

doi:10.6041/j.issn.1000-1298.2023.02.029

# 基于 D2 - YOLO 去模糊识别网络的果园障碍物检测

蔡舒平 潘文浩 刘慧 曾潇 孙仲鸣

(江苏大学电气信息工程学院, 镇江 212013)

**摘要:** 针对果园目标检测时相机抖动以及物体相对运动导致检测图像模糊的问题, 本文提出一种将 DeblurGAN - v2 去模糊网络和 YOLOv5s 目标检测网络相融合的 D2 - YOLO 一阶段去模糊识别深度网络, 用于检测识别果园模糊场景图像中的障碍物。为了减少融合网络的参数量并提升检测速度, 首先将 YOLOv5s 骨干网络中的标准卷积替换成深度可分离卷积, 并且在输出预测端使用 CIOU\_Loss 进行边界框回归预测。融合网络使用改进的 CSPDarknet 作为骨干网络进行特征提取, 将模糊图像恢复原始自然信息后, 结合多尺度特征进行模型预测。为了验证本文方法的有效性, 选取果园中 7 种常见的障碍物作为目标检测对象, 在 Pytorch 深度学习框架上进行模型训练和测试。试验结果表明, 本文提出的 D2 - YOLO 去模糊识别网络准确率和召回率分别为 91.33% 和 89.12%, 与分步式 DeblurGAN - v2 + YOLOv5s 相比提升 1.36、2.7 个百分点, 与 YOLOv5s 相比分别提升 9.54、9.99 个百分点, 能够满足果园机器人障碍物去模糊识别的准确性和实时性要求。

**关键词:** 果园机器人; 障碍物检测; 模糊图像; D2 - YOLO; 融合网络

中图分类号: TP391.4; S24

文献标识码: A

文章编号: 1000-1298(2023)02-0284-09

OSID:



## Orchard Obstacle Detection Based on D2 - YOLO Deblurring Recognition Network

CAI Shuping PAN Wenhao LIU Hui ZENG Xiao SUN Zhongming

(School of Electrical and Information Engineering, Jiangsu University, Zhenjiang 212013, China)

**Abstract:** Aiming at the problem of camera shake and relative motion of objects leading to blurred detection images during target detection in orchards, a D2 - YOLO one-stage deblurring recognition deep network that combined the DeblurGAN - v2 deblurring network and the YOLOv5s target detection network was proposed. It was used to detect and identify obstacles in orchard blurred scene images. To reduce the number of parameters of the fusion model and improve the detection speed, firstly the standard convolution used in the YOLOv5s backbone network with a deep separable convolution was replaced, then CIOU\_Loss was used as the bounding box regression loss function of prediction. The fusion network used the improved CSPDarknet as the backbone for feature extraction. After recovering the original natural information of the blurred image, it combined multi-scale features for model prediction. To verify the effectiveness of the proposed method, seven common obstacles in the real orchard settings were selected as the target detection objects, based on the chassis of the crawler mobile robot, the BUNKER was equipped with portable computers, cameras and other equipment to form a mobile platform for image acquisition, and the model training and testing were carried out on the Pytorch deep learning framework. The precision and recall rates of the proposed D2 - YOLO deblurring detection network were 91.33% and 89.12%, respectively, which were 1.36 percentage points and 2.7 percentage points higher than that of the step-by-step training DeblurGAN - v2 + YOLOv5s. Compared with YOLOv5s, there was an increase of 9.54 percentage points and 9.99 percentage points in precision and recall rates, which can meet the accuracy and real-time requirements of orchard robot obstacle deblurring recognition. The research result can provide a reference for obstacle detection of agricultural robots in orchard in the later stage.

**Key words:** agricultural robot in orchard; obstacle detection; blurred images; D2 - YOLO; fusion network

收稿日期: 2022-04-07 修回日期: 2022-07-18

基金项目: 国家自然科学基金项目(32171908)、江苏省现代农机装备与技术示范推广项目(NJ2021-14)和江苏高校优势学科项目(PAPD)

作者简介: 蔡舒平(1963—),男,副教授,主要从事人工智能算法在农业、电力应用研究, E-mail: spcai@ujs.edu.cn

通信作者: 刘慧(1980—),女,教授,博士生导师,主要从事农业电气化与自动化、智能控制与信号处理研究, E-mail: amity@ujs.edu.cn

## 0 引言

随着人们对果品类的消费需求日益增加,国内外果园种植面积不断扩大<sup>[1]</sup>。传统的果园种植模式耗时、费力、成本高。近些年,由于人工智能技术的高速发展,农业机器人<sup>[2-3]</sup>越来越多地应用于果园种植<sup>[4]</sup>,如修剪、采摘和喷洒农药等。智能农业机器人的自动驾驶技术受到了科研人员的广泛关注<sup>[5-7]</sup>。果园环境是一个复杂的非结构化环境,智能农业机器人在园内道路上作业时,难免会遇到各类障碍物<sup>[8]</sup>,包括人、树、灯杆等,如果不及时发现和规避,将会造成严重的安全问题。为了保证农业机器人在果园自主作业时的安全性和准确性,能够快速准确地检测周围的障碍物非常重要。

果园障碍物检测属于深度学习<sup>[9-10]</sup>中目标检测<sup>[11]</sup>的范畴,目标检测算法主要分为两类:一类是基于 Region Proposal 的目标检测算法,代表模型包括 R-CNN 系列,它们属于两阶段目标检测算法,即让算法先生成目标候选框,对候选框区域的图像进行特征提取,再对提取到的特征进行分类与回归;另一类是使用端到端思想的一阶段目标检测算法,代表模型包括 SSD<sup>[12]</sup> (Single shot multiBox detector)、YOLO<sup>[13-15]</sup> (You only look once) 系列等,此类算法仅使用卷积神经网络 (CNN) 直接提取图像特征后对特征进行分类和检测框坐标的回归。随着一阶段目标检测算法的不断更新,无论是检测速度还是检测精度都比两阶段目标检测算法更具优势。

基于稳定、高质量图像采集的目标检测技术日趋成熟<sup>[16]</sup>。然而,农用机器人在果园作业时,行驶经过的地面凹凸不平,车体容易出现剧烈颠簸,导致摄像头大幅不规则抖动,采集的图像模糊不清,这种模糊严重降低了图像质量,从而对后续基于图像的目标检测、语义分割等高级视觉任务产生负面影响,在实际应用中,会降低农业机器人在作业过程中的

检测精度,导致漏检或误检,作业效率低。

目前,林开颜等<sup>[17]</sup>提出了基于模糊逻辑的植物叶片边缘检测方法,该方法克服了常规算法未考虑叶片像素和背景像素的颜色差异,从而导致产生大量伪边缘的问题,但因为是基于传统的图像去模糊算法,过程较为复杂。马晓丹等<sup>[18]</sup>提出了一种将量子遗传算法的全局搜索能力和模糊推理神经网络的自适应性相结合的算法来识别苹果果实,该算法有效解决了图像边界像素的模糊性和不确定性,但耗时较长。FAN 等<sup>[19]</sup>提出一种基于深度学习的局部模糊图像去模糊方法,用于作物病虫害分类。陈斌等<sup>[20]</sup>提出一种改进 YOLO v3-tiny 的全景图像农田障碍物检测,该算法在 YOLO v3-tiny 的基础上增加一层  $52 \times 52$  的预测层来提升对小目标的检测能力,另外还在主干网络引入了残差模块提高网络的检测能力,但是该算法依旧存在着模型参数量较大且针对输入模糊图像检测能力不足的问题。薛金林等<sup>[21]</sup>提出了一种基于改进 Faster R-CNN 和 SSRN-DeblurNet 的两阶段检测方法,有效解决了农田中由于模糊障碍物产生的问题,但该算法在训练阶段, Faster R-CNN 和 SSRN-DeblurNet 相互独立进行训练,过程较为复杂且耗时。目前针对果园环境下,解决农业机器人在行进间作业时产生运动模糊图像的一阶段去模糊检测算法研究依旧不充分。

为了解决上述问题,本研究提出将改进的 DeblurGAN-v2 去模糊网络和 YOLOv5s 目标检测网络相融合,得到 D2-YOLO 一阶段去模糊识别网络。将模糊图像转换为相对清晰的图像再进行检测,以提高实际目标检测过程精度。

## 1 网络分析与改进

### 1.1 DeblurGAN-v2 去模糊网络

图 1 为 DeblurGAN-v2 的网络结构图,该网络由生成器和鉴别器组成<sup>[22-23]</sup>。生成器接收噪声

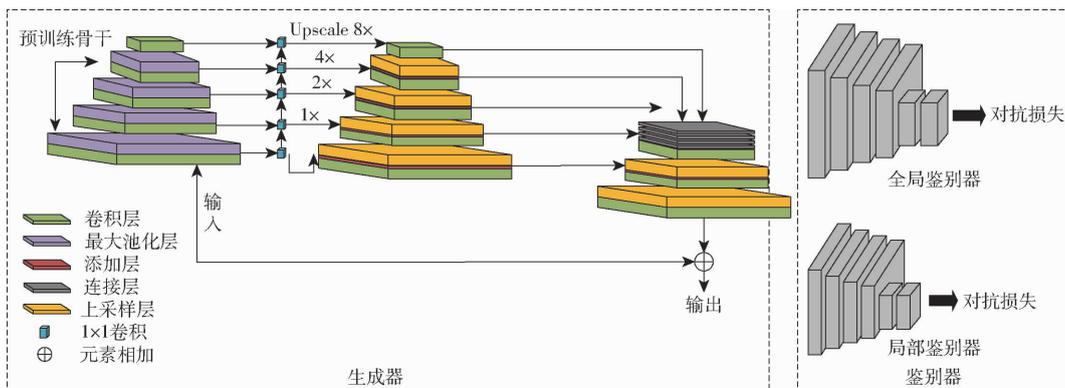


图 1 DeblurGAN-v2 结构图

Fig. 1 DeblurGAN-v2 structure diagram

作为输入并生成样本,鉴别器接收真实样本和生成样本并尝试将两者进行区分。生成器的目标是通过生成无法与真实样本区分开来的感知说服样本来迷惑判别器。两者始终处于最大最小博弈状态,直到判别器无法区分生成样本和真实样本,达到纳什平衡。其中,生成器使用的骨干网络为 Inception - ResNet - v2 或 MobileNet 及其变形。DeblurGAN - v2 首次将特征金字塔网络 (FPN) [24] 的思想引入到图像去模糊中。FPN 可以提高网络的多尺度特征融合能力,实现准确率和速度的平衡,其结构如图 2 所示。

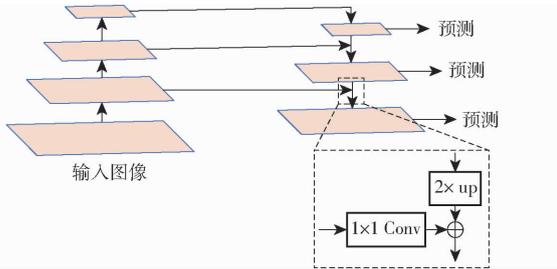


图 2 特征金字塔网络 (FPN) 框架

Fig. 2 Feature pyramid network (FPN) framework

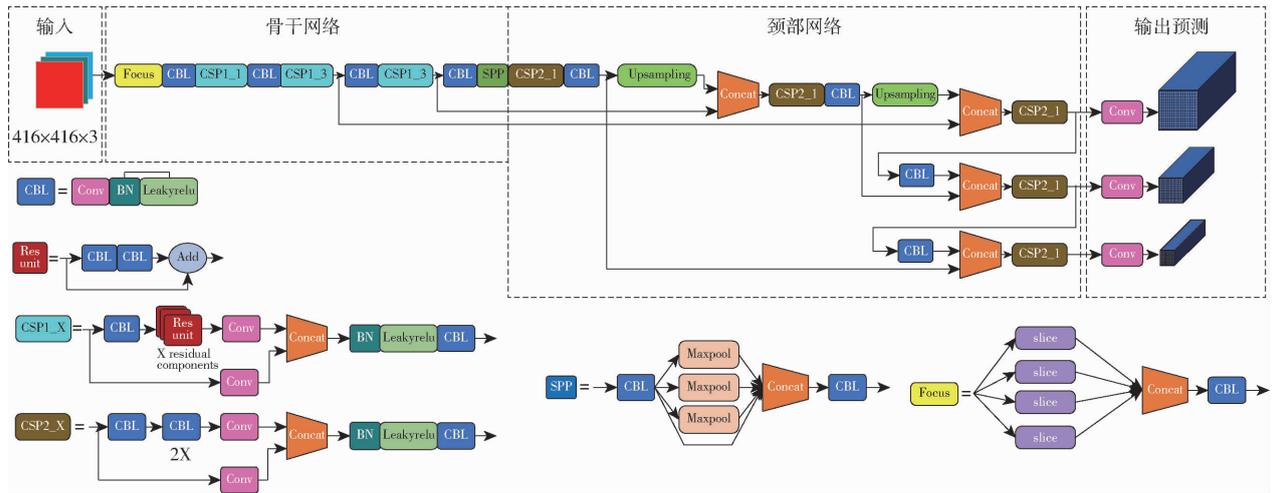


图 3 YOLOv5s 结构图

Fig. 3 YOLOv5s structure diagram

级。但是, CSPDarknet 参数量依旧存在可优化空间,本研究将 CSPDarknet 中的标准卷积替换为深度可分离卷积,并将其用作融合网络。

如图 4 所示,标准卷积过程是对每个通道的输入特征图和对应的卷积核进行卷积运算,然后将它们相加输出特征。计算量  $Q_1$  为

$$Q_1 = D_k^2 M N D_w D_h \quad (1)$$

式中  $D_k$  —— 卷积核大小

$M, N$  —— 输入、输出数据通道数

$D_w, D_h$  —— 输出数据长度和宽度

深度可分离卷积将标准卷积中的一步运算改为  $3 \times 3$  深度卷积和  $1 \times 1$  逐点卷积,计算量  $Q_2$  为

$$Q_2 = (D_k^2 M + M N) D_w D_h \quad (2)$$

因此,深度可分离卷积与标准卷积的计算量比为

### 1.2 YOLOv5s 目标检测网络

YOLOv5s 网络参数量较少、部署成本低,有利于快速部署至农业机器人,适用于中大型目标障碍物检测。YOLOv5s 的网络结构图如图 3 所示,图中, Conv 表示卷积层 (Convolution), CBL 综合模块由卷积层、批量归一化层 (Batch norm) 和激活函数 (Leaky ReLU) 组成, CSP (Cross stage partial) 表示跨阶段部分的结构, Residual unit 表示残差模块, Concat 表示通道数的特征融合。

该网络工作流程为:首先对输入图像进行自适应图像缩放,其次通过骨干网络 (Backbone) 对输入图像进行特征提取,然后经过颈部网络 (Neck) 对这些大小不同的特征图进行分层特征融合,最后在预测层 (Prediction) 上对多个特征图进行同时分类和位置回归。

### 1.3 骨干网络改进

YOLOv5s 使用 CSPDarknet 作为骨干网络,其优点是不仅增强了卷积神经网络的学习能力,还降低了计算瓶颈和内存成本,同时保持了准确性和轻量

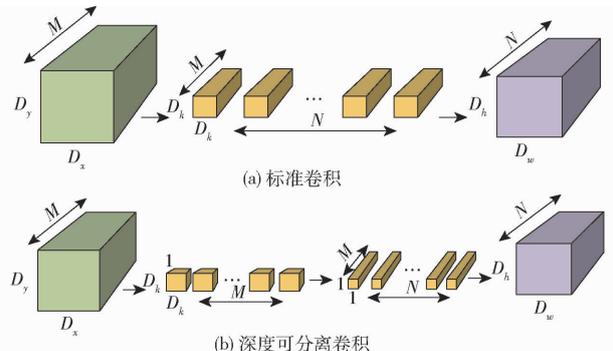


图 4 标准卷积和深度可分离卷积

Fig. 4 Standard convolution and depth separable convolution

$$\frac{Q_2}{Q_1} = \frac{MD_w D_h (D_k^2 + N)}{D_k^2 MND_w D_h} = \frac{1}{N} + \frac{1}{D_k^2} \quad (3)$$

通常  $D_k$  设为 3, 改进后网络计算量和参数实际减少约 1/3。

#### 1.4 损失函数改进

YOLOv5s 的输出预测端使用 GIoU\_Loss 作为边界框回归损失函数, 得到

$$\begin{cases} \text{GIoU\_Loss} = 1 - \text{GIoU} = 1 - \text{IoU} - \frac{A^c - u}{A^c} \\ \text{IoU\_Loss} = 1 - \text{IoU} \end{cases} \quad (4)$$

式中  $A^c$ ——边界框和真实框的最小外接矩阵面积  
 $u$ ——边界框和真实框的并集

当真实框包含边界框时, 两者的最小外接矩阵面积和并集相等, 此时 GIoU 退化为 IoU, 无法区分两者的相对位置。

本文使用 CIoU\_Loss 代替 GIoU\_Loss, CIoU\_Loss 取边界框和真实框的中心, 求两者的欧氏距离, 并引入权重  $\alpha$  来确定边界框和真实框的纵横比, 计算式为

$$\text{CIoU\_Loss} = 1 - \text{CIoU} = 1 - \text{IoU} + \frac{\rho^2}{c^2} + \alpha v \quad (5)$$

其中  $\alpha = \frac{v}{1 - \text{IoU} + v}$

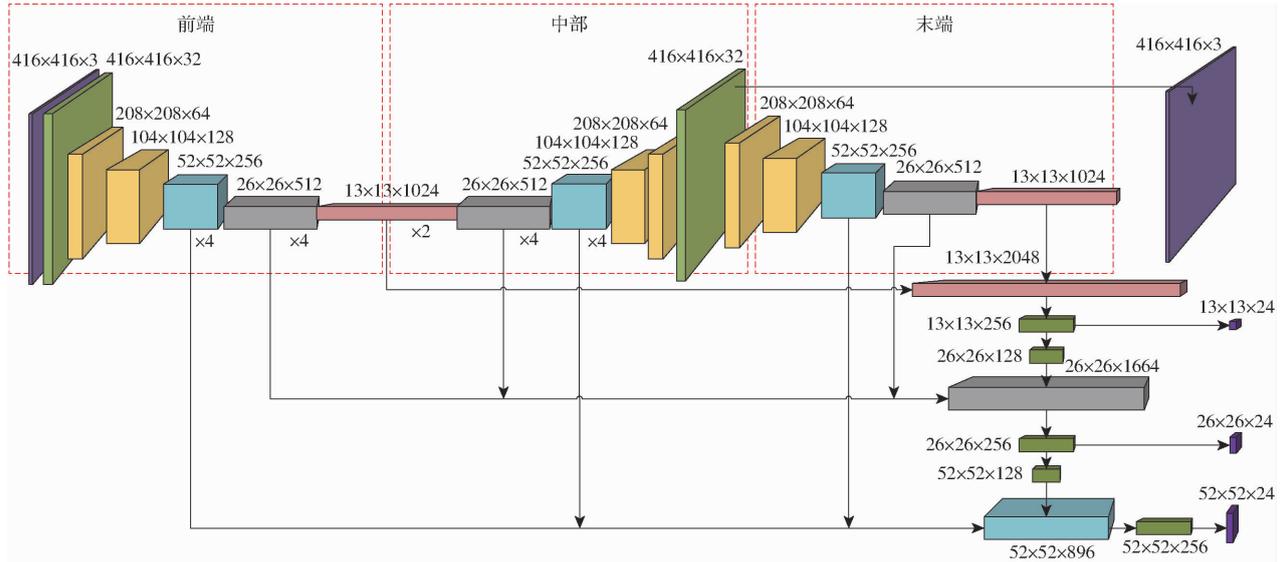


图 5 D2-YOLO 去模糊识别网络结构图

Fig. 5 D2-YOLO deblurring recognition network structure diagram

整个网络分为前端、中部和末端 3 部分。当网络前端接收到输入图像经过自适应图像缩放生成 416 像素 × 416 像素的帧图像后, 通过卷积和多个残差块进行下采样操作, 最终得到 13 像素 × 13 像素的帧图像, 此部分在 YOLOv5s 中将采用多尺度特征融合并进行预测。但在本文中为了恢复图像自然信息, 会在网络的中间阶段进行反卷积及上采样操作, 将图像尺寸重新恢复成 416 像素 × 416 像素并得到

$$v = \frac{4}{\pi^2} \left( \arctan \frac{w^{gt}}{h^{gt}} - \arctan \frac{w}{h} \right)^2 \quad (6)$$

式中  $\rho$ ——边界框与真实框中心点欧氏距离

$c$ ——边界框与真实框的最外接矩阵对角线长度

$w^{gt}, h^{gt}$ ——真实框宽、高

$w, h$ ——边界框宽、高

$v$ ——衡量边界框和真实框之间纵横比一致性参数

采用 CIoU\_Loss 作为边界框回归损失函数后, 一定程度上提高了预测速度和准确率。

#### 1.5 D2-YOLO 去模糊识别网络

为了解决模糊图像给果园障碍物检测带来的影响, 本研究将 DeblurGAN-v2 去模糊网络与 YOLOv5s 目标检测网络相融合。与图像去模糊和目标检测两步走的方法不同, 本研究提出的 D2-YOLO 去模糊检测网络属于一阶段检测网络。D2-YOLO 仅使用改进的 CSPDarknet 作为骨干网络进行特征提取, 采用特征金字塔思想将多个不同的特征图进行层级特征融合, 充分利用去模糊后的清晰图像进行预测, 提高了图像恢复质量和利用率。图 5 为 D2-YOLO 去模糊识别网络结构。

清晰的 RGB 三通道图像, 完成去模糊操作。在网络末端, 首先采用 Focus 切片操作将恢复过后的清晰图像进行处理, 再次通过卷积和多个残差块进行下采样操作, 将 3 个阶段的相同尺寸帧图像进行特征融合, 极大增强了图像自然信息的特征。

#### 1.6 D2-YOLO 训练方法

首先将采集的图像进行运动模糊处理。假设原始图像为  $\sigma(x, y)$ , 模糊图像为  $\varphi(x, y)$ , 则生成图

像为

$$\varphi(x, y) = \int_0^T [\sigma(x + C_{randx}t, y + C_{randy}t) + n(x, y)] dt \quad (7)$$

式中  $C_{randx}$ ——随机方向  $x$  上的投影

$C_{randy}$ ——随机方向  $y$  上的投影

$T$ ——模糊时间  $n$ ——随机噪声

网络分3个阶段进行训练。第1阶段,单独对去模糊过程进行训练,即只用网络的前端和中部,价值函数  $V(G, D)$  为

$$V(G, D) = \int_x p_{data}(x) \lg(D(x)) dx + \int_z p_z(z) \lg(1 - D(g(z))) dz = \int_x [p_{data}(x) \lg(D(x)) + p_g(x) \lg(1 - D(x))] dx \quad (8)$$

式中  $D(x)$ ——取判别器均方根误差

第2阶段,固定网络前端和中部的参数,不再进行去模糊,只训练网络的末端部分,即目标检测阶段。目标检测阶段总损失函数为

$$L_{objdet} = L_{cen} + L_{wh} + L_{cat} + L_{con} \quad (9)$$

其中

$$L_{cen} = \lambda_{coord} \sum_{i=0}^{s^2} \sum_{j=0}^B l_{ij}^{obj} [(x_i - \hat{x}_i)^2 + (y_i - \hat{y}_i)^2] \quad (10)$$

$$L_{wh} = \lambda_{coord} \sum_{i=0}^{s^2} \sum_{j=0}^B l_{ij}^{obj} [(\sqrt{w_i} - \sqrt{\hat{w}_i})^2 + (\sqrt{h_i} - \sqrt{\hat{h}_i})^2] \quad (11)$$

$$L_{cat} = \sum_{i=0}^{s^2} \sum_{j=0}^B l_i^{obj} (p_i(c) - \hat{p}_i(c))^2 \quad (12)$$

$$L_{con} = \sum_{i=0}^{s^2} \sum_{j=0}^B l_{ij}^{obj} (C_i - \hat{C}_i)^2 + \lambda_{noobj} \sum_{i=0}^{s^2} \sum_{j=0}^B l_{ij}^{noobj} (C_i - \hat{C}_i)^2 \quad (13)$$

式中  $L_{cen}$ ——预测中心坐标损失值

$L_{wh}$ ——预测边界框宽和高损失值

$L_{cat}$ ——预测类别损失值

$L_{con}$ ——预测置信度损失值

$\lambda_{coord}$ ——位置损失系数

$\hat{x}_i, \hat{y}_i$ ——目标真实中心坐标

$\hat{w}_i, \hat{h}_i$ ——目标宽度和高度

$p_i(c)$ ——目标类别概率

$\hat{p}_i(c)$ ——类别概率真实值

$C_i$ ——第1个网格中边界框置信度

$\hat{C}_i$ ——预测边界框置信度

如果  $(i, j)$  处的锚框包含目标,则  $l_{ij}^{obj}$  为 1, 否则为 0。

第3阶段,总体训练。在前面2个阶段完成之后,对整个网络再进行训练。总体损失函数为

$$L_{total} = L_{objdet} + D(x) \quad (14)$$

训练好的网络经过图像计算会产生4种结果,分别对应去模糊后  $416 \times 416$  的清晰图像,以及预测阶段为  $13 \times 13, 26 \times 26, 52 \times 52$  的目标检测结果。如果不需要中间去模糊结果,可以对中间图像的输出进行屏蔽,只使用最终的检测结果。

## 2 试验

### 2.1 图像采集

图像数据采集场景由两部分构成,一部分是距离江苏大学5 km处的果园,另一部分是江苏大学校内梅园。在2处场景中共拍摄了2 600幅原始图像,其中清晰图像1 500幅,模糊图像1 100幅,自制模糊图像1 500幅,共计4 100幅。图像由7种果园典型障碍物组成:树木、行人、电线杆、指示牌、支撑架、垃圾桶和围栏,其中,树木种类由桂花树、腊梅树、紫薇树、红枫树、樱花树和桑树等乔木以及金森女贞球等灌木组成。

### 2.2 试验平台

本试验训练处理平台所使用的计算机GPU为NVIDIA RTX 4000, CPU为Intel i9 10900K,内存为64 GB。计算机操作系统为Ubuntu 20.04, 搭载CUDA 11.1.1并行计算框架和CUDNN 8.2.1深度学习加速库。采用Python 3.7编程语言在Pytorch深度学习框架上进行试验。果园移动机器人搭载ZED相机作为果园场景图像采集设备,同时配套软件工具SDK和OpenCV库。

以BUNKER型履带式移动机器人底盘为基础,配备本试验使用的便携式计算机、摄像头等设备,形成图像采集移动平台。图像采集平台和采集场景如图6所示。

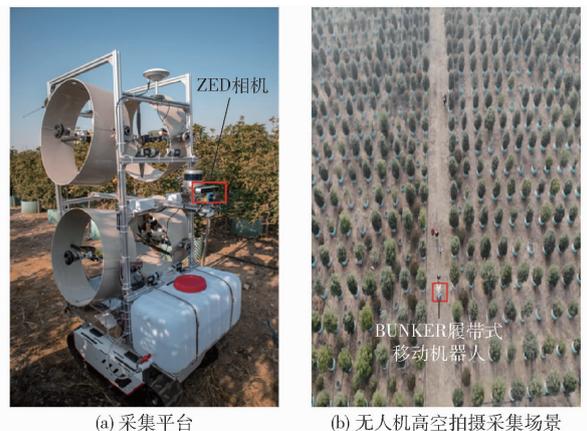


图6 图像采集平台及采集场景

Fig. 6 Experiment platform and data acquisition scene

### 2.3 模型训练

训练模型采用 YOLO 数据集格式,其中训练数据集、验证数据集和测试数据集比例为 8:1:1。使用 LabelImg 软件对 7 类目标的最终训练集进行标注。训练时,采集的图像尺寸先进行 padding 操作,然后尺寸调整为 416 像素  $\times$  416 像素作为输入图像,较小的输入尺寸可以加快网络的计算速度,将图像顺时针旋转  $90^\circ$ 、 $180^\circ$ ,然后裁剪、翻转、随机拉伸、随机扭曲、添加马赛克干扰等方法达到数据增强的目的。训练参数为:批量大小 32、动量 0.97、初始学习率 0.005、衰减系数 0.9。

### 2.4 结果分析

表 1 和图 7 为不同去模糊算法性能和结果对

比,去模糊评估指标包括峰值信噪比 (PSNR) 和结构相似度 (SSIM)<sup>[25]</sup>。

从表 1 可以看出,D2-YOLO 的去模糊用时约

表 1 不同去模糊算法性能对比

Tab.1 Performance comparison of different deblurring algorithms

模型	时间/ s	模型内存 占用量/MB	峰值信 噪比	结构相 似度
SRN-Deblur	0.760	31.5	25.78	0.839
DeblurGAN	0.192	43.5	22.78	0.772
DeblurGAN-v2	0.078	12.8	21.95	0.731
SSRN-DeblurNet	0.270	7.0	26.42	0.750
D2-YOLO	0.084	14.1	23.56	0.794



图 7 不同去模糊算法去模糊结果对比

Fig. 7 Comparison of deblurring results of different deblurring algorithms

为 SRN-Deblur 的 1/10,约为 SSRN-DeblurNet 的 1/3。在去模糊时间相近的情况下,整体性能优于 DeblurGAN 和 DeblurGAN-v2。从图 7 可以看出,与模糊图像相比,D2-YOLO 在一定程度上还原了图像原有的自然信息,使图像更加清晰。在比较模型中,去模糊结果优于 DeblurGAN 和 DeblurGAN-v2,虽与 SRN-Deblur 结果相似,但速度快了将近 10 倍。

本文所提出的 D2-YOLO 网络不仅可以去除图像模糊,而且可以同时障碍物进行识别,为验证 D2-YOLO 去模糊识别网络检测效果,分别与分步训练的 DeblurGAN-v2+YOLOv5s、DeblurGAN-v2+ Faster-RCNN 以及 YOLOv5s 检测网络用同样的数

据和训练参数进行训练,为防止过拟合,在代码中设置 30 Epoch 收敛停止训练。评价指标包括准确率、召回率、F1 值<sup>[26]</sup>。

训练完成后,使用果园障碍物数据集集中的测试集图像测试各个模型的性能,结果如表 2 所示。从表 2 可以看出,D2-YOLO 一阶段去模糊识别网络对果园障碍物进行检测时,准确率比 DeblurGAN-v2+YOLOv5s、DeblurGAN-v2+Faster-RCNN 以及 YOLOv5s 分别提高 1.36、6.82、9.54 个百分点,召回率分别提高 2.7、5.65、9.99 个百分点。同时,D2-YOLO 去模糊识别网络参数量比 DeblurGAN-v2+YOLOv5s 小 29.5%,比 DeblurGAN-v2+Faster-RCNN 小 92.91%。在检测速度方面,D2-YOLO 去

表 2 不同模型检测结果对比

Tab.2 Comparison of detection results among different models

模型	准确 率/%	召回 率/%	F1 值/%	检测速度/ ( $f \cdot s^{-1}$ )	模型内存 占用量/MB	具有去模糊 检测功能	障碍物种类
文献[27]方法	94.50	93.70	94.10	158.00	64.0	否	2
YOLOv5s	81.79	79.13	80.44	50.02	7.2	否	7
DeblurGAN-v2+YOLOv5s	89.97	86.42	88.16	35.73	20.0	是	7
DeblurGAN-v2+Faster-RCNN	84.51	83.47	83.99	15.38	198.8	是	7
D2-YOLO	91.33	89.12	90.21	43.61	14.1	是	7

模糊识别网络比 DeblurGAN - v2 + YOLOv5s 快 22.05%，比 DeblurGAN - v2 + Faster - RCNN 快 183.55%。虽然 D2 - YOLO 在参数量以及检测速度方面逊色于 YOLOv5s，但其准确率、召回率、F1 值均高于 YOLOv5s，在兼顾去模糊功能的前提下，模型的参数量相较于文献[27]方法缩小 77.97%，很大程度上解决了针对模糊图像检测精度较差的问题，表明 D2 - YOLO 去模糊识别网络具有更优秀的性能。图 8 为不同模型平均精度均值(mAP)变化曲线。

为进一步验证本文所提方法的优越性，于真实果园和江苏大学校内共选取 12 处不同场景进行检测试验。将试验分为 2 个类别：

第 1 类，在江苏大学校内选取 3 处场景用以模

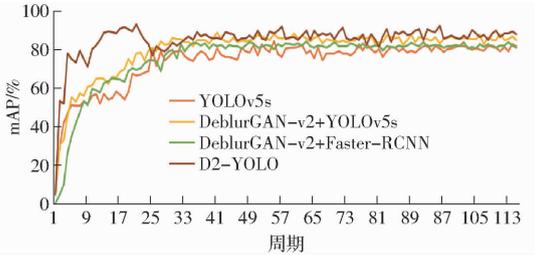


图 8 不同模型平均精度均值(mAP)变化曲线

Fig.8 Changing curves of mean average precision

拟真实果园场景检测，如图 9a ~ 9c 所示。从图中可以看出，YOLOv5s 在接收到模糊图像后，对被遮挡的行人和部分树木以及电线杆出现漏检和误检现象，但使用 D2 - YOLO 进行去模糊后的目标检测，有效解决了这一问题。



图 9 不同场景下不同模型的检测结果对比

Fig.9 Comparison of detection results of different models in different scenes

第 2 类，真实果园中选取 9 处不同场景进行试验，其中 3 处场景中的树种为金森女贞球等灌木，如图 9d ~ 9f 所示，6 处场景中的树种为桂花树、腊梅树、紫薇树、红枫树、桑树等乔木，如图 9g ~ 9l 所示，可以看出，无论是乔木还是灌木，YOLOv5s 在接收到模糊图像后，对支撑杆和部分树木均出现误检和漏检现象，对围栏位置检测也不够准确，但是，同样使用 D2 - YOLO 进行去模糊后的目标检测，极大程度上解决了输入模糊图像造成的漏检和误检问题，与模拟场景下的检测结果一致，充分验证了本文所提方法的优越性。

### 3 结论

(1)提出了一种 D2 - YOLO 一阶段去模糊识别

网络。为了减少融合网络的参数量并提高检测速度，首先将 YOLOv5s 骨干中使用的标准卷积替换为深度可分离卷积，并使用 CIoU\_Loss 作为边界框回归损失函数。融合网络使用改进的 CSPDarknet 进行特征提取，将模糊图像恢复图像原始自然信息后，结合多尺度特征进行模型预测。

(2)将 D2 - YOLO 一阶段去模糊识别网络与 DeblurGAN - v2 + YOLOv5s、DeblurGAN - v2 + Faster - RCNN、YOLOv5s 模型进行对比试验，结果表明 D2 - YOLO 具有更好的检测性能，准确率和召回率分别达到 91.33% 和 89.12%，且网络参数量更少，极大程度上解决了相机抖动和物体相对运动等原因所导致的图像模糊给障碍物检测带来干扰的问题，能够满足果园环境下作业要求。

## 参 考 文 献

- [1] 秦喜田,刘学峰,任冬梅,等.我国果园生产机械化现状及其发展趋势[J].农业装备与车辆工程,2019,57(增刊1):35-38.  
QIN Xitian, LIU Xuefeng, REN Dongmei, et al. Current situation and development prospect of orchard mechanization in China [J]. Agricultural Equipment & Vehicle Engineering, 2019, 57(Supp. 1): 35-38. (in Chinese)
- [2] JIN Yucheng, LIU Jizhan, XU Zhujie, et al. Development status and trend of agricultural robot technology[J]. International Journal of Agricultural and Biological Engineering, 2021, 14(4): 1-19.
- [3] 毛文菊,刘恒,王小乐,等.双导航模式果园运输机器人设计与试验[J].农业机械学报,2022,53(3):27-39,49.  
MAO Wenju, LIU Heng, WANG Xiaole, et al. Design and experiment of dual navigation mode orchard transport robot[J]. Transactions of the Chinese Society for Agricultural Machinery, 2022, 53(3): 27-39, 49. (in Chinese)
- [4] 周俊,胡晨.密植果园作业机器人行间定位方法[J].农业机械学报,2015,46(11):22-28.  
ZHOU Jun, HU Chen. Inter-row localization method for agricultural robot working in close planting orchard[J]. Transactions of the Chinese Society for Agricultural Machinery, 2015, 46(11): 22-28. (in Chinese)
- [5] 毕松,王宇豪.果园机器人视觉导航行间姿态估计与果树目标定位方法[J].农业机械学报,2021,52(8):16-26,39.  
BI Song, WANG Yuhao. Inter-line pose estimation and fruit tree location method for orchard robot[J]. Transactions of the Chinese Society for Agricultural Machinery, 2021, 52(8): 16-26, 39. (in Chinese)
- [6] 李秋洁,丁旭东,邓贤.基于激光雷达的果园行间路径提取与导航[J].农业机械学报,2020,51(增刊2):344-350.  
LI Qiujie, DING Xudong, DENG Xian. Intra-row path extraction and navigation for orchards based on LiDAR[J]. Transactions of the Chinese Society for Agricultural Machinery, 2020, 51(Supp. 2): 344-350. (in Chinese)
- [7] 熊斌,张俊雄,曲峰,等.基于BDS的果园施药机自动导航控制系统[J].农业机械学报,2017,48(2):45-50.  
XIONG Bin, ZHANG Junxiong, QU Feng, et al. Navigation control system for orchard spraying machine based on Beidou navigation satellite system[J]. Transactions of the Chinese Society for Agricultural Machinery, 2017, 48(2): 45-50. (in Chinese)
- [8] LI Y, LI M, QI J, et al. Detection of typical obstacles in orchards based on deep convolutional neural network[J]. Computers and Electronics in Agriculture, 2021, 181: 105932.
- [9] 谢忠红,徐焕良,黄秋桂,等.基于高光谱图像和深度学习的菠菜新鲜度检测[J].农业工程学报,2019,35(13):277-284.  
XIE Zhonghong, XU Huanliang, HUANG Qiugui, et al. Spinach freshness detection based on hyperspectral image and deep learning method[J]. Transactions of the CSAE, 2019, 35(13): 277-284. (in Chinese)
- [10] 宁政通,罗陆锋,廖嘉欣,等.基于深度学习的葡萄果梗识别与最优采摘定位[J].农业工程学报,2021,37(9):222-229.  
NING Zhengtong, LOU Lufeng, LIAO Jiixin, et al. Recognition and the optimal picking point location of grape stems based on deep learning[J]. Transactions of the CSAE, 2021, 37(9): 222-229. (in Chinese)
- [11] XU W, ZHAO L, LI J, et al. Detection and classification of tea buds based on deep learning[J]. Computers and Electronics in Agriculture, 2022, 192: 106547.
- [12] 刘慧,张礼帅,沈跃,等.基于改进SSD的果园行人实时检测方法[J].农业机械学报,2019,50(4):29-35,101.  
LIU Hui, ZHANG Lishuai, SHEN Yue, et al. Real-time pedestrian detection in orchard based on improved SSD[J]. Transactions of the Chinese Society for Agricultural Machinery, 2019, 50(4): 29-35, 101. (in Chinese)
- [13] 景亮,王瑞,刘慧,等.基于双目相机与改进YOLOv3算法的果园行人检测与定位[J].农业机械学报,2020,51(9):34-39,25.  
JING Liang, WANG Rui, LIU Hui, et al. Orchard pedestrian detection and location based on binocular camera and improved YOLOv3 algorithm[J]. Transactions of the Chinese Society for Agricultural Machinery, 2020, 51(9): 34-39, 25. (in Chinese)
- [14] 蔡舒平,孙仲鸣,刘慧,等.基于改进型YOLOv4的果园障碍物实时检测方法[J].农业工程学报,2021,37(2):36-43.  
CAI Shuping, SUN Zhongming, LIU Hui, et al. Real-time detection methodology for obstacles in orchards using improved YOLOv4[J]. Transactions of the CSAE, 2021, 37(2): 36-43. (in Chinese)
- [15] 胡根生,吴继甜,鲍文霞,等.基于改进YOLOv5网络的复杂背景图像中茶尺蠖检测[J].农业工程学报,2021,37(21):191-198.  
HU Gensheng, WU Jitian, BAO Wenxia, et al. Detection of ectropis oblique in complex background images using improved YOLOv5[J]. Transactions of the CSAE, 2021, 37(21): 191-198. (in Chinese)
- [16] 王昱潭,薛君蕊.改进SSD的灵武长枣图像轻量化目标检测方法[J].农业工程学报,2021,37(19):173-182.  
WANG Yutan, XUE Junrui. Lightweight object detection method for Lingwu long jujube images based on improved SSD[J]. Transactions of the CSAE, 2021, 37(19): 173-182. (in Chinese)
- [17] 林开颜,司慧萍,周强,等.基于模糊逻辑的植物叶片边缘检测方法[J].农业机械学报,2013,44(6):227-231.  
LIN Kaiyan, SI Huiping, ZHOU Qiang, et al. Plant leaf edge detection based on fuzzy logic[J]. Transactions of the Chinese Society for Agricultural Machinery, 2013, 44(6): 227-231. (in Chinese)
- [18] 马晓丹,刘刚,周薇,等.基于量子遗传模糊神经网络的苹果果实识别[J].农业机械学报,2013,44(12):227-232,251.  
MA Xiaodan, LIU Gang, ZHOU Wei, et al. Apple recognition based on fuzzy neural network and quantum genetic algorithm [J]. Transactions of the Chinese Society for Agricultural Machinery, 2013, 44(12): 227-232, 251. (in Chinese)
- [19] FAN T, XU J. Image classification of crop diseases and pests based on deep learning and fuzzy system[J]. International Journal of Data Warehousing and Mining, 2020, 16(2): 34-47.
- [20] 陈斌,张漫,徐弘祯,等.基于改进YOLO v3-tiny的全景图像农田障碍物检测[J].农业机械学报,2021,52(增刊):58-65.  
CHEN Bin, ZHANG Man, XU Hongzhen, et al. Farmland obstacle detection in panoramic image based on improved YOLO v3-tiny[J]. Transactions of the Chinese Society for Agricultural Machinery, 2021, 52(Supp.): 58-65. (in Chinese)
- [21] 薛金林,李雨晴,曹梓建.基于深度学习的模糊农田图像中障碍物检测[J].农业机械学报,2022,53(3):234-242.

- XUE Jinlin, LI Yuqing, CAO Zijian. Obstacle detection based on deep learning for blurred farmland images[J]. Transactions of the Chinese Society for Agricultural Machinery, 2022, 53(3): 234–242. (in Chinese)
- [22] KUPYN O, MARTYNIUK T, WU J, et al. Deblurgan - v2: deblurring (orders-of-magnitude) faster and better [C] // Proceedings of the IEEE/CVF International Conference on Computer Vision, 2019: 8878–8887.
- [23] KUPYN O, BUDZAN V, MYKHAILYCH M, et al. Deblurgan: blind motion deblurring using conditional adversarial networks [C] // Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2018: 8183–8192.
- [24] ZHAO Y, HAN R, RAO Y. A new feature pyramid network for object detection [C] // 2019 International Conference on Virtual Reality and Intelligent Systems (ICVRIS). IEEE, 2019: 428–431.
- [25] HORE A, ZIOU D. Image quality metrics: PSNR vs. SSIM [C] // 2010 20th International Conference on Pattern Recognition. IEEE, 2010: 2366–2369.
- [26] 沈跃, 庄珍珍, 刘慧, 等. 基于 RealSense 深度相机的多特征树干快速识别方法[J]. 农业机械学报, 2022, 53(4): 304–312. SHEN Yue, ZHUANG Zhenzhen, LIU Hui, et al. Fast recognition method of multi-feature trunk based on RealSense depth camera [J]. Transactions of the Chinese Society for Agricultural Machinery, 2022, 53(4): 304–312. (in Chinese)
- [27] ZHENG S, WU Y, JIANG S, et al. Deblur-YOLO: real-time object detection with efficient blind motion deblurring [C] // 2021 International Joint Conference on Neural Networks (IJCNN). IEEE, 2021: 1–8.

(上接第 255 页)

- [14] 李卓, 杜晓冬, 毛涛涛, 等. 基于深度图像的猪体尺检测系统[J]. 农业机械学报, 2016, 47(3): 311–318. LI Zhuo, DU Xiaodong, MAO Taotao, et al. Pig dimension detection system based on depth image [J]. Transactions of the Chinese Society for Agricultural Machinery, 2016, 47(3): 311–318. (in Chinese)
- [15] PEZZUOLO A, GUARINO M, SARTORI L, et al. On-barn pig weight estimation based on body measurements by a Kinect v1 depth camera [J]. Computers and Electronics in Agriculture, 2018, 148: 29–36.
- [16] 司永胜, 安露露, 刘刚, 等. 基于 Kinect 相机的猪体理想姿态检测与体尺测量[J]. 农业机械学报, 2019, 50(1): 58–65. SI Yongsheng, AN Lulu, LIU Gang, et al. Ideal posture detection and body size measurement of pig based on Kinect [J]. Transactions of the Chinese Society for Agricultural Machinery, 2019, 50(1): 58–65. (in Chinese)
- [17] ZHANG J, ZHUANG Y, JI H, et al. Pig weight and body size estimation using a multiple output regression convolutional neural network: a fast and fully automatic method [J]. Sensors, 2021, 21(9): 3218.
- [18] TOSHEV A, SZEGEDY C. DeepPose: human pose estimation via deep neural networks [C] // Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2014: 1653–1660.
- [19] INSAFUTDINOV E, PISHCHULIN L, ANDRES B, et al. DeeperCut: a deeper, stronger, and faster multi-person pose estimation model [C] // European Conference on Computer Vision, 2016: 34–50.
- [20] INSAFUTDINOV E, ANDRILUKA M, PISHCHULIN L, et al. ArtTrack: articulated multi-person tracking in the wild [C] // Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2017: 6457–6465.
- [21] CAO Z, HIDALGO G, SIMOM T, et al. OpenPose: realtime multi-person 2D pose estimation using part affinity fields [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2019, 43(1): 172–186.
- [22] MATHIS A, MAMIDANNA P, CURY K M, et al. DeepLabCut: markerless pose estimation of user-defined body parts with deep learning [J]. Nature Neuroscience, 2018, 21(9): 1281–1289.
- [23] BOVA A, KERNODLE K, MULLIGAN K, et al. Automated rat single-pellet reaching with 3-Dimensional reconstruction of paw and digit trajectories [J]. Journal of Visualized Experiments, 2019(149): e59979.
- [24] LABUGUEN R, BARDELOZA D K, NEGRETE S B, et al. Primate markerless pose estimation and movement analysis using DeepLabCut [C] // 2019 Joint 8th International Conference on Informatics, Electronics & Vision (ICIEV) and 2019 3rd International Conference on Imaging, Vision & Pattern Recognition (icIVPR). IEEE, 2019: 297–300.
- [25] NATH T, MATHIS A, CHEN A C, et al. Using DeepLabCut for 3D markerless pose estimation across species and behaviors [J]. Nature Protocols, 2019, 14(7): 2152–2176.
- [26] MATHIS A, BIASI T, SCHNEIDER S, et al. Pretraining boosts out-of-domain robustness for pose estimation [C] // Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision, 2021: 1859–1868.
- [27] WU Z, SHEN C, VAN DEN HENGEL A. Wider or deeper: revisiting the resnet model for visual recognition [J]. Pattern Recognition, 2019, 90: 119–133.
- [28] HE K, ZHANG X, REN S, et al. Deep residual learning for image recognition [C] // Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2016: 770–778.
- [29] 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.
- [30] TAN M, LE Q. EfficientNet: rethinking model scaling for convolutional neural networks [C] // International Conference on Machine Learning, 2019: 6105–6114.