示例如图2所示,经过数据增强后的训练集共有1 528幅图像。



A:原图 Original;B:翻转 Flipped;C:高斯模糊 Gaussian blur;D:椒盐噪声 Salt-pepper noise.

图2 数据增强结果示例

Fig.2 Example of data augmentation results

1.2 检测方法

YOLOv8是当前主流的目标检测网络之一,具备高效的训练和推理能力,能够准确快速地分类和定位果实目标,根据网络深度和特征图宽度的不同有n、s、m、l、x 5个版本,以满足不同应用场景的需求。为保证相关机器视觉设备的工作效率,本研究采用复杂度较低的YOLOv8n模型,并在此基础上搭

建适合夜间苹果检测的网络,通过零参考深度曲线估计(zero-reference deep curve estimation, Zero-DCE)^[18]模块增强夜间图像,并利用SPD-Conv^[19]提取细粒度特征,同时结合迁移学习(transfer learning)策略进一步优化模型性能。

1) Zero-DCE增强夜间图像。夜间采集的图像整体偏暗,导致部分果实边缘与背景的颜色差异微

小,给机器视觉识别带来较大困难,鉴于此,引入Zero-DCE增强夜间图像,更清晰地展现苹果的轮廓和

细节,以降低夜间图像识别难度,从而为后续图像处理任务提供基础,Zero-DCE的框架如图3所示。

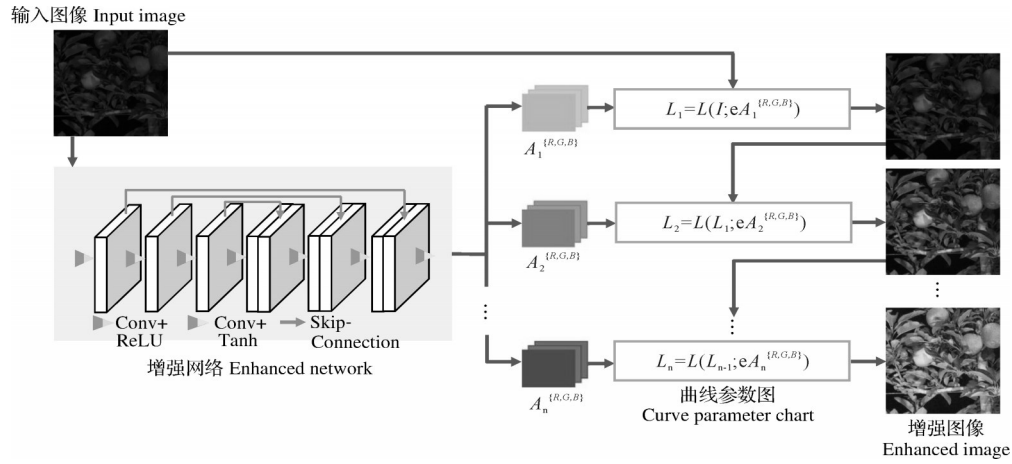


图3 Zero-DCE的框架

Fig.3 The framework of Zero-DCE

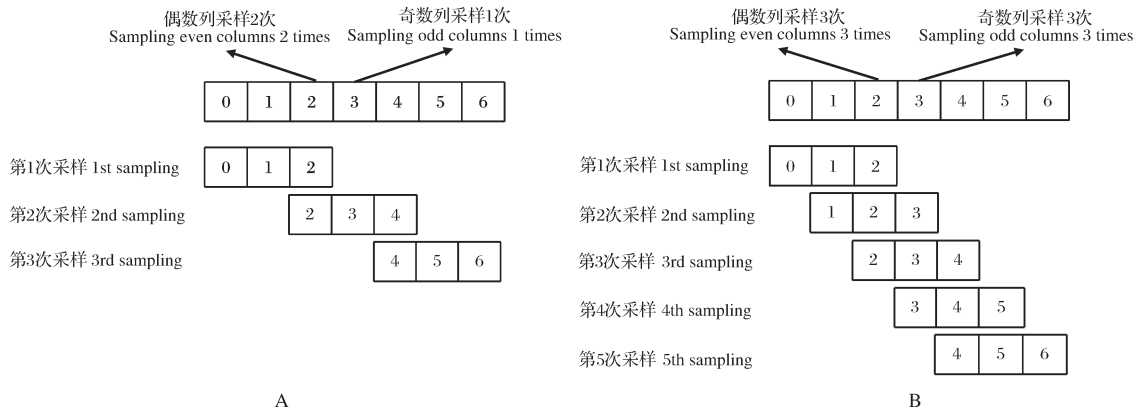
其中,Zero-DCE以低光图像作为输入,利用DCE-Net学习曲线参数图,通过亮度增强曲线(light enhancement curve)对输入的夜间图像进行像素级调整,并经过多次迭代优化输出最终的增强图像,提亮曲线对应的函数表达式为:

$$L_n(x) = L_{n-1}(x) + A_n(x)L_{n-1}(x)(1 - L_{n-1}(x)) \quad (1)$$

式(1)中, x 表示像素坐标, n 为迭代次数, A_n 是与输入图像尺寸相同的参数图。为得到准确的提亮曲线,Zero-DCE利用DCE-Net神经网络进行拟合,DCE-Net包含7个卷积层和3个对称跳跃连接层,其中所有卷积层都采用步长为1的常规 3×3 卷积核,为保持相邻像素间的联系,卷积层之后未使用BN层,隐藏层激活函数为ReLU,为确保输出范围落在 $[-1, 1]$ 内,输出层采用Tanh激活函数。

2)SPD-Conv提取细粒度特征。图像增强方法

可以有效降低夜间图像的检测难度,但由于图像分辨率较低,复杂情况下的识别仍然较为困难,因此需要加强模型细粒度特征提取能力。卷积核的步长大小对特征提取的细致程度有重要影响,相较于步长为2的卷积,步长为1的卷积能够使特征图的每个像素点被采样多次,并且奇数列和偶数列的采样次数保持一致,这有助于模型提取到细粒度特征,步长为2与步长为1的卷积核采样特点如图4所示。现有的YOLOv8n网络使用步长为2的卷积进行下采样和特征提取,这会造成判别特征的丢失。为解决这一问题,在YOLOv8n中引入SPD-Conv,以使模型提取到更充分、更细致的判别特征,SPD-Conv通过SPD(space to depth)保持下采样过程中信息的完整性并配合步长为1的卷积提取细粒度特征,其操作过程如图5所示。



A:步长为2的卷积 Convolution with a stride of 2;B:步长为1的卷积 Convolution with a stride of 1.

图4 步长为2与步长为1的卷积核采样特点(以列为例)

Fig.4 Sampling characteristics of convolutional kernels with steps of 2 and 1(using columns as an example)

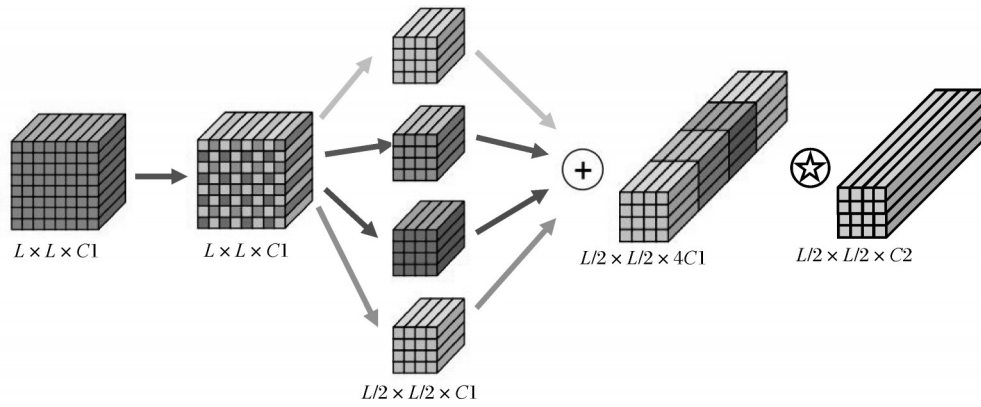


图5 SPD-Conv的操作过程

Fig.5 Operation process of SPD-Conv

图5中,设输入图像形状为 $L \times L \times C1$,SPD-Conv首先通过SPD得到4个形状为 $L/2 \times L/2 \times C1$ 的子特征图,其次将它们沿着通道维度拼接,获得形状为 $L/2 \times L/2 \times 4C1$ 的特征图,最后使用步长为1的卷积核提取细粒度特征并调整输出特征图的通道数。

3)改进后的网络。改进后网络的结构如图6所示,其中选择了一些层次的特征图进行可视化展示。首先Input通过Zero-DCE对夜间苹果图像增强,提升夜间图像的清晰度和辨识度,降低夜间场景下苹果的识别难度,为后续的图像处理任务奠定基础;其次Backbone使用融合SPD-Conv的CBS、C2f和SP-

PF等模块对增强后的图像进行多层次的特征提取,浅层特征图(14层)包含丰富的纹理特征,提供较多的位置信息,而深层特征图(20层)主要包含语义特征,提供较多的类别信息;然后Neck依次通过FPN、PAN对Backbone提取的不同层次的特征图加以利用,将纹理特征与语义特征融合,从而产生更高质量的综合特征;最后通过解耦头(doubled head)分别生成预测目标的类别概率和位置坐标,包括3个具有不同感受野的检测分支,分别对应25、28和31层的特征图,用以检测不同尺度的果实目标,并经过非极大值抑制得到符合需求的检测结果。

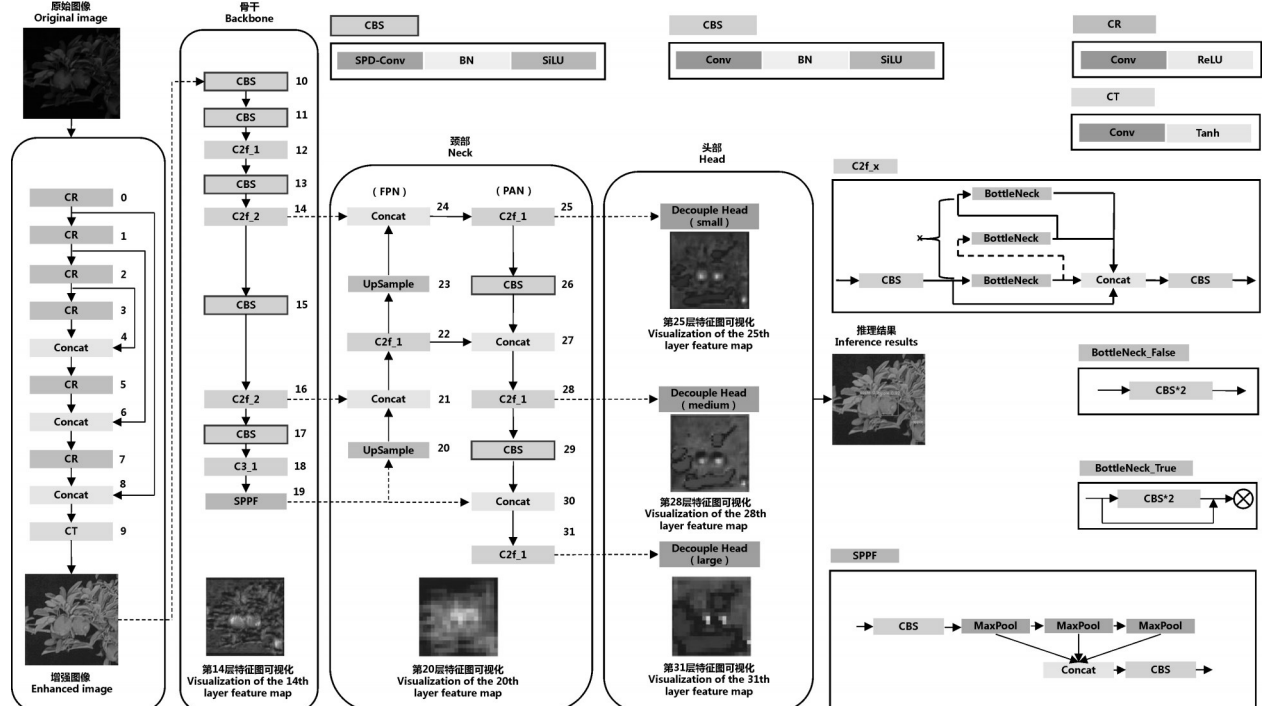


图6 改进的网络结构

Fig.6 Improved network structure

4) 迁移学习训练策略。在复杂的夜间环境中, 确保模型的通用性和稳定性显得至关重要, 这需要大量的、多样的训练数据作为支撑, 然而实际场景中数据获取比较困难, 且人工标注数据耗时费力, 迁移学习可以将一个领域上的预训练模型复用在新领域中, 从而降低数据获取和标注的成本。迁移学习的

核心思路是借助源域中已有的标注数据, 通过算法的开发, 最大化地利用这些知识, 该过程需要寻找源域和目标域之间的相似性, 以便将源域的知识有效地迁移到目标域中^[20-21]。为提升迁移学习效果, 本研究从源域和目标域对迁移学习策略分别做了改进, 本研究所提的迁移学习方法的过程如图7所示。

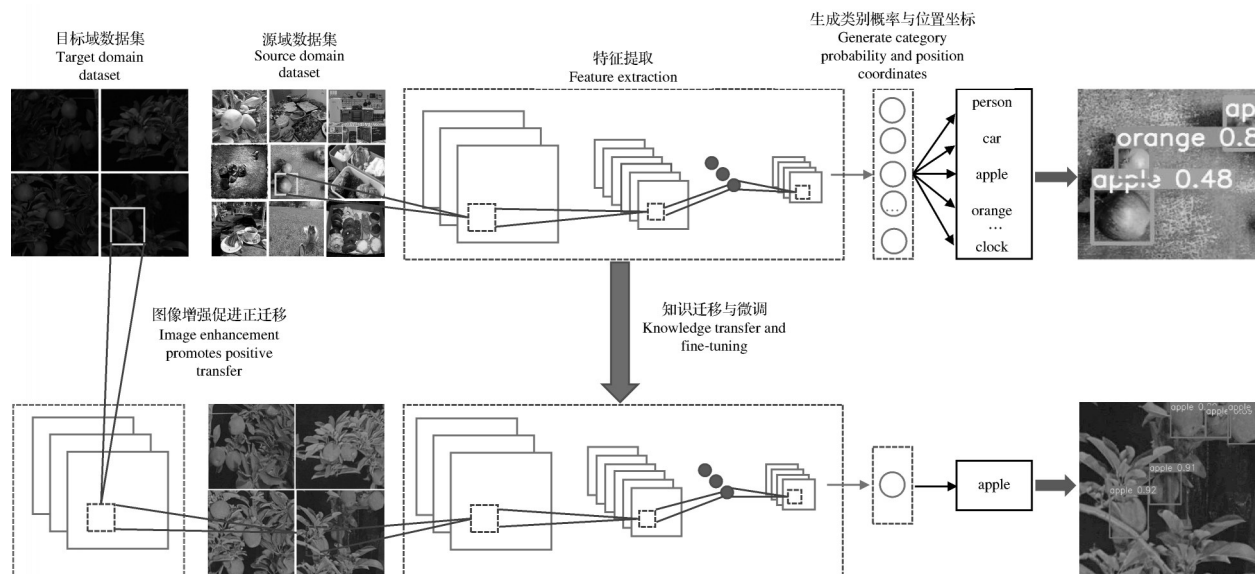


图7 本研究方法迁移学习过程

Fig.7 Transfer learning process of the method in this article

其中, 针对源域, 选取MS COCO数据集作为源域数据集, 与PASCAL VOC^[22]等数据集不同, MS COCO拥有80个类别的丰富数据, 将MS COCO作为源域有助于目标域模型习得底层的、普遍性的规则, 同时苹果属于其中的一个类别, 因此, 源域数据中也涵盖了大量的、各种环境下的日间苹果图像, 这使得目标域模型在源域模型上微调是容易泛化的。对于目标域, 则先将夜间苹果图像数据集送入ZeroDCE增强, 以增加其与日间苹果图像的相似度, 促进正迁移的发生, 进而在源域模型上微调目标域模型, 从而充分利用源域模型中的知识。

1.3 评估指标

本研究将平均精度均值(mean average precision, mAP)作为算法质量有效性的评估指标, mAP是基于精确率(percision, P)和召回率(recall, R)计算出来的, 精确率 P 表示预测的正确样本占总预测样本的比例, 召回率 R 表示预测的正确样本占实际正确样本的比例, mAP@0.5:0.95则是模型在不同交并比(intersection over union, IoU)阈值(0.5至0.95, 步长为0.05)下的平均值。另外, 将推理速度(单位为帧/s)作为算法实时性的评估指标。

1.4 试验环境与参数设置

模型训练硬件配置为Intel(R) Core(TM) i9-10900KF CPU @ 3.70 GHz, RAM为64 GB, GPU为NVIDIA GeForce RTX 3090 24 GB; 软件运行环境基于Windows 10(x64)操作系统, 采用Python编程语言和Pytorch深度学习框架, Python 3.9.18, Pytorch 1.13.1, Torchvision 0.14.1, Torchaudio 0.13.1, Cuda 11.7.0, Cudnn 8.4.0。

试验采用随机梯度下降法(stochastic gradient descent, SGD)作为优化器, 初始学习率和最终学习率设为0.01, SGD动量设为0.937, 权值衰减参数设为0.000 5, 批次大小设为16。运用warm up训练策略, warm up轮数设为3, 总训练轮数设为100。

2 结果与分析

2.1 训练过程

YOLOv8n基线模型和应用本研究方法得到的改进模型训练过程中在测试集上的mAP0.5:0.95变化如图8所示。由图8可知, 对于基线模型, 训练初期的较大学习率使mAP@0.5:0.95快速提升, 但也导致了模型的震荡, 当训练轮次达到20左右时, mAP@

0.5:0.95 增长速度明显放缓,模型逐步达到饱和,并随着学习率不断减小,曲线逐渐趋于平稳。相比之下,本研究所提方法采用了迁移学习策略,因此曲线起点更高、坡度更陡且渐近线更高,表明本研究方法初始性能更为出色、收敛速度更快且能达到更高的精度。此外,由于迁移学习允许使用较大的学习率训练,本研究方法从训练初期就表现出更好的稳定性。

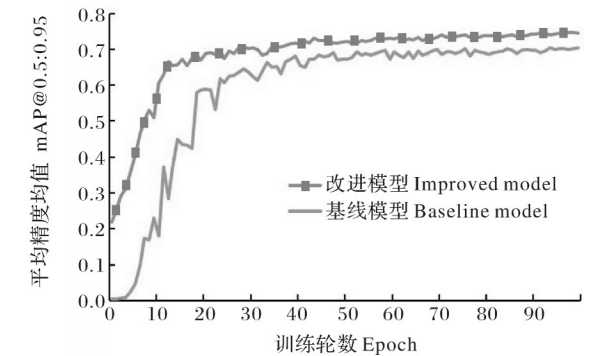


图8 基线模型和改进模型训练过程中在测试集上的 mAP@0.5:0.95 变化
Fig.8 Changes in mAP@0.5:0.95 on the test set during training of baseline and improved model

2.2 消融试验

本部分重点就图像增强和迁移学习的改进方法进行了消融试验研究。其中,将YOLOv8n作为基线模型,记为模型1(YOLOv8n);在模型1中应用SPD-Conv,得到模型2(YOLO-SPD);在模型1中应用Zero-DCE,得到模型3(YOLO-Zero);在模型1中应用迁移学习策略,得到模型4(YOLO-TL);在模型1中同时应用Zero-DCE和迁移学习策略,得到模型5(YOLO-Zero-TL);在模型5中应用SPD-Conv,得到模型6(YOLO-SPD-Zero-TL)。测试结果如表1所示。

由表1可知,模型2与模型1(YOLOv8n)相比,mAP@0.5:0.95提升0.7个百分点,这是因为SPD-Conv可以增强模型细粒度特征的提取能力,从而更准确地识别低分辨率的果实目标;模型3与模型1相比,mAP@0.5:0.95提升了0.9个百分点,说明Zero-CDE提升了夜间图像的清晰度和辨识度,降低苹果对象的检测难度;模型4与模型1相比,mAP@0.5:0.95提升2.3个百分点,表明迁移学习增强了模型的稳定性和通用性,提升了模型在复杂场景下的检测能力;模型5与模型1相比,在mAP@0.5:0.95表现上提升3.7个百分点,且提升幅度大于模型3与模型4增益的总和,这是因为经Zero-DCE处理的夜间苹果图像

表 1 消融试验结果				
Table 1 Results of ablation test				
模型 Model	精确率/% Precision	召回率/% Recall	平均精度 均值/% mAP@0.5: 0.95	推理速度/ (帧/s) Infer- ence speed
YOLOv8n	94.7	91.5	70.3	35
YOLO-SPD	94.8	92.0	71.0	30
YOLO-Zero	95.2	91.7	71.2	26
YOLO-TL	95.9	92.4	72.6	35
YOLO-Zero-TL	96.8	92.9	74.0	26
YOLO-SPD-Zero-TL	97.0	93.4	74.6	22

有效促进了正迁移;模型6与模型1相比,mAP@0.5:0.95提升4.3个百分点,说明SPD-Conv、Zero-CDE和迁移学习等改进是可以相互兼容的。此外,从各模型推理速度可以看出,迁移学习策略并未影响模型的推理速度,即使Zero-DCE和SPD-Conv在一定程度上降低了推理速度,但是在可接受范围内,而将两者结合使用却能够显著提高模型的精度,这为夜间环境下的果实检测提供了更有效的解决思路。

2.3 鲁棒性试验

为探究复杂场景下模型的检测效果,从测试集中随机挑选85幅图像分别输入到基线模型和应用本研究方法得到的改进模型中进行检测,其中含有重叠、遮挡、绿果、光线过暗的苹果实例个数分别为109、185、93和76,结果如表2所示。

由表2可知,相较于基线模型,改进后的模型在重叠情况下的正确识别率提升3.7个百分点,在遮挡情

表 2 鲁棒性试验结果				
Table 2 Robustness test results				
模型 Model	场景 Situation	总数 Total	正确识别数 Correct	正确识别率/% Accuracy
基线模型 Baseline model	重叠 Overlap	109	97	89.0
	遮挡 Occlusion	185	167	90.3
	绿果 Green fruit	93	81	87.1
	光线过暗 Dim light	76	68	89.5
改进模型 Improved model	重叠 Overlap	109	101	92.7
	遮挡 Occlusion	185	172	93.0
	绿果 Green fruit	93	85	91.4
	光线过暗 Dim light	76	72	94.7



A:重叠 Overlap;B:遮挡 Occlusion;C:绿果 Green fruit;D:光线过暗 Dim light.

图9 基线模型和本研究模型的复杂场景检测示例

Fig.9 Examples of complex scene detection using baseline and improved model

况下的正确识别率提升2.7个百分点,在绿果情况下的正确识别率提升4.3个百分点,在光线过暗情况下的正确识别率提升5.2个百分点。对应的检测示例如图9所示,其中正确检测结果采用红色矩形框标记,漏检、误检结果分别采用蓝色和黄色椭圆框标记(上侧、下侧图像分别表示基线模型和本研究所提模型的推理结果)。

由以上结果可知,在面对重叠、遮挡、绿果和光线过暗等复杂情形时,基线模型更容易出现误检和漏检,而本研究所提模型仍然保持较高的检测精度,证明本研究所提模型在稳定性和鲁棒性上表现更好,能够有效地降低夜间苹果检测的误检率和漏检率。

3 讨论

本研究提出了一种融合图像增强与迁移学习的YOLOv8n苹果夜间检测方法,该方法采用YO-

LOv8n模型,通过Zero-DCE增强夜间图像,并利用SPD-Conv提取细粒度特征,同时结合迁移学习策略进一步优化模型性能。试验结果表明,本研究所提方法的模型在夜间苹果数据集上的精确率为97.0%,召回率为93.4%,平均精度均值mAP@0.5:0.95为74.6%,推理速度为22帧/s,可以满足实时准确检测的需求。

苹果夜间检测具有一定的复杂性,尽管深度学习方法在夜间检测中展现出一定优势,但仍有一些不足。熊俊涛等^[13]和何斌等^[16]检测夜间果实时仅对YOLO模型本身改进,尽管取得了一定优化效果,但在密集、重叠、暗光或绿果等情景的果实识别率较低。相比之下,本研究在对YOLOv8n模型改进的基础上,针对性地引入输入端图像增强方法和迁移学习训练策略,有效地提升了模型在夜间复杂场景下检测苹果的稳定性的。

此外,本研究采用的Zero-DCE夜间图像增强方

法基于深度学习技术,可以避免传统图像处理方法^[11]导致的模型推理速度慢的问题,而迁移学习策略不增加模型的复杂度,因此本研究模型保持较高检测精度的同时,也具备较好的实时性。

消融试验显示,结合图像增强和迁移学习的方法的效果超过单独应用两者效果的总和。另外,改进后的模型展现出了较强的抗干扰能力,即使在重叠、遮挡、绿果和光线过暗等困难情形下,也能达到良好的识别效果。本研究所提方法为解决夜间果实检测问题提供了一种新的研究思路和技术手段,可为后续的相关研究提供参考。

参考文献 References

- [1] 李会宾,史云.果园采摘机器人研究综述[J].中国农业信息,2019,31(6):1-9.LI H B,SHI Y.Review on orchard harvesting robots[J].China agricultural informatics,2019,31(6):1-9(in Chinese with English abstract).
- [2] 中国苹果产业协会,中国农业大学农业规划设计院.2021年度中国苹果产业发展报告(总篇)精简版[J].中国果菜,2023,43(1):1-8.China Apple Industry Association,China Agricultural University Agricultural Planning and Design Institute.China apple industry development report in 2021(general),condensed version[J].China fruit & vegetable,2023,43(1):1-8(in Chinese with English abstract).
- [3] 宋怀波,尚钰莹,何东健.果实目标深度学习识别技术研究进展[J].农业机械学报,2023,54(1):1-19.SONG H B,SHANG Y Y,HE D J.Review on deep learning technology for fruit target recognition[J].Transactions of the CSAM,2023,54(1):1-19(in Chinese with English abstract).
- [4] WEI X Q,JIA K,LAN J H,et al.Automatic method of fruit object extraction under complex agricultural background for vision system of fruit picking robot[J].Optik,2014,125(19):5684-5689.
- [5] 周桂红,马帅,梁芳芳.基于改进YOLOv4模型的全景图像苹果识别[J].农业工程学报,2022,38(21):159-168.ZHOU G H,MA S,LIANG F F.Recognition of the apple in panoramic images based on improved YOLOv4 model[J].Transactions of the CSAE,2022,38(21):159-168(in Chinese with English abstract).
- [6] 袁紫薇.基于机器视觉的番茄收割机实时分拣系统研究[D].西安:长安大学,2017.YUAN Z W.Research on real-time sorting system of tomato harvester based on machine vision[D].Xi'an:Changan University,2017(in Chinese with English abstract).
- [7] 岳琳茜,李文宽,杨晓峰,等.基于改进YOLOv4的苹果检测与果径估测方法[J].激光杂志,2022,43(2):58-65.YUE L X,LI W K,YANG X F,et al.Apple detection and fruit diameter estimation method based on improved YOLOv4[J].Laser journal,2022,43(2):58-65(in Chinese with English abstract).
- [8] 陈青,殷程凯,郭自良,等.苹果采摘机器人关键技术研究现状与发展趋势[J].农业工程学报,2023,39(4):1-15.CHEN Q,YIN C K,GUO Z L,et al.Current status and future development of the key technologies for apple picking robots[J].Transactions of the CSAE,2023,39(4):1-15(in Chinese with English abstract).
- [9] FONT D,PALLEJÀ T,TRESANCHEZ M,et al.Counting red grapes in vineyards by detecting specular spherical reflection peaks in RGB images obtained at night with artificial illumination[J].Computers and electronics in agriculture,2014,108:105-111.
- [10] 赵德安,刘晓洋,陈玉,等.苹果采摘机器人夜间识别方法[J].农业机械学报,2015,46(3):15-22.ZHAO D A,LIU X Y,CHEN Y,et al.Image recognition at night for apple picking robot[J].Transactions of the CSAM,2015,46(3):15-22(in Chinese with English abstract).
- [11] 戴家裕.苹果采摘机器人夜间青苹果识别方法[D].杭州:浙江工业大学,2020.DAI J Y.Method for identify green apples at night by apple pic robot[D].Hangzhou:Zhejiang University of Technology,2020(in Chinese with English abstract).
- [12] REDMON J,DIVVALA S,GIRSHICK R,et al.You only look once: unified, real-time object detection[C]//2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR).June 27-30,2016.Las Vegas,NV,USA.Las Vegas: IEEE,2016:779-788.
- [13] 熊俊涛,郑镇辉,梁嘉恩,等.基于改进YOLO v3网络的夜间环境柑橘识别方法[J].农业机械学报,2020,51(4):199-206.XIONG J T,ZHENG Z H,LIANG J E,et al.Citrus detection method in night environment based on improved YOLO v3 network[J].Transactions of the CSAM,2020,51(4):199-206(in Chinese with English abstract).
- [14] HUANG G,LIU Z,VAN DER MAATEN L,et al.Densely connected convolutional networks [DB/OL]. arXiv, 2016: 1608.06993 [2023-12-01]. <http://arxiv.org/abs/1608.06993v5>.
- [15] 孙宝霞,梁翠晓,刘凯,等.夜间环境下树上柑橘表征缺陷深度学习检测方法[J].林业工程学报,2021,6(6):148-155.SUN B X,LIANG C X,LIU K,et al.Fluorescence detection of citrus characterization defects using SVM[J].Journal of forestry engineering,2021,6(6):148-155(in Chinese with English abstract).
- [16] 何斌,张亦博,龚健林,等.基于改进YOLOv5的夜间温室番茄果实快速识别[J].农业机械学报,2022,53(5):201-208.HE B,ZHANG Y B,GONG J L,et al.Fast recognition of tomato fruit in greenhouse at night based on improved YOLOv5[J].Transactions of the CSAM,2022,53(5):201-208(in Chinese with English abstract).
- [17] LIN T Y,MARIRE M,BELONGIE S,et al.Microsoft CO-

- CO: Common objects in context [C]//Computer Vision-EC CV 2014 Workshops. Zurich, Switzerland, September 6-7 and 12, 2014. Cham: Springer, 2014.
- [18] GUO C L, LI C Y, GUO J C, et al. Zero-reference deep curve estimation for low-light image enhancement [DB/OL]. arXiv, 2020: 2001.06826 [2023-12-01]. <https://doi.org/10.48550/arXiv.2001.06826>
- [19] SUNKARA R, LUO T. No more strided convolutions or pooling: a new CNN building block for low-resolution images and small objects [DB/OL]. arXiv, 2022: 2208.03641 [2023-12-01]. <https://doi.org/10.48550/arXiv.2208.03641>.
- [20] 庄福振, 罗平, 何清, 等. 迁移学习研究进展 [J]. 软件学报, 2015, 26(1): 26-39. ZHUANG F Z, LUO P, HE Q, et al. Survey on transfer learning research [J]. Journal of software, 2015, 26(1): 26-39 (in Chinese with English abstract).
- [21] PAN S J, YANG Q. A survey on transfer learning [J]. IEEE transactions on knowledge and data engineering, 2010, 22(10): 1345-1359.
- [22] EVERINGHAM M, ALI ESLAMI S M, VAN GOOL L, et al. The pascal visual object classes challenge: a retrospective [J]. International journal of computer vision, 2015, 111(1): 98-136.

A method for detecting apple at night based on YOLOv8n with fusion of image enhancement and transfer learning

TONG Zhaomao¹, CHEN Xuehai^{1,2}, MA Zhiyan^{1,2}, YANG Guangyou^{1,2}, ZHANG Can¹

1. *Institute of Agricultural Machinery, Hubei University of Technology, Wuhan 430068, China;*

2. *Hubei Engineering Research Center for Intellectualization of Agricultural Equipment, Wuhan 430068, China*

Abstract This paper proposed a method for detecting apple at night based on YOLOv8n with fusion of image enhancement and transfer learning to address the issues of low recognition rate and poor real-time performance of apples in nighttime scenarios. Firstly, embedding a Zero-DCE module in the front-end of YOLOv8n enhanced images of apple at night, presented the contours and details of apples more clearly, and reduced the difficulty of recognizing images of apple at night. Secondly, using SPD-Conv for down-sampling enhanced the ability of the model to extract fine-grained features. On this basis, transfer learning training strategy was used to solve the problem of small sample size in the dataset of apple at night. The MS COCO dataset containing categories of apple was selected as the source domain dataset. In term of the target domain dataset in nighttime scenarios, Zero-DCE was used to increase its similarity with images of apple during the day and finely tune the model of target domain on the model of source domain. Experiments were conducted on the image dataset of apple at night based on the method above. The results showed that the model accuracy P , a recall R , and an average accuracy mean $mAP@0.5$: 0.95 of method proposed was 97.0%, 93.4% and 74.6%, being 2.3, 1.9, and 4.3 percentages higher than that of the YOLOv8n original model. The inference speed of this model was 22 frames/s, meeting requirements of real-time detection. The results of the ablation experiment showed that the combined effect of image enhancement and transfer learning exceeded the sum of the effects when applied separately. The improved model performed better than the original model in dealing with complex situations including overlap, occlusion, green fruits, and dim lighting, and had good robustness.

Keywords apple; detection at night; image enhancement; transfer learning; YOLOv8

(责任编辑:陆文昌)