

郑旭康,李志忠,秦俊豪. 基于半监督学习的梨叶病害检测[J]. 江苏农业科学,2024,52(5):192-201.

doi:10.15889/j.issn.1002-1302.2024.05.029

# 基于半监督学习的梨叶病害检测

郑旭康,李志忠,秦俊豪

(广东工业大学信息工程学院,广东广州 510006)

**摘要:**植物病害的检测与识别是一个日益发展的研究领域,随着机器学习和深度学习概念的不断介入,为农业的发展提供了重要的技术支持。然而,目标检测技术存在着带标注数据获取成本高,且需要大量的人工来对数据进行标注等问题,给技术的实际应用造成了一定的阻碍。为解决在使用少量已标注数据及大量未标注数据进行训练模型从而提高准确率的问题,提出一种 YOLO 目标检测结合 self-training 半监督学习的方法,并且针对现有的 YOLO v3-Tiny 目标检测网络在半监督学习基础上准确率相比于监督学习较低的问题,对原有的 YOLO v3-Tiny 模型进行了改进。首先,使用空间金字塔池化结构对主干网络的多尺度特征进行融合;其次,将 YOLO v3-Tiny 检测头部分的标准卷积层替换成 GSConv;最后,运用 BiFPN 结构对中间部分的特征与检测头部分的多尺度特征进行双向融合。本研究提出的基于半监督学习的改进型 YOLO v3-Tiny 网络可以快速准确地检测出梨叶上的病斑,在试验中,准确度、召回率、平均精度分别达到 97.07%、93.78%、97.51%,对于快速准确地诊断出梨叶病斑的危害程度并且及时进行防治具有十分重要的意义。

**关键词:**深度学习;半监督学习;目标检测;梨;无标签数据;

**中图分类号:**TP391.41 **文献标志码:**A **文章编号:**1002-1302(2024)05-0192-09

目前,我国的水果产业发展非常迅速,其中梨是仅次于柑橘和苹果的世界第三大水果产业,而且种植面积以及产量均居于世界首位<sup>[1]</sup>。然而,近年来全球气候不断变化,由于我国南方地区雨季加长,降水量增大,导致梨主产区出现不同程度的早期梨树病害现象,从而致使产量的下降。因此,提前做出预测和采取有效的防治措施来及时准确地对梨叶发生病变程度进行诊断,对梨的生产具有重要的现实意义。梨树病害的前期症状大多由尺寸较小、颜色不均衡、纹理不够明显的类圆形斑点构成,随着后期病变程度的加重,病斑会越来越多,也越来越大,多个病斑通常还会融合形成显著且形状不规则的深色斑块分布在整個叶片上。因此,防治的最好时机是病害前期。

近年来,目标检测在计算机视觉领域不断发展,成为比较热门的图像识别方法。目标检测在自

动导航、人脸识别、行人检测、农业病害检测等领域得到广泛的应用<sup>[2-4]</sup>,例如,胡瓜顿绥螨自动检测、苹果病害识别与麦冬叶片病害识别等<sup>[5-7]</sup>。然而,当前的目标检测算法大多严重依赖已经标注好的数据来训练模型,并且这些数据必须具备多样性和丰富的信息,以提供更高的准确度,才能将训练好的模型部署进行实际的预测。在实际的检测任务中,存在着数据采集困难、数据通常是复杂多变的、无标签数据大量存在,需要大量的人工耗时数周甚至更长的时间来完成图像中梨叶病害种类和位置等信息的标注。这些问题都在很大程度上影响了检测技术在农作物病害检测中的应用。为解决这些无标签数据大量存在以及有标签数据集获取成本较高的问题,同时,要让已有的模型快速地利用少量的无标签数据来应对各种变化的梨叶病害图像数据,即使用少量已标注数据对模型进行训练,降低模型训练的成本,并能够对梨叶的各种病害进行精准检测。因此,构建出适用于梨叶病斑检测数据集也是十分有意义的。而无论是具有标签数据还是无标签数据,都具有相似的特征分布,因此,使用半监督学习(SSL)策略能够充分利用大量无标签数据来提高模型的表征学习能力<sup>[8]</sup>。

目前,半监督学习(SSL)的研究取得了显著成

收稿日期:2023-03-27

基金项目:广东省自然科学基金(编号:2019A151011371);广东省省级科技计划(产学研)(编号:2016B090918031)。

作者简介:郑旭康(1999—),男,广西柳州人,硕士,研究方向为计算机视觉领域。E-mail:1412154348@qq.com。

通信作者:李志忠,博士,副教授,硕士生导师,研究方向为数字电源技术、人工智能。E-mail:leezzl@163.com。

果<sup>[8]</sup>。这些方法中有许多类似共享的基本技术,比如深度学习中的熵最小化、伪标签或一致性正则化<sup>[9]</sup>。伪标签指的是使用从模型本身预测中产生的伪标签目标训练带有未标注数据的分类器<sup>[10]</sup>。一致性正则化的核心思想是对于模型的一个输入,即使其受到了微小的干扰,模型对其的预测结果应该是一致的<sup>[9]</sup>,即将模型输出的 softmax 概率分布作为软伪标签<sup>[11]</sup>,将经过 argmax 或者 one\_hot 得到的预测视为硬伪标签<sup>[12]</sup>。无论是通过软伪标签<sup>[11]</sup>还是硬伪标签<sup>[12]</sup>,生成多视图的有效方法包括:不同强度的输入数据增强<sup>[13]</sup>、网络层内的标准 dropout<sup>[14]</sup>和随机深度<sup>[15]</sup>。

目前主流目标检测框架主要为单阶段检测的 YOLO 系列算法<sup>[16]</sup>、两阶段检测的 Faster RCNN<sup>[17]</sup>和 SSD<sup>[18]</sup>等系列算法。虽然这些算法在有监督学习的技术上取得了很大的成功,但这些检测器是用大量完全注释的数据训练的,注释成本很高。因此,人们提出了许多降低标注成本的方法。比如引入半监督检测 (SSOD)<sup>[19]</sup>和弱监督检测 (WSOD)<sup>[20]</sup>,以降低数据标注的巨大成本。半监督目标检测方法是一种学习少量标记图像和大量未标记图像的检测器。Jeong 等使用一致性约束利用未标记数据进行目标检测<sup>[19]</sup>,而弱监督检测方法利用带有弱标注的大量数据,比如图像标签。此外,为了追求有监督检测的性能和保持较低的标注成本,Bilen 等通过结合区域分类和选择,研究了弱半监督检测方法 (WSSOD),该方法使用小的盒级标记图像和大量的弱标记图像训练检测器<sup>[20]</sup>。

自训练在半监督学习中被广泛研究<sup>[21]</sup>。文献[22]概述了不同的自训练方法。最近,深度神经网络重新激起了人们对自训练的兴趣。固定特征的自训练和深度自训练之间的细小区别是深度自训练涉及到嵌入的学习,在领域对齐方面比分类层适应具有更大的灵活性。在此背景下,本研究探讨了半监督学习方法结合改进后的 YOLO v3 - Tiny 在梨叶病斑检测中的实际应用,提出了一种可称为改进型 YOLO v3 - Tiny 结合半监督学习检测的算法,可以利用大约 5% 的已标注数据,与现有最新的检测网络相比,改进后的模型精度最高。

## 1 方法

### 1.1 基于改进 YOLO v3 - Tiny 目标检测原理

YOLO v3 - Tiny<sup>[23]</sup>是 YOLO v3<sup>[24]</sup>网络的轻量

化版本,具有高精度、轻量且实时性高的特点。不同于常规的 YOLO v3 应用中需要用很深的模型来检测几十个甚至几百个目标,对于农业上算力差的嵌入式设备而言,常规的 YOLO v3 检测算法明显满足不了实时性检测要求,然而轻量化的 YOLO v3 - Tiny 算法,检测速度显著提升,满足设备的实时性。在梨叶病斑检测任务当中,除了要满足实时性等技术指标之外,也应该充分地考虑技术在实际应用场景中的情况,尤其是在半监督学习下采用较少样本对模型进行训练来满足高准确率的要求。

由图 1 可知,模型会将输入的图像划分为  $N \times N$  个大小相等的栅格,每个栅格负责检测落在该格子中的病斑区域以及预测框和其置信度,若栅格包含检测目标的中心位置,则判定该栅格能识别该目标,置信度则表示该格子含有病斑区域的可能性以及坐标预测的准确性。可以通过以下公式计算检测框的位置与大小以及置信度。

$$B_x = \sigma(t_x) + c_x; \quad (1)$$

$$B_y = \sigma(t_y) + c_y; \quad (2)$$

$$B_w = P_w e^{t_w}; \quad (3)$$

$$B_h = P_h e^{t_h}; \quad (4)$$

$$F = P_r(obj) \times IOU(box, obj)。 \quad (5)$$

式中:  $(B_x, B_y)$  是预测边界框的中心坐标;  $(B_w, B_h)$  为预测边界框 (Anchor) 的宽和高;  $(c_x, c_y)$  表示栅格的边距;  $(P_w, P_h)$  为预设边界框的宽和高;  $(t_x, t_y)$  和  $(t_w, t_h)$  分别为网络预测的边界框中心偏移量以及宽高缩放比;  $IOU$  为预测的边界框和真实框之间的交并比,  $F$  表示预测边界框的置信值。

### 1.2 改进 YOLO v3 - Tiny 网络

完全去除 YOLO v3 中原有残差模块的 YOLO v3 - Tiny 不仅减少了大量多尺度特征融合,而且深层特征图中缺少浅层的细节特征,在种植园环境下整个网络在前向传播过程中,被遮挡部分以及受背景干扰的目标本身特征就较少,而 YOLO v3 - Tiny 由于残差结构的缺失,深层的特征图中对图像某些细节信息表达能力减弱,导致目标中某些细节特征在整个模型的深层网络中消失,使得 YOLO v3 - Tiny 在半监督学习情况下很难达到监督学习的效果<sup>[23]</sup>。因此,本研究对 YOLO v3 - Tiny 进行改进和增强,加强 YOLO v3 - Tiny 目标检测模型的特征融合以及检测能力<sup>[23]</sup>,使得该模型不断接近监督学习情况下所能达到的效果。整体改进后的 YOLO v3 - Tiny 结构如图 2 所示,本研究主要对检测头部分进

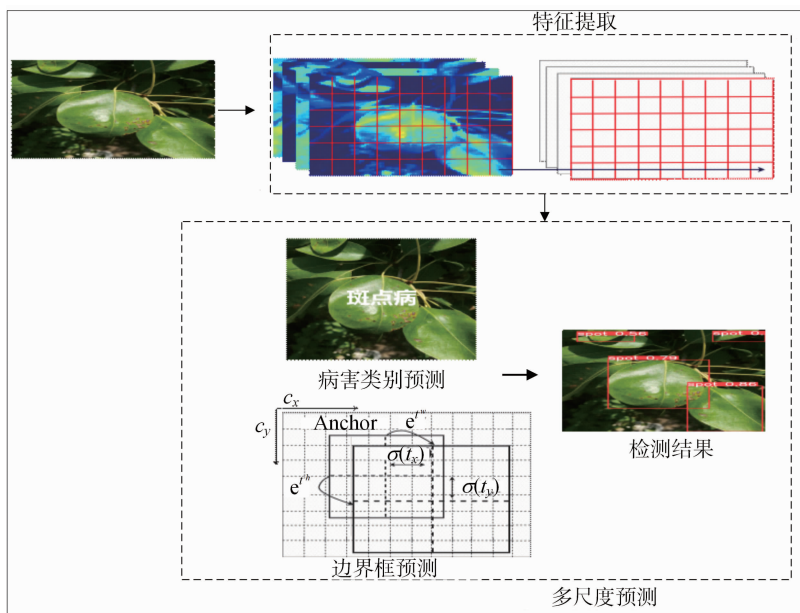


图1 改进 YOLOv3-tiny 检测原理

行了改进,在检测头和主干网络之间加入空间金字塔池化模块 SPPCSPC,对主干网络中的多尺度特征进行融合,紧接着再连向检测头部分。然后使用 BiFPN 结构替换原有的 Concat 操作,将主干网的第 5 个卷积块以及检测头第 3 个 GSConv 经上采样后提取到的特征进行通道数压缩统一后送入 BiFPN 模

块中进行深浅层特征双向融合,其目的是将主干网络中中间部分的特征与检测头部分的多尺度特征进行再次融合。最终,使用改进后的模型在半监督的基础上与原有的 YOLO v3 - Tiny<sup>[23]</sup> 以及 YOLO v7 - Tiny<sup>[25]</sup>、YOLO v5s<sup>[26]</sup> 等轻量型的目标检测网络进行对比试验,并且取得了不弱于这些网络的性能。

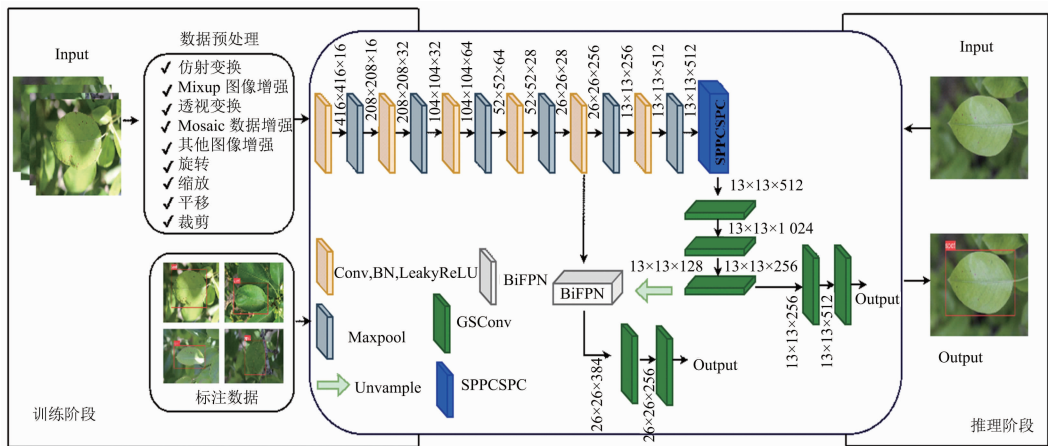


图2 改进 YOLO v3-tiny 结构

1.2.1 SPPCSPC 空间金字塔池化 虽然 YOLO v3 - Tiny 中存在 FPN 特征融合结构<sup>[23]</sup>,但对于部分面积较小的梨叶而言,模型很难对其特征进行充分的提取,从而导致漏检与错检等精度低的情况。因此,为提高检测精度,在 YOLO v3 - Tiny 的主干网络后面,即检测头前引入 SPPCSPC 空间金字塔池化结构。该结构如图 3 所示,该结构主要是通过最大池化层来获取不同感受野,使得算法能够适应于不同尺度的特征图。整体结构首先将特征分为 2 个部

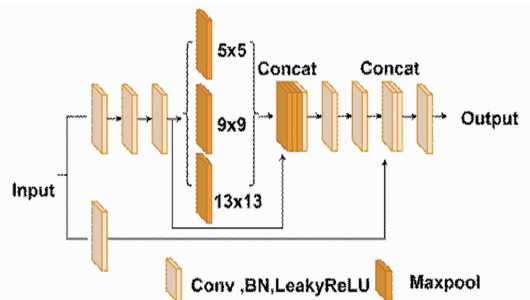


图3 SPPCSPC 结构

分,一个部分进行常规的卷积处理,另外一个部分进行 SPP 结构的处理,最后将这 2 个部分进行 Concat 操作合并在一起,这样便能够减少将近一半的计算量,使得速度变快的同时,精度有所提升。

1.2.2 BiFPN 模型结构 随着网络模型的层数不断叠加,特征语义也由低维度转向高维度,每添加 1 层网络都会导致一定程度上的特征丢失,因此,需要

从骨干网络中间的各层提取出不同尺度的特征,进行自上向下的特征融合。构造出一种直接有效的特征金字塔结构的特征融合方式对不同网络层中的特征进行融合,以此来保存和丰富特征语义信息<sup>[27]</sup>。本研究使用一种运用双向融合结构的 BiFPN<sup>[28]</sup>对深浅层特征进行由上而下与由下而上的双向融合,其结构如图 4 所示。

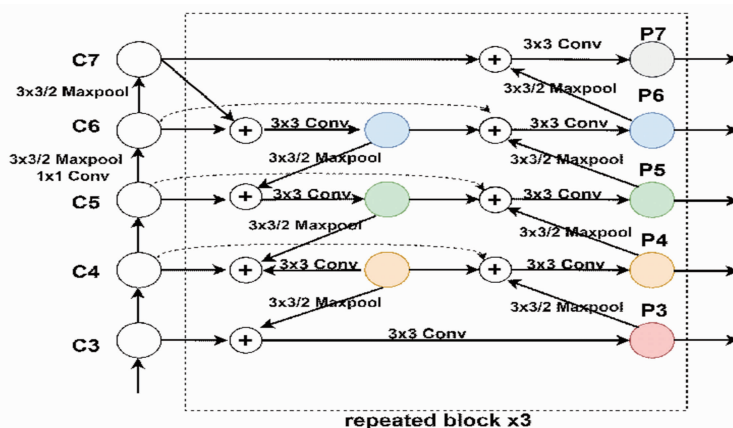


图4 BiFPN 结构

该模型通过构造由顶向底以及由底向顶的双向通道,来自骨干特征提取层中不同尺度的特征信息进行融合,再通过上、下采样操作来统一特征图的尺度,并在相同尺度的特征图之间添加横向连接,减少因层数过多而造成的特征信息丢失。

1.2.3 GSConv 模型 对于农业设备来说,一个巨大的模型很难满足实时性检测的要求。而且,使用大量可分离卷积层来构造轻量级模型很难满足精度要求。轻量化模型的设计主要是通过使用深度可分离卷积(DWConv)操作来减少参数量和降低计算复杂度,许多轻量化模型从神经网络搭建开始到结束只使用了 DWConv 这样类似的思维来设计基本架构,效果虽然明显,但是缺点也很明显:输入图像的通道信息在运算过程中是分离的,从而导致 DWConv 的特征提取和融合能力比标准卷积(Conv)低得多,且这种缺陷会直接在模型中放大,无论是用于分类还是检测,通过通道混洗(Shuffle)的输出通道生成的特征图仍然是深度分离的。因此,本研究使用了一种新卷积 GSConv<sup>[29]</sup>。由图 5 可知,使用 Shuffle 的方式将来自 Conv 的特征信息渗透到 DWConv 所生成特征信息的每个部分,使得 Conv 所生成的特征信息完全融入到 DWConv 的输出中。但如果将其在加目标检测模型的所有阶段的话,会使得网络层会加深,而阻碍数据流的流动,导致推理

速度变慢。因此只在检测头部分使用该模型,使得模型在保持模型精度的前提下减轻模型的重量,在模型的准确性和实时性之间实现了极好的权衡。

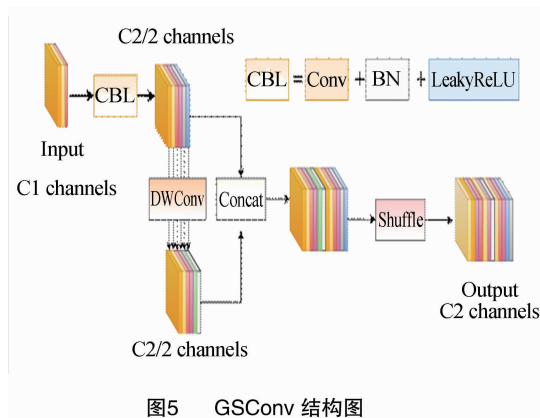


图5 GSConv 结构图

### 1.3 半监督 self-training 训练策略

半监督学习作为一种能够同时利用大量的未标注数据以及少量的已标注数据来改善深度学习性能的方法,是基于监督学习与无监督学习之间的一种学习范式。self-training 方法是半监督学习中最常用的方法之一<sup>[22]</sup>,在深度学习和模式识别等领域有着广泛的应用,本研究将该算法引入梨叶病斑检测训练过程。该算法基本思想是先在少量已标注的数据上训练得到一个基础模型,然后使用该模型对未标注数据进行预测,根据某种算法将预测结果中准确度较高的结果保留作为伪标签,再将伪标



签数据和已标注数据结合起来使用标准的监督学

习方法对模型进行再训练,其流程如图 6 所示。

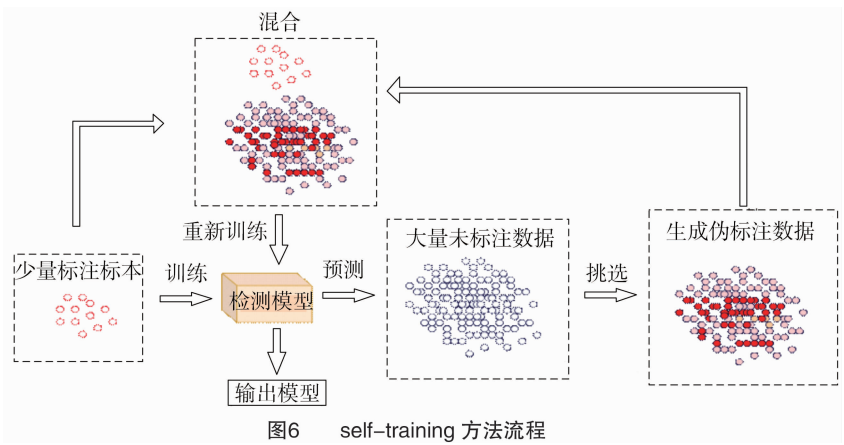


图6 self-training 方法流程

通常,为保证模型能够从已标注数据中学到更多的特征信息,在最初的  $N$  个迭代次数中,将  $\omega(t)$  设置为 0,随后慢慢增加到  $M$  个迭代次数后保持不变。如下式:

$$\omega(t) = \begin{cases} 0 & t < T_1 \\ \frac{t - T_1}{T_2 - T_1} & T_1 < t < T_2。 \\ \omega_f & T_2 \leq t \end{cases} \quad (6)$$

为提高大量未标注数据的利用率,从而提取出有效的图像特征信息,可使用 self - training 策略<sup>[22]</sup>。首先使用训练好的模型在未标注的数据上进行检测,选择置信值比较高的伪标注数据加入训练数据集中,直到无标注的数据集中的数据不再变化为止,详细流程可以参考以下算法。

输入:  $L = \{x_{\text{label}}, y_{\text{label}}\}$ ,  $U = \{x_{\text{unlabel}}\}$ ,  $L$  表示标签数据,  $U$  表示未标注数据。

步骤 1:  $U \gg L$ ;

步骤 2: 进入循环: 当未标注数据不包含在  $U$  中或者未标注数据不再变化时作出以下行为;

步骤 3: 从  $(x_{\text{label}}, y_{\text{label}})$  中选取数据训练模型;

步骤 4: 使用模型预测  $x_{\text{unlabel}}$ , 获得伪标签数据  $\{(x_{\text{unlabel}}, y_{\text{unlabel}})\}$ ;

步骤 5: 筛选出置信样本  $(x_{\text{conf}}, y_{\text{conf}})$ ,  $(x_{\text{conf}}, y_{\text{conf}}) \in (x_{\text{unlabel}}, y_{\text{unlabel}})$ ;

步骤 6: 合并最新数据  $(x_{\text{label}}, y_{\text{label}}) \cup (x_{\text{conf}}, y_{\text{conf}})$ ;

步骤 7: 结束循环。

以上的流程可以简单理解为将初始的带标注数据集作为初始训练集  $L$ , 根据训练集训练得到一个初始模型。利用模型对无标签数据集  $U$  中的样本进行检测, 选出置信度较高的伪标注样本  $(X_{\text{conf}}, Y_{\text{conf}})$ , 并将其加入到有标签数据集中, 根据新得到

的带有伪标注数据的训练集再次训练成新的模型, 重复步骤 2 至步骤 6 直到所有无标签样本都被标记完成, 最后得到最优的模型。总损失函数为有标注损失和无标注损失的加权和, 公式如下:

$$Loss = \frac{1}{n} \sum_{m=1}^n \sum_{i=1}^L CE\_Loss(y_i^m, f_i^m) + \omega(t) \frac{1}{n'} \sum_{m=1}^{n'} \sum_{i=1}^L CE\_Loss(y_i'^m, f_i'^m); \quad (7)$$

式中:  $n$  表示标记数据中的小批量数据;  $n'$  表示未标记数据;  $f$  表示标记数据中  $m$  个样本的输出单位;  $y$  表示其标签;  $f'$  表示未标记的数据;  $y'$  表示非标记数据的伪标签;  $CE\_Loss$  为交叉熵损失函数。

## 2 试验

### 2.1 图像数据集预处理

本研究引入了一个用于诊断和检测植物症状的现场数据集, 称为 DiaMOS Plant<sup>[30]</sup>。这是一个试点数据集, 包含梨树整个生长季节(2—7 月)不同光照条件下所拍摄的梨叶图像的照片, 该数据集适用于检测任务执行深度学习方法。该数据集已标注和未标注的数据划分情况见表 1、表 2。本试验选取该数据集当中季节性病斑的图像来验证该算法的可行性以及使用蛭蟥损伤梨叶数据集验证该方法的泛化能力。本试验中图像的分辨率统一设置为  $512 \times 512$ , 为增加图像信息, 使模型的初始训练能够获得更好的效果, 在训练时都采用了随机遮挡、裁

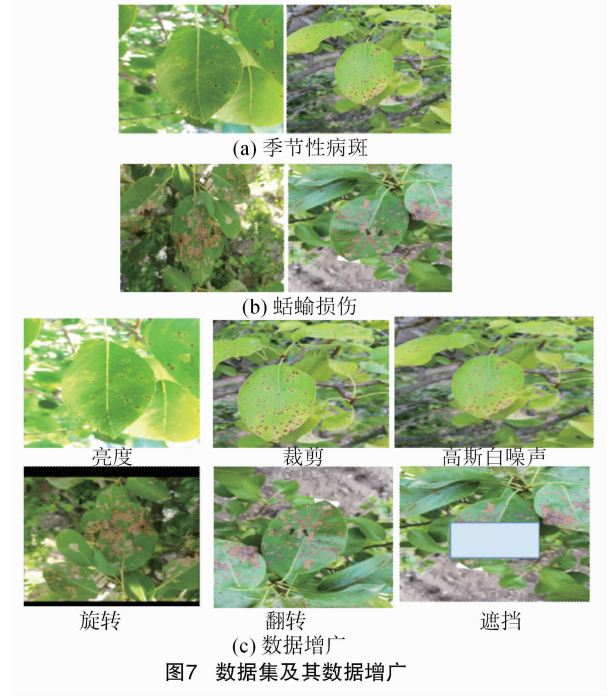
表 1 季节性病斑的标注数据划分

数据集名称	图片数量 (张)	标注形式
训练集	50	已标注
验证集	178	已标注
测试集	50	未标注
无标签数据集	606	未标注

表 2 蛭螭损伤标注数据划分

数据集名称	图片数量(张)	标注形式
训练集	100	已标注
验证集	508	已标注
测试集	100	未标注
无标签数据集	1 316	未标注

剪、色彩扭曲、加噪、旋转、图像归一化等图像增强技术来扩展未标注的训练集(图7)。



2.2 实施细节

本试验的时间为2022年12月5日,在笔者所在实验室搭建的深度学习平台:64位的Window 10操作系统下进行,编程环境Python 3.7,CUDA11.0的GPU加速库。采用深度学习框架Pytorch,显卡NVIDIA GeForce RTX-1050Ti,显存4 GB。批量大小设置为4,学习速率为 $1\times10^{-2}$ ,动量设置为0.937,衰减权重为 $1\times10^{-5}$ 的SGD优化器,迭代次数设为400次,训练过程使用Warm up方法加快模型收敛速度。

2.3 评价指标

本研究使用的是目标检测模型最常用的评价指标对试验结果进行评估,如准确度( precision,  $P, \%$ )和召回率( recall,  $R, \%$ ),为了度量准确度和召回率在实际检测过程中的表现能力,本研究使用  $IOU$  最常用的阈值为0.5的平均精度均值( mean average precision,  $mAP, \%$ ),计算公式如下:

$$P = \frac{TP}{TP + FP} \times 100\%; \tag{8}$$

$$R = \frac{TP}{TP + FN} \times 100\%; \tag{9}$$

$$mAP = \frac{\int_0^1 P(R) d(R)}{N} \tag{10}$$

其中,公式中的  $TP$  为检测正确样本的数量;  $FP$  为检测错误样本的数量;  $FN$  为漏检的目标样本数量;  $N$  为检测样本的类别总数;  $P(x)$  为以召回率  $x$  为自变量,精确率  $P$  为因变量的函数。

3 结果与分析

3.1 试验结果

为验证本试验改进后模型的检测效果,在同等条件下进行了模型对比试验,以验证该方法的可行性。采用现有模型中较为先进的目标检测模型网络分别在有监督学习和半监督学习的基础上进行对比试验,最后用测试集进行检测,并进行结果可视化。由表3可知,在监督学习的情况下,各项指标的差距并不大,基本都能准确地检测出病斑。但在半监督学习下即只用50(5%)张的带标注数据进行模型训练时,各项指标都有显著下降,因此,本研究对相比于其他模型结构较为简单的YOLO v3-Tiny进行了改进,经过改进之后,模型的各项指标在不断接近监督学习。

由图8可知,当迭代次数达到400次时,各个模型指标的数值已经达到了稳定状态,模型训练过程

表 3 季节性病斑的试验结果对比

训练方式	网络模型	准确度 (%)	召回率 (%)	平均精度 (%)	推理速度 (张/s)	训练平均时长 (h)
监督学习	YOLO v7-Tiny	97.71	97.18	99.32	0.0 118	60
	YOLO v3-Tiny	96.57	95.47	98.72	0.0 094	
	YOLO v5s	97.12	96.61	99.02	0.0 103	
半监督学习( self-training)	YOLO v7-Tiny	90.95	90.93	96.53	0.0 118	36
	YOLO v3-Tiny	86.97	87.06	92.92	0.0 094	
	YOLO v5s	77.57	72.29	77.48	0.0 103	
	改进 YOLO v3-Tiny	97.07	93.78	97.51	0.0 113	

完成。指标结果如表 3 所示,改进的 YOLO v3 - Tiny 所带来的效果很明显,各项指标都有显著提升,比如平均精度提升了 4.59 百分点,准确度提升了超过 10 百分点,召回率提升了 6.72 百分点。比起目前较新的 YOLO v7 - Tiny 分别提升了 0.98、6.12、2.85 百分点。测试集中季节性病斑试验检测结果

见图 9,图中检测框左上方的单词代表检测病斑类别,以及左上方的数字表示置信度。由图 9 可知,改进后的模型无论是在较暗的环境还是有些叶面部分被遮挡的情况下都能较为精准地检测出病斑位置且置信度逼近监督学习。

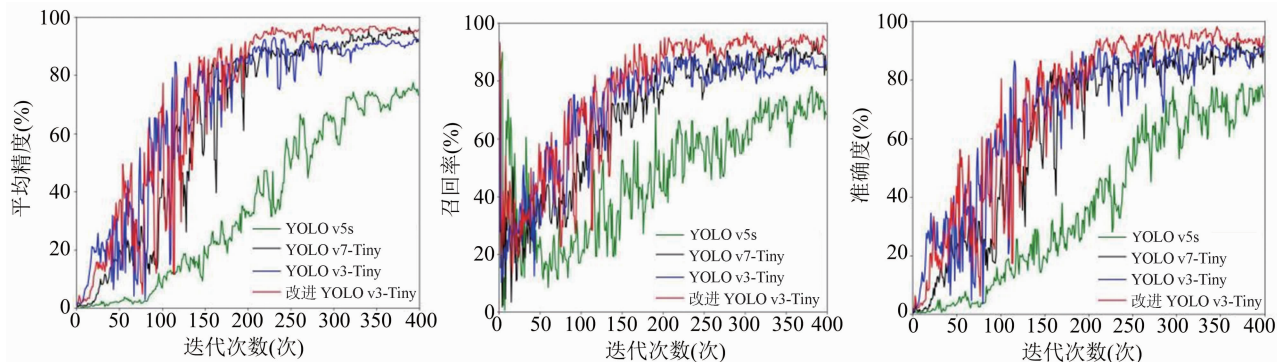


图8 季节性病斑试验指标对比

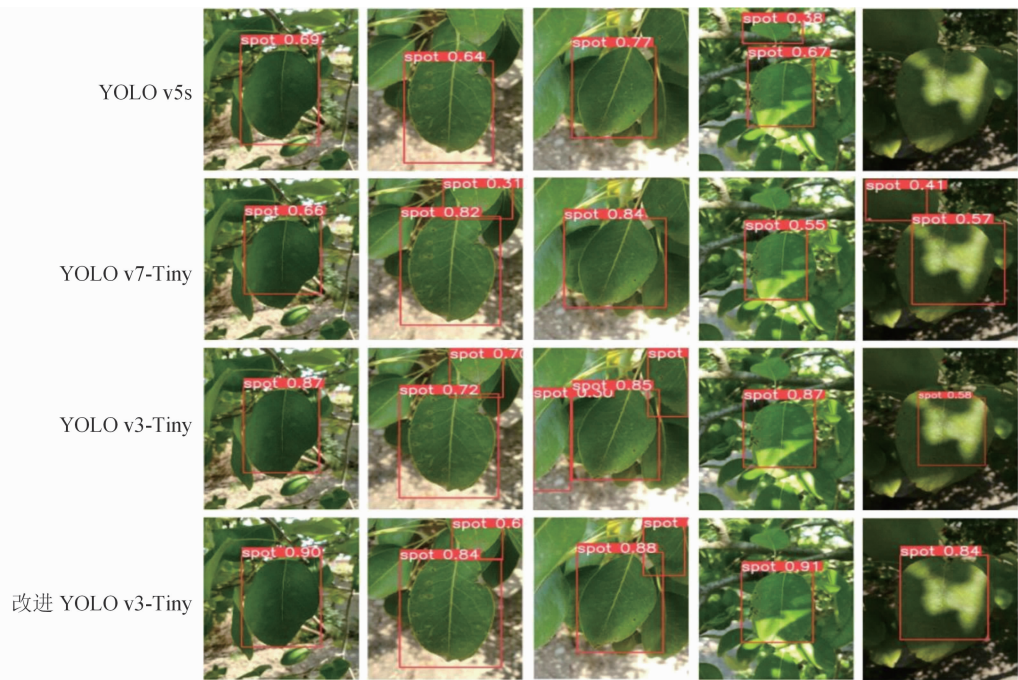


图9 测试集中季节性病斑试验结果对比

为进一步验证该方法的可行性以及泛化能力,继续使用表 2 的蛭螭损伤数据集进行试验,结果如图 10、表 4 所示。随着数据集数目的增大,以及已标注数据的增多,各个模型都能达到极好的效果,且本研究的模型依然有着优于其他模型的效果。其中准确度、召回率、平均精度均值分别达到了 97.46%、98.42%、99.20%,均超过 YOLO v7 - Tiny 以及原来的 YOLO v3 - Tiny 模型。由图 11 可知,在仅使用 100(5%)张带标注数据的情况下,各个模型

基本都能精确检测出蛭螭损伤的叶片,且本研究的模型检测出损伤叶片的置信度基本上都高于其他模型,从第 4 列的图片来看,本研究模型未出现漏检的情况,检测能力明显增强。显然本试验进一步证明了半监督学习算法以及本研究改进模型的可行性。

### 3.2 试验分析

消融试验通常用于研究网络的特定子模块或者策略及参数等对模型产生的作用,在神经网络结构的设计过程中具有重要的意义<sup>[31]</sup>。为验证在半



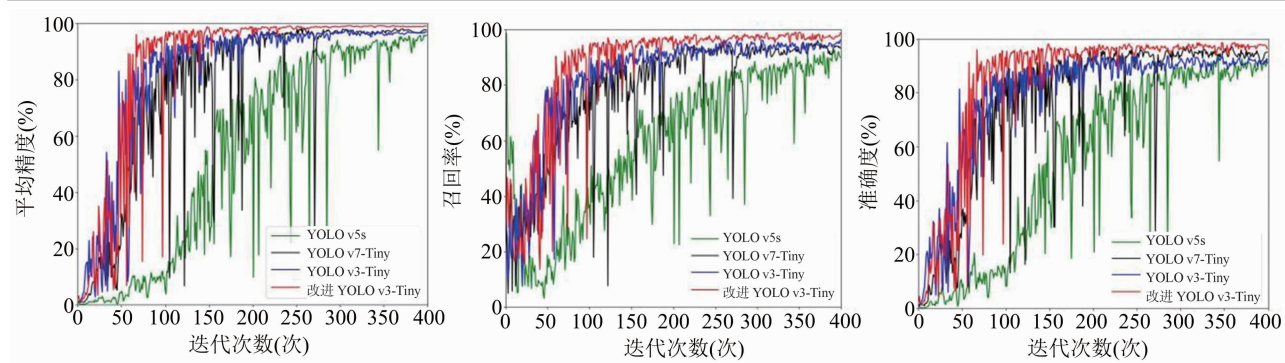


图10 蛭螭损伤试验指标对比

表 4 季节性病斑的试验结果对比

网络模型	准确度 (%)	召回率 (%)	平均精度 (%)	推理速度 (张/s)
YOLO v7 - Tiny	95.44	95.07	98.35	0.012 3
YOLO v3 - Tiny	92.42	96.04	97.90	0.011 2
YOLO v5s	92.67	89.96	95.95	0.012 0
改进 YOLO v3 - Tiny	97.46	98.42	99.20	0.011 3

监督学习中模型的改进效果,本研究设置了不同模块的消融试验,即嵌入不同模块来改进 YOLO v3 - Tiny 目标检测模型的策略进行训练,以验证改进的策略是否具有促进作用。最终,试验结果见表 5。首先,在原 YOLO v3 - Tiny 主干网络的基础上引入 SPPCSPC(空间金字塔池化结构模块),以此来促进主干网络之间的特征信息进行更加充分的融合。



图11 测试集中的蛭螭损伤梨叶数据集对比结果

其次,将检测头部分的 CBL 模块使用 GSConv 来替换,使得模型的准确性和实时性之间实现了极好的权衡。最后还加入了 BiFPN 模块来使得主干网络的特征信息与检测头网络的特征信息进一步地融合。本研究继续使用季节性病斑数据集进行试验,以每个目标检测模型的平均精度均值、准确度以及

召回率为主要指标,从表中不难发现,在结合不同模块之后各项指标都有了显著的提升。

4 结论

基于半监督学习框架的梨叶病斑检测算法在一定程度上解决了大量未标注的梨叶数据无法利



表 5 加入不同模块的试验结果

网络模型	准确度 (%)	召回率 (%)	平均精度 (%)
YOLO v3 - Tiny( darknet - 19 + FPN)	87.06	92.92	86.97
darknet - 19 + BiFPN	85.04	90.39	92.23
darknet - 19 + BiFPN + GSConv	90.47	94.35	96.28
darknet - 19 + BiFPN + GSConv + SPPCSPC	97.07	93.78	97.51

用、已标注的梨叶数据不足等实际问题。同时,针对半监督学习下的模型准确度比有监督学习下低的问题,对 YOLO v3 - Tiny 模型进行了改进,试验结果的各项指标都有了显著提高,为基于图像的植物病害检测应用提供了一种有效的解决方案。但依然存在着在恶劣环境下出现漏检误检的情况,以及难以达到甚至超过监督学习所能达到的准确度以及农业设备的部署等,这些问题也将成为下一阶段的研究重点。

参考文献:

[1] Wu J, Wang Z W, Shi Z B, et al. The genome of the pear (*Pyrus bretschneideri* Rehd.)[J]. Genome Research, 2013, 23(2): 396 - 408.

[2] Yang W L, Fan S S, Xu S X, et al. Autonomous underwater vehicle navigation using sonar image matching based on convolutional neural network[J]. IFAC - Papers OnLine, 2019, 52(21): 156 - 162.

[3] 马博宇, 尉寅玮. 基于 AdaBoost 算法的人脸识别系统的研究与实现[J]. 仪器仪表学报, 2016, 37(增刊 1): 162 - 167.

[4] Zhang L L, Lin L, Liang X D, et al. Is faster R - CNN doing well for pedestrian detection? [C]//European Conference on Computer Vision. Cham; Springer, 2016: 443 - 457.

[5] 李建兴, 刘振宇, 马莹, 等. 基于 YOLO v4 的胡瓜钝绥螨自动检测计数[J]. 江苏农业科学, 2022, 50(14): 199 - 207.

[6] Yan Q A, Yang B H, Wang W Y, et al. Apple leaf diseases recognition based on an improved convolutional neural network[J]. Sensors, 2020, 20(12): 3535.

[7] Tao Y, Cuicu L. Recognition system for leaf diseases of ophiopogon japonicus based on PCA - SVM[J]. Plant Diseases and Pests, 2020, 11(2): 9 - 13.

[8] Zeng L X, Irwin K. Introduction to semi - supervised learning[M]. Taylor and Francis; CRC Press, 2013: 1 - 250.

[9] Sajjadi M, Javanmardi M, Tasdizen T. Regularization with stochastic transformations and perturbations for deep semi - supervised learning [C]//Proceedings of the 30th International Conference on Neural Information Processing Systems. December 5 - 10, 2016, Barcelona, Spain. New York; ACM, 2016: 1171 - 1179.

[10] Zhang X, He B, Luo T J. Training query filtering for semi - supervised learning to rank with pseudo labels[J]. World Wide Web, 2016, 19(5): 833 - 864.

[11] Laine S, Aila T. Temporal ensembling for semi - supervised learning [J]. CoRR, 2016, abs/1610.02242.

[12] Sohn K, Berthelot D, Li C L, et al. FixMatch: simplifying semi - supervised learning with consistency and confidence [EB/OL]. 2020; arXiv: 2001.07685. <https://arxiv.org/abs/2001.07685>. pdf.

[13] Zhang Y, Shi B M. Improving pooling method for regularization of convolutional networks based on the failure probability density[J]. Optik, 2017, 145: 258 - 265.

[14] Srivastava N, Hinton G, Krizhevsky A, et al. Dropout: a simple way to prevent neural networks from overfitting[J]. The Journal of Machine Learning Research, 2014, 15(1): 1929 - 1958.

[15] Huang G, Sun Y, Liu Z A, et al. Deep networks with stochastic depth [M]//Computer Vision - ECCV 2016. Cham; Springer International Publishing, 2016: 646 - 661.

[16] 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. IEEE, 2016: 779 - 788.

[17] Ren S Q, He K M, Girshick R, et al. Faster R - CNN: towards real - time object detection with region proposal networks [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2017, 39(6): 1137 - 1149.

[18] Liu W, Anguelov D, Erhan D, et al. SSD: single shot MultiBox detector [EB/OL]. 2015; arXiv: 1512.02325. <https://arxiv.org/abs/1512.02325>. pdf.

[19] Jeong J, Lee S, Kim J, et al. Consistency - based semi - supervised learning for object detection [C]. Proceedings of the 33rd International Conference on Neural Information Processing Systems, 2019, 965: 10759 - 10768.

[20] Bilen H, Vedaldi A. Weakly supervised deep detection networks [C]//2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). June 27 - 30, 2016, Las Vegas, NV, USA. IEEE, 2016: 2846 - 2854.

[21] Zhu Y, Zhang Z Y, Wu C R, et al. Improving semantic segmentation via efficient self - training[J]. arxiv: 2004.14960v2.

[22] Triguero I, García S, Herrera F. Self - labeled techniques for semi - supervised learning: taxonomy, software and empirical study [J]. Knowledge and Information Systems, 2015, 42(2): 245 - 284.

[23] 刘军, 后士浩, 张凯, 等. 基于增强 Tiny YOLO v3 算法的车辆实时检测与跟踪[J]. 农业工程学报, 2019, 35(8): 118 - 125.

[24] Farhadi A, Redmon J. YOLO v3: an incremental improvement [C]//Computer vision and pattern recognition. Berlin/Heidelberg, Germany: Springer, 2018, 1804: 1 - 6.

[25] Wang C Y, Bochkovskiy A, Liao H Y M. YOLO v7: trainable bag - of - freebies sets new state - of - the - art for real - time object detectors [C]//2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). June 17 - 24, 2023, Vancouver, BC, Canada. IEEE, 2023: 7464 - 7475.

[26] Yuan X, Fang S T, Li N, et al. Performance comparison of sea cucumber detection by the YOLO v5 and DETR approach [J]. JMSE, 2023, 11(11): 2043.

[27] Zhang Y N, Kong J, Qi M A, et al. Object detection based on multiple information fusion net [J]. Applied Sciences, 2020, 10(1): 418.

史宝明,贺元香,赵霞.一种轻量级 CNN 玉米病害图像识别方法[J].江苏农业科学,2024,52(5):201-207.  
doi:10.15889/j.issn.1002-1302.2024.05.030

# 一种轻量级 CNN 玉米病害图像识别方法

史宝明<sup>1</sup>,贺元香<sup>1</sup>,赵霞<sup>2</sup>

(1. 兰州文理学院数字媒体学院,甘肃兰州 730010; 2. 甘肃农业大学信息科学技术学院,甘肃兰州 730070)

**摘要:**针对传统卷积神经网络模型参数和运算量超大,难以部署在资源受限的移动终端或嵌入式设备上的问题,以 VGG16 作为基础框架,结合 MobileNet v3 模型思想,提出了一种轻量级卷积神经网络玉米病害图像识别方法,通过逐级渐进的方式建立网络模型,用线性瓶颈的倒残差深度可分离卷积代替标准卷积,用卷积层来代替全连接层,大幅度降低了模型的参数量和运算量。在深度卷积和点卷积之间加入改进的 squeeze and excitation 通道注意力模块,来增强模型精度。注意力模块的第 2 个全连接层的激活函数使用 hard-swish 代替 sigmoid,可以大幅度提高运算速度。试验样本数据为 PlantVillage 数据集的玉米病害子数据集,由于样本数据集偏小,通过随机旋转、随机缩放大小、随机宽度高度偏移、水平翻转、垂直翻转、随机错切变换、随机亮度变化、样本零均值化等方式对样本数据进行了增强和扩充,扩充后的数据集在改进模型上进行试验。试验结果表明,和 VGG16 对比,改进模型的准确率提高了 1.48 百分点,参数量是原模型的 1/5,运算量是原模型的 1/15;在不降低准确率的前提下,模型的参数量和运算量大幅度降低,实现了模型的轻量化。改进模型可以部署在移动终端等手持设备上,为农业病害识别提供指导和参考。

**关键词:**玉米病害;轻量级卷积神经网络;倒残差结构;深度可分离卷积;VGG;病害图像识别

**中图分类号:**TP391.41;S435.131 **文献标志码:**A **文章编号:**1002-1302(2024)05-0201-07

随着深度学习技术的快速发展,相继诞生了一大批优秀的卷积神经网络,如 VGG<sup>[1]</sup>、ReNet<sup>[2]</sup>、InceptionNet<sup>[3]</sup>、DenseNet<sup>[4]</sup>等,模型的准确率不断提升,但同时网络模型的深度越来越深,参数量和运算量越来越大,对硬件设备的算力要求也越来越高。这些网络模型适合在 PC 端进行部署应用,但在移动设备和嵌入式设备中却不适合部署应用。国内学者通过将上述经典卷积神经网络引入农业病害识别中<sup>[5-11]</sup>,取得了不错的效果,但实用性不高。对于农业病害识别来说,开发能够部署在诸如手机移动设备或嵌入式设备的轻量级网络模型<sup>[12]</sup>

具有更好的市场前景。

轻量化卷积神经网络的设计主要从卷积结构轻量化、卷积模块轻量化和卷积运算轻量化 3 个方面来进行<sup>[13]</sup>。在轻量化卷积网络研究和应用方面,2016 年最早公开的轻量化网络 SqueezeNet<sup>[14]</sup>,在 ImageNet 数据集上的表现和 AlexNet 相当,但其参数量仅为 AlexNet 的 2%。后来随着 Xception<sup>[15]</sup>、ShuffleNet<sup>[16-18]</sup>等网络的提出,轻量级神经网络得到了快速发展,并在农业病害防治方面进行了应用和改进<sup>[19-21]</sup>。2017 年 Google 研究团队首次提出 MobileNet<sup>[22]</sup>,使用深度可分离卷积技术,大大降低了模型的参数和运算量,在 MobileNet v1 的基础上,对网络模型的主体模块引入了线性瓶颈的倒残差结构,提出了 MobileNet v2<sup>[23]</sup>。相比之前的模型,它的准确率更高,模型更小。MobileNet v3<sup>[24]</sup>更新了层块的设计,加入了通道注意力机制,使用 H-

收稿日期:2023-04-20

基金项目:甘肃省教育厅创新基金(编号:2023A-181)。

作者简介:史宝明(1981—),男,甘肃平凉人,硕士,副教授,主要研究方向为深度学习、计算机视觉、虚拟现实技术。E-mail:zlsbm@qq.com。

[28] Tan M X, Pang R M, Le Q V. EfficientDet: scalable and efficient object detection [C]//2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). June 13-19, 2020, Seattle, WA, USA. IEEE, 2020:10778-10787.

[29] Li H L, Li J, Wei H B, et al. Slim-neck by GSConv: a better design paradigm of detector architectures for autonomous vehicles [EB/OL]. 2022; arXiv: 2206.02424. [https://arxiv.org/abs/](https://arxiv.org/abs/2206.02424)

2206.02424. pdf.

[30] Fenu G, Mallocci F M. DiaMOS plant: a dataset for diagnosis and monitoring plant disease[J]. Agronomy, 2021, 11(11):2107.

[31] Peng S D, Jiang W, Pi H J, et al. Deep snake for real-time instance segmentation [C]//2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). June 13-19, 2020, Seattle, WA, USA. IEEE, 2020:8530-8539.