

基于 MSRCR – YOLOv4 – tiny 的田间玉米杂草检测模型

刘莫尘^{1,2} 高甜甜¹ 马宗旭¹ 宋占华^{1,3} 李法德^{1,3} 闫银发^{1,2}

(1. 山东农业大学机械与电子工程学院, 泰安 271018; 2. 山东省农业装备智能化工程实验室, 泰安 271018;
3. 山东省园艺机械与装备重点实验室, 泰安 271018)

摘要: 为实现田间环境下对玉米苗和杂草的高精度实时检测, 本文提出一种融合带色彩恢复的多尺度视网膜 (Multi-scale retinex with color restoration, MSRCR) 增强算法的改进 YOLOv4 – tiny 模型。首先, 针对田间环境的图像特点采用 MSRCR 算法进行图像特征增强预处理, 提高图像的对比度和细节质量; 然后使用 Mosaic 在线数据增强方式, 丰富目标检测背景, 提高训练效率和小目标的检测精度; 最后对 YOLOv4 – tiny 模型使用 K-means + + 聚类算法进行先验框聚类分析和通道剪枝处理。改进和简化后的模型总参数量降低了45.3%, 模型占用内存减少了45.8%, 平均精度均值 (Mean average precision, mAP) 提高了2.5个百分点, 在 Jetson Nano 嵌入式平台上平均检测帧耗时减少了22.4%。本文提出的 Prune – YOLOv4 – tiny 模型与 Faster RCNN、YOLOv3 – tiny、YOLOv4 3 种常用的目标检测模型进行比较, 结果表明: Prune – YOLOv4 – tiny 的 mAP 为 96.6%, 分别比 Faster RCNN 和 YOLOv3 – tiny 高 22.1 个百分点和 3.6 个百分点, 比 YOLOv4 低 1.2 个百分点; 模型占用内存为 12.2 MB, 是 Faster RCNN 的 3.4%, YOLOv3 – tiny 的 36.9%, YOLOv4 的 5%; 在 Jetson Nano 嵌入式平台上平均检测帧耗时为 131 ms, 分别是 YOLOv3 – tiny 和 YOLOv4 模型的 32.1% 和 7.6%。可知本文提出的优化方法在模型占用内存、检测耗时和检测精度等方面优于其他常用目标检测算法, 能够为硬件资源有限的田间精准除草的系统提供可行的实时杂草识别方法。

关键词: 杂草识别; YOLOv4 – tiny; 带色彩恢复的多尺度视网膜增强算法; 模型剪枝; 嵌入式设备
中图分类号: S887.33⁺1 **文献标识码:** A **文章编号:** 1000-1298(2022)02-0246-10 **OSID:** 

Target Detection Model of Corn Weeds in Field Environment Based on MSRCR Algorithm and YOLOv4 – tiny

LIU Mochen^{1,2} GAO Tiantian¹ MA Zongxu¹ SONG Zhanhua^{1,3} LI Fade^{1,3} YAN Yinfa^{1,2}

(1. College of Mechanical and Electronic Engineering, Shandong Agricultural University, Taian 271018, China
2. Shandong Engineering Laboratory of Agricultural Equipment Intelligence, Taian 271018, China
3. Shandong Provincial Key Laboratory of Horticultural Machinery and Equipment, Taian 271018, China)

Abstract: To solve the problem of low accuracy and poor real-time performance of weed recognition in corn field, a detection method of weed based on multi-scale retinex with color restoration (MSRCR) and improved YOLOv4 – tiny algorithm was proposed. Firstly, according to the image characteristics of weed in corn field environment, the MSRCR algorithm was used for image feature enhancement preprocessing to improve the image contrast and detail quality. Then, Mosaic online data augmentation method was used to enrich the object detection background, improve the training efficiency and the detection accuracy of small objects. Finally, The K-means + + was used for a priori anchor boxes clustering analysis and channel pruning for the YOLOv4 – tiny model. The total parameters of the improved and simplified model were reduced by 45.3%, the model size was reduced by 45.8%, the mean average precision (mAP) was increased by 2.5 percentage points, and the average detection frame time on the Jetson Nano embedded platform was reduced by 22.4%. The proposed Prune – YOLOv4 – tiny model was compared with Faster

收稿日期: 2021 – 11 – 04 修回日期: 2021 – 12 – 13

基金项目: 国家自然科学基金青年科学基金项目 (32001419)、财政部和农业农村部: 国家现代农业产业技术体系项目 (CARS – 18 – ZJ0402) 和山东省现代农业产业技术体系建设项目 (SDAIT – 18 – 06)

作者简介: 刘莫尘 (1980—), 女, 副教授, 博士, 主要从事现代农业机械装备研究, E-mail: liu_mochen@126.com

通信作者: 闫银发 (1976—), 男, 教授, 博士, 主要从事机器视觉在现代农业中的应用研究, E-mail: sd28@163.com

RCNN, YOLOv3 - tiny, and YOLOv4, the experimental results showed that the mAP of the Prune - YOLOv4 - tiny model was 96.6%, which was 22.1 percentage points and 3.6 percentage points higher than that of the Faster RCNN and YOLOv3 - tiny, and 1.2 percentage points lower than that of the YOLOv4 model; the model size of the Prune - YOLOv4 - tiny was 12.2 MB, which was 3.4% of the Faster RCNN, 36.9% of the YOLOv3 - tiny, and 5% of the YOLOv4; the average detection frame time on the Jetson Nano embedded platform was 131 ms, which was 32.1% of the YOLOv3 - tiny and 7.6% of YOLOv4. The optimization method proposed was superior to other commonly used object detection algorithms in model size, detection time and detection accuracy, which could provide a feasible real-time weed recognition method for the field precision weeding system with limited hardware resources.

Key words: weed recognition; YOLOv4 - tiny model; multi-scale retinex with color restoration (MSRCR) enhancement algorithm; model pruning; embedded device

0 引言

农田杂草是农业生产中最严重的威胁之一,杂草与农作物争夺阳光、水分和养料,加剧了土地退化,严重影响农作物的产量和品质^[1]。传统的人工除草方法劳动强度大、费时费力且除草效率低,已经不适应现代农业的发展要求^[2]。目前主要的除草方式是大量无差别喷洒除草剂,这种方式不但会造成大量除草剂的浪费,还会伤害目标作物或下茬轮作物^[3]。因此,在玉米苗期进行杂草抑制,开发杂草的精准管理系统非常必要,首先需要正确检测和识别杂草^[4]。

在作物杂草识别的研究中,机器视觉技术已经得到广泛应用,通常包括特征提取与模式识别 2 个步骤^[5]。文献[6 - 9]均是先提取特征再通过分类器对杂草进行识别,虽然识别精度较高,但分类耗时过长,泛化性差,难以实际应用。

随着机器学习的发展,深度学习^[10 - 13]被广泛应用在目标检测领域。基于深度学习的目标检测算法主要分为两类:一类是两阶段算法,通过卷积神经网络生成一系列可能存在潜在目标的候选区域,然后根据候选区域的特征对每一个区域进行分类和边界回归^[14],典型的算法有 R - CNN^[15]、Faster RCNN^[16]、Mask R - CNN^[17]、SPP - Net^[18]等;另一类是单阶段算法,省略了候选区域生成的步骤,且不再采用滑动窗口来寻找目标,仅将目标检测设计为一个回归问题,使用单个卷积神经网络直接完成输入的整幅图像上所有目标的定位与分类,典型的算法有 YOLO^[19]、SSD^[20]、Retina - Net^[21]、CenterNet^[22]等。文献[23 - 27]均是在带有高性能图形处理器(Graphics processing unit, GPU)的个人计算机(Personal computer, PC)上进行测试,而在实际工作环境下除草机器人大多采用嵌入式设备,如何在有限硬件配置下节省计算资源,达到实时精确的识别效果是需要解决的一个关键问题。

为解决上述问题,本文以田间环境下玉米苗和杂草为研究对象,将带色彩恢复的多尺度视网膜(Multi-scale retinex with color restoration, MSRCR)增强算法应用到数据集上,采用基于改进 YOLOv4 - tiny 的玉米杂草识别方法,通过模型剪枝减少模型参数量与计算量,在保证识别精度的前提下有效提升检测速度,大幅降低模型占用空间,为机器学习在硬件资源有限的嵌入式设备上应用提供理论依据和方法参考。

1 图像获取及预处理

1.1 图像获取

本研究以 2 ~ 5 叶期玉米及其早期伴生杂草作为研究对象进行识别试验,于 2021 年 5 月进行图像采集,分别选择晴天、多云和阴天(时间段为 08:00—12:00、14:00—17:30)在山东农业大学泮河校区玉米试验田中完成,采集设备为 Raspberry Camera V2 型摄像头,800 万像素,CCD 传感器,最大分辨率为 3 280 像素 × 2 464 像素。

本文建立的数据集均从玉米苗期田间环境中采集,距地面 40 cm 处垂直拍摄,共采集图像 800 幅,为了后续应用到嵌入式平台减少计算成本,存为 608 像素 × 608 像素的低像素图像。本文数据集图像具有以下 1 种或多种特点:玉米和杂草互相遮挡(图 1a)、玉米和杂草叶片的颜色和纹理具有相似性(图 1b)、自然光线下玉米和杂草阴影(图 1c)、田间其他物体干扰(图 1d)等复杂背景,与实际应用环境相同,部分图像样本如图 1 所示。

1.2 图像数据增强

为了增强试验数据集的丰富性,使用图像数据增强技术来扩充数据集的规模,降低杂草识别模型对某些图像属性的依赖,减少训练模型过拟合、加强模型的稳定性^[28]。本文随机进行逆时针 90°、180°、270°旋转,水平翻转,以及添加适量的高斯噪声、椒

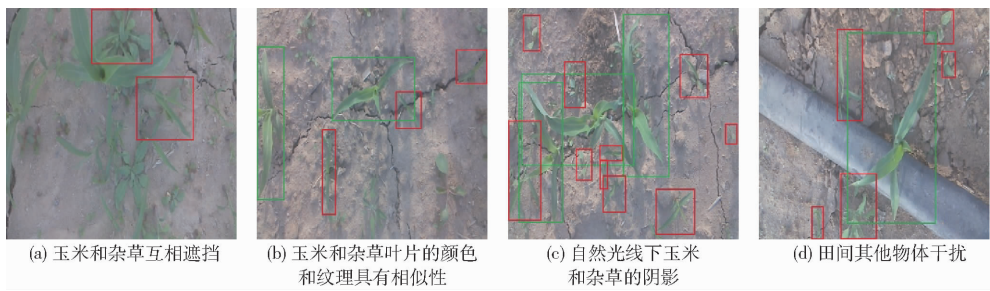


图 1 玉米田间杂草示例图像(绿色框代表玉米,红色框代表杂草)

Fig.1 Example images of weeds in corn filed (green boxes indicated crons and red boxes indicated weeds)

盐噪声和散点噪声等常见方式进行数据增强。使用上述方法将数据集增加到 9 600 幅,随机划分数据

集中的 80% 用于训练,20% 用于验证。数据增强后的示例样本图像如图 2 所示。

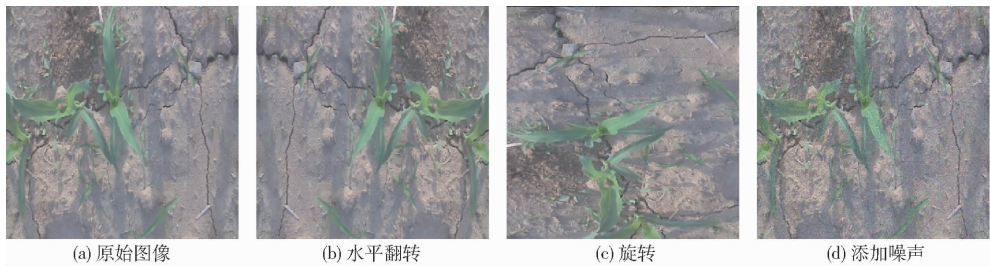


图 2 示例图像

Fig.2 Example images

1.3 图像特征增强预处理

采集的图像不仅具有 1.1 节中提到的复杂特征,在实际工作环境下采集的图像还会由于振动、光线、天气等原因造成图像的细节模糊、对比度差等问题。图像增强技术能够有效提高原始图像中杂草和作物的特征清晰度,有利于增加后续网络的检测精度,没有使检测的速度明显降低^[29]。

分别采取限制对比度自适应直方图均衡化(Contrast limited adaptive histogram equalization, CLAHE)算法^[30]、同态滤波算法、MSRCR^[31]算法进行图像特征增强。CLAHE 算法通过对局部直方图的高度进行限制来提升图像对比度,增强有效信号和抑制噪声信号,并利用线性插补算法提高速度。同态滤波算法把频率滤波和空域灰度变换结合起来,将图像的照度/反射率模型作为频域处理的基础,利用压缩亮度范围和增强对比度来改善图像的品质。MSRCR 算法利用彩色恢复因子调节图像中 3 个颜色通道之间的比例关系,进而凸显相对较暗区域的信息,消除图像色彩失真。

各算法图像特征增强效果如图 3 所示。由图 3b 可以看出,经过 CLAHE 处理过的图像,对比度有明显提升,但图像色偏问题严重并且图像偏暗;由图 3c 可以看出,经过同态滤波处理后的图像中绿叶特征变得突出,但整体上呈红绿色调且亮度较低;由图 3d 可以看出,经过 MSRCR(取尺度数为 3,Dynamic 为 2.5)处理后的图像叶片与背景区分明

显,图像边缘纹理等细节更加清晰,整体亮度提升,饱和度较高。

1.4 图像标注

本文使用图像可视化标注工具 LabelImg 对数据集进行人工标注玉米及杂草标签,以目标的最小外接矩形进行标注,标注为 xml 标准格式并以 PASCAL VOC 格式进行存储,存储信息包括:图像的路径、宽高尺寸、通道数以及玉米、杂草标注框的位置信息。

2 基于 YOLOv4 - tiny 模型的玉米杂草识别网络

2.1 YOLOv4 - tiny 模型

YOLOv4 - tiny^[32]是目前主流检测网络中的轻量化模型,更适合部署在硬件资源有限的移动端或嵌入式设备上实时作业。其网络结构如图 4 所示,该模型共 38 层,输入为 608 像素 × 608 像素 × 3 通道的玉米田间苗草图像,Backbone 模块是网络的骨干部分,采用的特征提取网络是 CSPDarknet53 - tiny。CSP 结构将基础层的特征映射划分为两部分,通过跨层连接将它们合并。CBL 结构主要对图像进行下采样处理,由卷积层(Convolutional)、批量标准化层(BN)以及激活函数层(Leaky ReLU)构成。输出层对 YOLO 的 2 个头部的处理,同时得到玉米和杂草的置信度及位置信息。

YOLOv4 - tiny 进行目标检测的基本思路如图 5

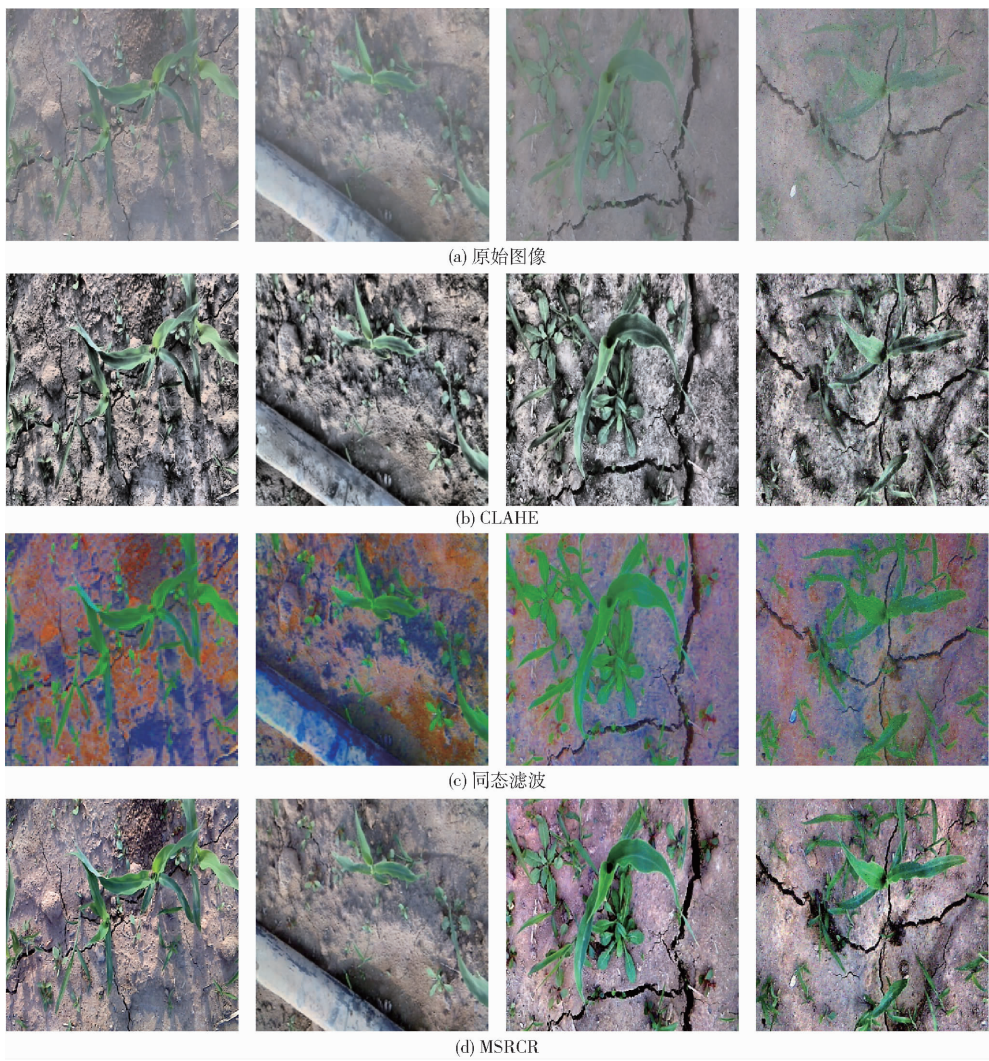


图 3 不同图像特征增强算法的效果对比

Fig. 3 Effect comparison of different image feature enhancement algorithm

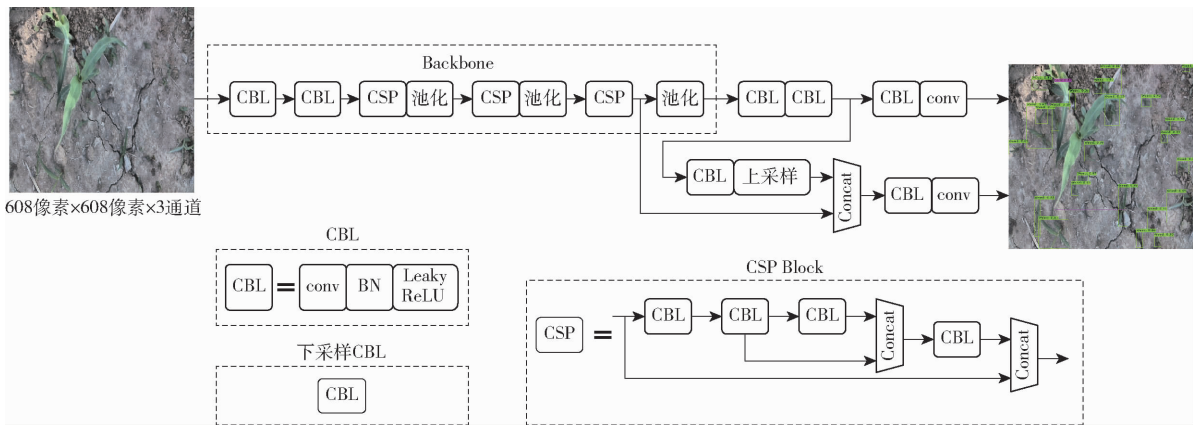


图 4 YOLOv4 - tiny 网络结构图

Fig. 4 YOLOv4 - tiny network structure diagram

所示。YOLOv4 - tiny 算法会对输入的图像划分候选框,将 1 幅图像分为 $S \times S$ 个小单元格,在每个单元格中会预测 B 个边界框和 D 类物体,预测边界框主要包含类别名称、识别目标在边界框的中心坐标 (x,y) 、边界框的宽度和高度 (w,h) 以及置信度。本文选取置信阈值为 0.5,如果边界框的置信度得分

高于 0.5,则边界框保持;否则删除边界框。边界框的置信度数学表达式为

$$C_i^j = P_r(\text{Object}) \text{IOU} \tag{1}$$

式中 C_i^j ——第 i 个单元格中的第 j 个边界框的置信度

$P_r(\text{Object})$ ——单元格包含某类目标的概

率, $P_r \in \{0, 1\}$

IOU ——预测边界框与真实边界框的交并比

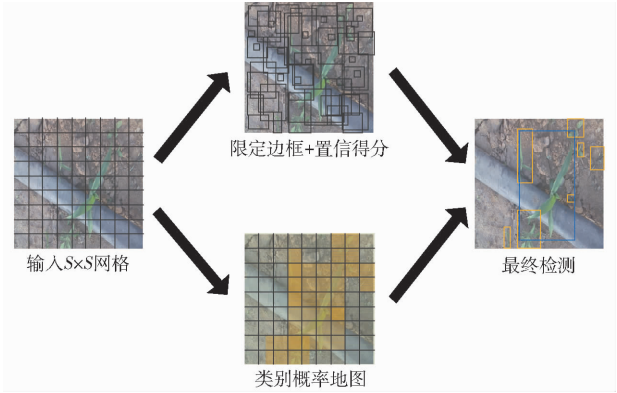


图5 YOLOv4 - tiny 目标检测过程

Fig. 5 Object detection process of YOLOv4 - tiny

2.2 在线数据增强

为提高训练效率和提升小目标的检测精度,本文在训练中采用 Mosaic 在线数据增强策略。主要思想是利用多个样本来产生新的样本,将 4 幅图像进行随机裁剪,再拼接成 1 幅图像,既丰富了物体检测的背景,又可以将中大目标的尺度缩小,进而更好地权衡训练数据的尺度分布,使得模型能够有效提高小目标的检测精度;同时在进行批归一化时,一次便可计算 4 幅图像,减少了对批大小的需求。Mosaic 数据增强图像示例如图 6 所示。



图6 Mosaic 数据增强图像

Fig. 6 Image augmentation image with Mosaic

2.3 先验框聚类分析

原始 YOLOv4 - tiny 模型中采用典型的 K-means 聚类算法对训练图像进行维度聚类分析获得先验框,但 K-means 算法对初始聚类中心选取的随机性会对聚类效果造成一定的影响^[33]。为获得最佳的训练模型,本文中使⤵用 K-means + + 聚类算法代替 K-means 聚类算法,K-means + + 聚类算法的聚类随机性较小,能够有效降低随机选择初始聚类中心所造成的聚类结果偏差。K-means + + 聚类算法流程如下:

(1)输入数据集中玉米及杂草的宽高集合 N 和聚类中心数量 k 。

(2)随机从集合 N 中选取 1 个点,作为初始聚类中心 O_1 。

(3)计算集合 N 中其余各点 x 到其最近聚类中心 O_x 的距离 $D(x)$,距离越远的先验框成为下一个聚类中心的概率 $P(x)$ 越大。重复此步骤,直到找到 k 个聚类中心。

$$D(x) = 1 - IOU(x, O_x) \tag{2}$$

$$P(x) = \frac{D(x)^2}{\sum_{x \in N} D(x)^2} \tag{3}$$

式中 $IOU(x, O_x)$ ——聚类中心与每个标注框的交并比

(4)计算集合 N 中所有点分别到 k 个聚类中心的距离 $D(x)$,将该点划分到距离最小的聚类中心类别中。针对聚类结果,重新计算每个聚类类别中心 C_i 。

$$C_i = \frac{\sum_{x \in C_i} x}{|C_i|} \tag{4}$$

(5)重复步骤(3),当每个聚类类别的聚类中心 C_i 不再变化时,输出 k 个聚类中心结果。

经过 K-means + + 聚类算法,得到了 6 组先验框:(31, 36)、(45, 78)、(90, 83)、(87, 209)、(201, 130)、(215, 363),此时选取的先验框与待识别目标的交并比达到 87.67%。

2.4 模型训练

2.4.1 试验平台

训练平台为 Lenovo legion 台式计算机,主要硬件配置为: Intel(R) Core(TM) i7 - 10700k CPU@ 3.8 GHz, 16 GB 内存, NVIDIA GeForce RTX 2080 SUPER GPU。软件系统环境为 Windows 10 64 位操作系统, CUDA 10.0 版本, CUDNN 7.1 版本, Python 3.7.11 版本, Pytorch 1.2.0 版本。

测试平台为 Jetson Nano 嵌入式平台,主要硬件配置为: 4 - core ARM A57 CPU@ 1.43 GHz, 4 GB 内存, 128 - core Maxwell GPU。软件系统环境为 Ubuntu 18.04 系统, CUDA 10.1 版本, CUDNN 7.5 版本, Python 3.8.10 版本, Pytorch 1.7.1 版本。在训练平台训练好模型后,将模型导入 Jetson Nano 嵌入式平台进行测试。

2.4.2 参数设置

在训练的过程中, YOLOv4 - tiny 模型使用随机梯度下降法(Stochastic gradient descent, SGD)对卷积神经网络的权重参数进行优化, YOLOv4 - tiny 模型经过对训练集图像的多次稳定迭代学习,完成对权重参数的更新和优化,模型训练的具体参数设置如表 1 所示。

表 1 训练参数设置
Tab.1 Training parameter setting

参数	数值
批大小	32
动量	0.9
权重衰减	0.0005
学习率	0.001
迭代次数	300

2.4.3 评价指标

为了分析模型的性能,采用准确率 (Precision, P)、召回率 (Recall, R)、F1 值、类别精度 (Average precision, AP) 及类别精度均值 mAP 作为模型评价指标。

2.4.4 K-means + 聚类算法有效性检验

为检验本文采用的先验框聚类算法的有效性,将 YOLOv4 - tiny 模型分别采用 K-means 聚类算法和 K-means + 聚类算法对原始图像数据集进行训练并测试,检测结果如表 2 所示。

表 2 不同聚类算法对 YOLOv4 - tiny 检测结果的影响
Tab.2 Influence of different clustering algorithms on detection results of YOLOv4 - tiny %

聚类算法	准确 率	召回 率	F1 值	AP		mAP
				玉米	杂草	
K-means	68.2	94.0	79.0	96.7	86.9	91.8
K-means +	68.9	94.1	79.6	97.5	91.8	94.7

由表 2 可知,与采用 K-means 聚类算法的 YOLOv4 - tiny 模型相比,采用 K-means + 聚类算法的 YOLOv4 - tiny 模型准确率提高了 0.7 个百分点,召回率提高了 0.1 个百分点,F1 值提高了 0.6 个百分点,玉米的 AP 提高了 0.8 个百分点,杂草的 AP 提高了 4.9 个百分点,mAP 提高了 2.9 个百分点。采用 K-means + 聚类算法对 YOLOv4 - tiny 模型的各项指标均有所提升,故本文采用 K-means + 聚类算法的 YOLOv4 - tiny 模型进行后续试验。

2.4.5 不同图像增强算法的检测结果对比

为了验证不同的图像特征增强算法对检测结果的影响,本文将采用不同算法增强后的图像分别输入 YOLOv4 - tiny 模型中进行训练并测试,模型检测结果如表 3 所示,mAP 曲线如图 7 所示,训练中损失函数曲线如图 8 所示,PR 曲线如图 9 所示。

由图 7 和表 3 可知,与原始图像相比,经过 CLAHE 图像特征增强后的图像,只有玉米的 AP 提高了 0.4 个百分点,准确率、召回率等其他指标均降低(准确率降低了 1.1 个百分点,召回率降低了 0.3 个百分点,F1 值降低了 0.9 百分点,杂草的 AP 降低了 0.9 个百分点,mAP 降低了 0.3 个百分点);经过

表 3 不同图像特征增强算法对 YOLOv4 - tiny 检测结果的影响
Tab.3 Influence of different image feature enhancement algorithms on detection results of YOLOv4 - tiny %

算法	准确 率	召回 率	F1 值	AP		mAP
				玉米	杂草	
原始图像	68.9	94.1	79.6	97.5	91.8	94.7
CLAHE	67.8	93.8	78.7	97.9	90.9	94.4
同态滤波	68.7	94.1	79.4	97.6	92.1	94.9
MSRCR	70.8	95.2	81.2	97.7	93.7	95.7

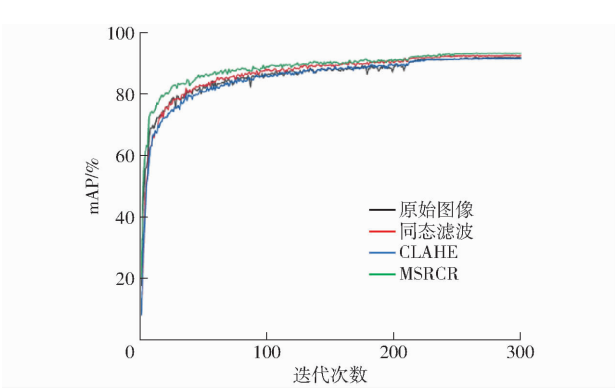


图 7 不同图像特征增强算法的 mAP 曲线

Fig. 7 The mAP curves of different image feature enhancement algorithms

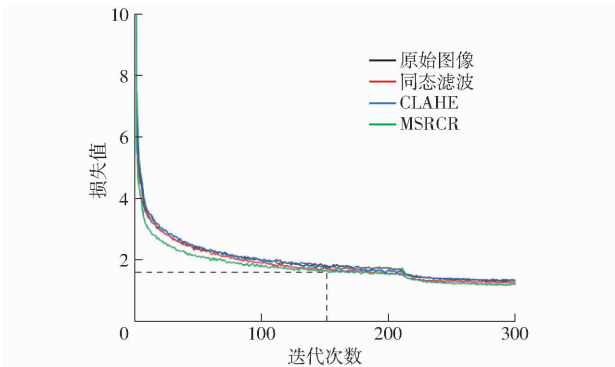


图 8 不同图像特征增强算法的训练损失曲线

Fig. 8 Training loss curves of different image feature enhancement algorithms

同态滤波图像特征增强后的图像,召回率不变,玉米和杂草的 AP、mAP 提高了 0.1 ~ 0.3 个百分点,准确率和 F1 均降低了 0.2 个百分点;经过 MSRCR 图像特征增强后的图像,准确率提高了 1.9 个百分点,召回率提高了 1.1 个百分点,F1 值提高了 1.6 个百分点,玉米的 AP 提高了 0.2 个百分点,杂草的 AP 提高了 1.9 个百分点,mAP 提高了 1.0 个百分点。由以上数据可知,CLAHE 算法和同态滤波算法对各指标的影响较小,对 F1 值和 mAP 等关键指标略有降低,而 MSRCR 算法对玉米和杂草的 AP 等各指标均有明显提升。

由图 8 可知,各损失函数在迭代的初期下降迅速,随着迭代的进行,约在 50 次时,模型损失值下降

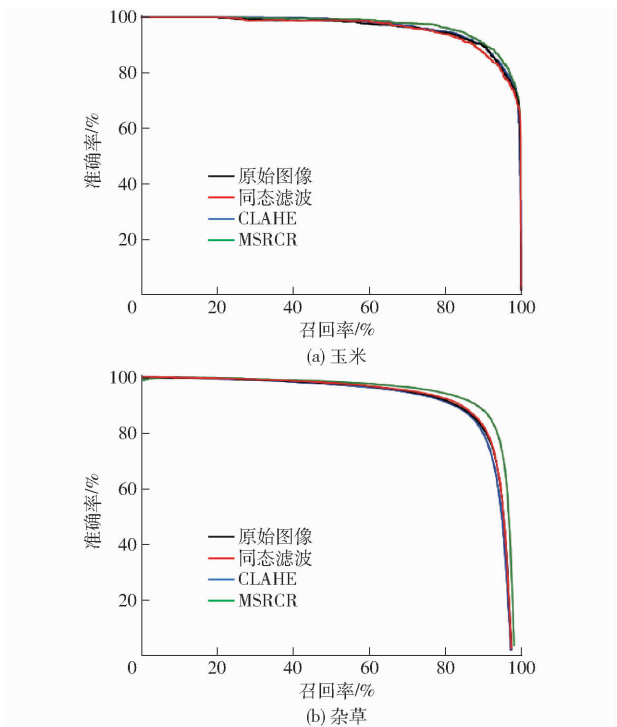


图 9 不同图像特征增强算法的 PR 曲线
Fig.9 PR curves of different image feature enhancement algorithms

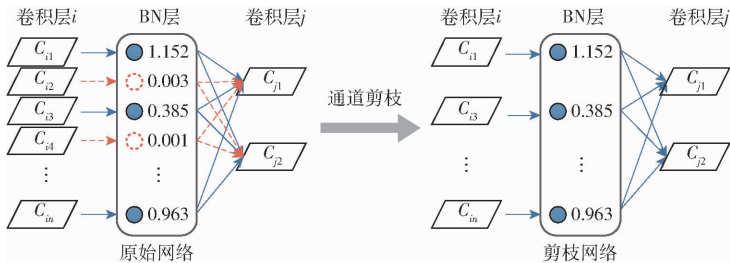


图 10 模型剪枝原理图
Fig.10 Schematic of model pruning

YOLOv4 - tiny 模型剪枝的具体步骤如下：

- (1)基础训练获得原始的 YOLOv4 - tiny 玉米杂草识别模型。
- (2)对识别模型进行稀疏化训练,采用 BN 层中可训练的尺度因子作为指标,并在稀疏训练的过程中进行 L_1 正则化操作,添加 L_1 正则化操作后的损失函数 L 的计算公式为

$$L = l_{\text{yolo}} + \vartheta \sum_{\gamma \in \Gamma} f(\gamma)$$

其中

$f(\gamma) = |\gamma|$

式中 γ ——可训练的尺度因子

Γ —— γ 的集合

$f(\gamma)$ —— L_1 范数正则化函数

l_{yolo} ——训练过程中的损失

ϑ ——平衡损失的惩罚因子

(3)在稀疏训练之后,模型按照设定的剪枝率进行修剪,从而降低模型占用的存储空间。

(4)剪枝后,模型的检测精度等指标会有所下
- 速度开始变缓,迭代至 150 次时,损失值在 1.6 附近波动,模型达到稳定状态。其中,采用 MSRCR 图像特征增强后的图像损失函数曲线收敛最快且最终损失值最低。从图 9 的 PR 曲线来看,MSRCR 图像特征增强后的图像中玉米和杂草的 PR 曲线下包围的面积 (Area under curve, AUC) 明显高于其他 3 组图像,这也体现了 MSRCR 具有良好的效果,因此本文最终选用 MSRCR 图像特征增强后的玉米田间图像进行后续试验。
- 2.5 YOLOv4 - tiny 模型剪枝
- 为了进一步提高运算速度、减少算法对硬件资源的占用,本文对训练后的 YOLOv4 - tiny 模型进行通道剪枝,采用在 BN 层的尺度因子上进行 L_1 正则化的方法。

如图 10 所示,原始网络结构中 BN 层上的尺度因子与卷积层中的每个通道或全连接的神经元关联起来。在训练过程中对这些尺度因子进行稀疏正则化,自动识别尺度因子较小的信道或神经元,然后将这些信道或神经元 (橙色) 去除,得到右侧紧凑的模型,使得压缩后的模型能更好地部署在嵌入式设备中。
- 降,需要对修剪后的模型进行再次训练,微调参数以补偿精度的暂时退化。
- YOLOv4 - tiny 模型剪枝过程的基本参数设置如表 4 所示。
- 表 4 YOLOv4 - tiny 模型剪枝的基本参数设置
Tab.4 Paramaters setting for YOLOv4 - tiny model pruning
- | 处理方式 | 参数 | 数值 |
|------|------|-------|
| 稀疏训练 | 批大小 | 32 |
| | 学习率 | 0.001 |
| | 迭代次数 | 500 |
| 通道剪枝 | 稀疏率 | 0.001 |
| | 剪枝率 | 0.35 |
| 模型微调 | 迭代次数 | 300 |
| | 批大小 | 32 |
- 通道剪枝后模型命名为 Slim - YOLOv4 - tiny, 通道剪枝并微调后模型命名为 Prune - YOLOv4 - tiny, 分别在台式计算机和 Jetson Nano 嵌入式平台

上运行评估,评估结果如表 5 所示。与 YOLOv4 - tiny 模型相比,Slim - YOLOv4 - tiny 模型的总参数量发生明显变化,降为 YOLOv4 - tiny 模型的 54.3%,模型占用内存减少为 YOLOv4 - tiny 模型的 53.8%,在台式计算机检测平均帧耗时降为 YOLOv4 - tiny 模型的 75%,在 Jetson Nano 检测平均帧耗时降为 YOLOv4 - tiny 模型的 67.3%,达到了简化模型的目的,但模型的准确率、召回率、F1 值及 mAP 均有下降,分别降低了 5.1、11.4、7.5、24.6 个

百分点,这是由于剪枝后暂时导致的精度损失,需要后续的微调过程进行补偿,经过微调的模型在许多情况下可以达到比原始未剪枝模型更高的精度^[34]。微调后 Prune - YOLOv4 - tiny 模型总参数量、模型占用内存及在台式计算机检测平均帧耗时略有增加,但较 YOLOv4 - tiny 模型可忽略不计,较 Slim - YOLOv4 - tiny 模型的准确率、召回率、F1 值及 mAP 值均有提高,分别提高了 8.4、12.1、9.9、27.1 个百分点,达到了比 YOLOv4 - tiny 模型更好的精度,在

表 5 模型剪枝的评估结果
Tab.5 Test evaluation results of model pruning

模型	准确率/ %	召回率/ %	F1 值/ %	mAP/ %	总参数量/ 个	模型占用 内存/MB	台式计算机检测 平均帧耗时/ms	Jetson Nano 检测 平均帧耗时/ms
YOLOv4 - tiny	70.8	95.2	81.2	94.1	5 918 006	22.5	4.0	168
Slim - YOLOv4 - tiny	65.7	83.8	73.7	69.5	3 212 961	12.1	3.0	113
Prune - YOLOv4 - tiny	74.1	95.9	83.6	96.6	3 239 297	12.2	3.2	131

Jetson Nano 检测平均帧耗时降为 YOLOv4 - tiny 模型的 77.9%。结果表明,采用模型剪枝方法后 YOLOv4 - tiny 模型的性能得到了提高。

2.6 不同算法检测效果对比

为验证 Prune - YOLOv4 - tiny 模型的有效性,

在台式计算机上使用本文训练集和 MSRCR 图像特征增强算法对常用的目标检测模型 Faster RCNN、YOLOv3 - tiny、YOLOv4 进行训练,并在台式计算机和 Jetson Nano 嵌入式平台上分别使用测试集对 4 种算法性能进行评估,评估结果如表 6 所示。

表 6 不同算法检测结果比较
Tab.6 Detection results comparison of different algorithms

模型	准确率/%	召回率/%	F1 值/%	AP/%		mAP/%	模型占用 内存/MB	台式计算机检测 平均帧耗时/ms	Jetson Nano 检测 平均帧耗时/ms
				玉米	杂草				
Faster RCNN	68.8	89.2	77.7	88.2	60.7	74.5	359	35.4	
YOLOv3 - tiny	71.4	92.6	80.6	94.6	91.3	93.0	33.1	4.3	408
YOLOv4	88.9	98.2	93.3	99.0	96.6	97.8	244	24.7	1 725
Prune - YOLOv4 - tiny	74.1	95.9	83.6	98.2	94.9	96.6	12.2	3.2	131

Faster RCNN 的模型占用内存是 Prune - YOLOv4 - tiny 的 29.4 倍,mAP 比 Prune - YOLOv4 - tiny 低 22.1 个百分点,并且由于参数过多无法在 Jetson Nano 嵌入式平台检测图像。YOLOv3 - tiny 在台式计算机检测平均帧耗时与 Prune - YOLOv4 - tiny 模型相差不大,但在 Jetson Nano 嵌入式平台检测平均帧耗时是 Prune - YOLOv4 - tiny 模型的 3 倍。YOLOv4 模型的 mAP 虽然比 Prune - YOLOv4 - tiny 模型高 1.2 个百分点,但模型占用内存是 Prune - YOLOv4 - tiny 模型的 20 倍,在 Jetson Nano 嵌入式平台检测平均帧耗时是 Prune - YOLOv4 - tiny 模型的 13 倍。

与其他 3 种模型相比,Prune - YOLOv4 - tiny 模型具有较高的准确率、召回率、F1 值和 mAP,模型剪枝后占用内存小,实现了速度与精度的平衡,并适用

于嵌入式平台。

图 11 展示了不同算法的检测效果,为了直观比较检测结果,本文采用原始标注图像作为对比。从图 11 可以看出,Faster RCNN 模型存在漏检情况,如图 11a、11b 中漏检了部分小目标杂草,而 YOLOv3 - tiny、YOLOv4、Prune - YOLOv4 - tiny 的检测结果与原始标注图像基本相同。图 11c、11d 中存在多株杂草重叠及玉米与杂草互相遮挡,Faster RCNN 对于重叠的杂草有漏检情况,YOLOv3 - tiny 和 YOLOv4 对于小目标多株杂草重叠检测存在冗余框。相比于 Faster RCNN、YOLOv3 - tiny 和 YOLOv4 这 3 种模型,Prune - YOLOv4 - tiny 模型的检测效果更好,对大目标和小目标均能有效识别,且不受玉米与杂草互相遮挡、多株杂草重叠、背景存在杂物以及自然光线下玉米及杂草阴影等影响。

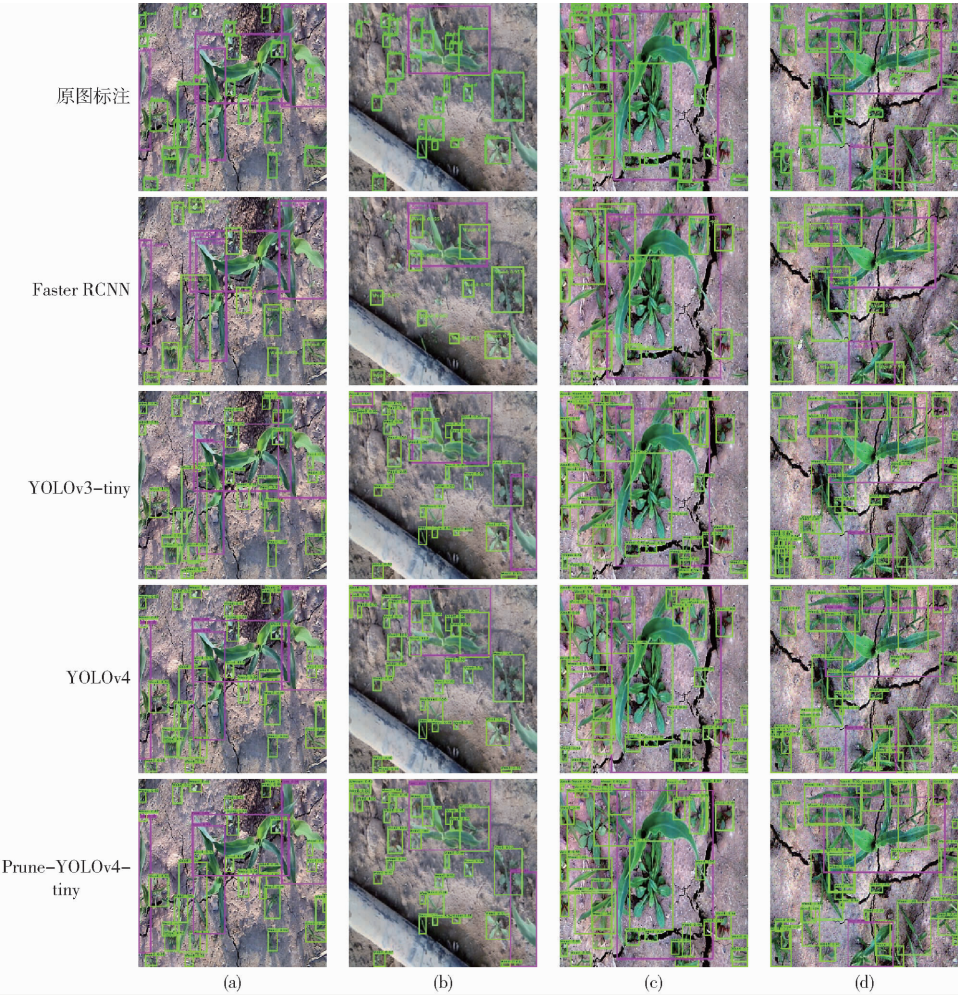


图 11 不同算法检测效果对比

Fig. 11 Detection effect comparison of different algorithms

3 结论

(1) 提出了一种融合 MSRCR 算法和改进 YOLOv4-tiny 目标检测模型的田间环境下玉米苗和杂草的识别方法。试验结果表明,在嵌入式平台上,Prune-YOLOv4-tiny 模型的综合评价指标 F1 值为 83.6%,玉米苗的 AP 为 98.2%,杂草的 AP 为 94.9%,mAP 为 96.6%,检测平均帧耗时 131 ms。满足实际生产需要,可用于实际的田间环境进行玉米苗和杂草的识别,为后续除草机器人的视觉系统提供良好的技术支持。

(2)MSRCR 算法克服了原始图像对比度低、细节模糊等问题,使得低像素的图像中特征更易提取,图像特征增强后 F1 值提高了 1.6 个百分点,杂草的 AP 提高了 1.9 个百分点,mAP 提高了 1.0 个百

分点。
(3)采用 Mosaic 在线数据增强、K-means++ 聚类算法重构先验框及模型剪枝处理对 YOLOv4-tiny 模型进行改进,得到的 Prune-YOLOv4-tiny 模型占用内存为 12.2 MB,与 Faster RCNN、YOLOv3-tiny 和 YOLOv4 模型进行对比,结果表明:Prune-YOLOv4-tiny 的 mAP 分别比 Faster RCNN 和 YOLOv3-tiny 高 22.1 个百分点和 3.6 个百分点,比 YOLOv4 低 1.2 个百分点;Prune-YOLOv4-tiny 模型占用内存是 Faster RCNN 的 3.4%、YOLOv3-tiny 的 36.9%、YOLOv4 的 5%;在 Jetson Nano 移动平台的检测平均帧耗时分别是 YOLOv3-tiny 和 YOLOv4 的 32.1% 和 7.6%。综合对比以上数据,本文提出的方法在保证检测精度的前提下,具有模型占用内存小、运算速度快的显著优势。

参 考 文 献

[1] PATEL D, KUMBHAR B. Weed and its management: a major threats to crop economy[J]. Journal of Pharmaceutical Science and Bioscientific Research, 2016,6(6): 753-758.
[2] 伍同,曾山,赵润茂,等. 智能化作物株间机械除草技术分析与研究[J]. 农机化研究, 2019, 41(6): 1-6, 12.
WU Tong, ZENG Shan, ZHAO Runmao, et al. Analysis and study on intelligentized technologies of intra-row weeding

- mechanical control[J]. Journal of Agricultural Mechanization Research, 2019, 41(6): 1 - 6, 12. (in Chinese)
- [3] 李香菊. 近年我国农田杂草防控中的突出问题与治理对策[J]. 植物保护, 2018, 44(5): 77 - 84.
LI Xiangju. Main problems and management strategies of weeds in agricultural fields in China in recent years[J]. Plant Protection, 2018, 44(5): 77 - 84. (in Chinese)
- [4] LIU B, BRUCH R. Weed detection for selective spraying: a review[J]. Current Robotics Reports, 2020, 1(1): 19 - 26.
- [5] SHANMUGAM S, ASSUNCAO E, MESQUITA R, et al. Automated weed detection systems: a review [J/OL]. KnE Engineering, 2020: 271 - 284. DOI:10.18502/keg.v5i6.7046.
- [6] 王璨, 李志伟. 利用融合高度与单目图像特征的支持向量机模型识别杂草[J]. 农业工程学报, 2016, 32(15): 165 - 174.
WANG Can, LI Zhiwei. Weed recognition using SVM model with fusion height and monocular image features[J]. Transactions of the CSAE, 2016, 32(15): 165 - 174. (in Chinese)
- [7] 李慧, 祁力钧, 张建华, 等. 基于 PCA - SVM 的棉花出苗期杂草类型识别[J]. 农业机械学报, 2012, 43(9): 184 - 189, 196.
LI Hui, QI Lijun, ZHANG Jianhua, et al. Recognition of weed during cotton emergence based on principal component analysis and support vector machine[J]. Transactions of the Chinese Society for Agricultural Machinery, 2012, 43(9): 184 - 189, 196. (in Chinese)
- [8] 白敬, 徐友, 魏新华, 等. 基于光谱特性分析的冬油菜苗期田间杂草识别[J]. 农业工程学报, 2013, 29(20): 128 - 134.
BAI Jing, XU You, WEI Xinhua, et al. Weed identification from winter rape at seedling stage based on spectrum characteristics analysis[J]. Transactions of the CSAE, 2013, 29(20): 128 - 134. (in Chinese)
- [9] PULIDO C, SOLAQUE L, VELASCO N. Weed recognition by SVM texture feature classification in outdoor vegetable crop images[J]. Ingeniería e Investigación, 2017, 37(1): 68 - 74.
- [10] KRIZHEVSKY A, SUTSKEVER I, HINTON G E. ImageNet classification with deep convolutional neural networks[C] // Advances in Neural Information Processing Systems, 2012: 1097 - 1105.
- [11] 李丹, 张凯锋, 李行健, 等. 基于 Mask R - CNN 的猪只爬跨行为识别[J]. 农业机械学报, 2019, 50(增刊): 261 - 266, 275.
LI Dan, ZHANG Kaifeng, LI Xingjian, et al. Mounting behavior recognition for pigs based on Mask R - CNN [J]. Transactions of the Chinese Society for Agricultural Machinery, 2019, 50(Supp.): 261 - 266, 275. (in Chinese)
- [12] 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 Juntao, ZHENG Zhenhui, LIANG Jiaen, et al. Citrus detection method in night environment based on improved YOLO v3 network[J]. Transactions of the Chinese Society for Agricultural Machinery, 2020, 51(4): 199 - 206. (in Chinese)
- [13] 樊湘鹏, 周建平, 许燕, 等. 基于改进卷积神经网络的复杂背景下玉米病害识别[J]. 农业机械学报, 2021, 52(3): 210 - 217.
FAN Xiangpeng, ZHOU Jianping, XU Yan, et al. Corn disease recognition under complicated background based on improved convolutional neural network[J]. Transactions of the Chinese Society for Agricultural Machinery, 2021, 52(3): 210 - 217. (in Chinese)
- [14] GIRSHICK R, DONAHUE J, DARRELL T, et al. Region-based convolutional networks for accurate object detection and segmentation[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2015, 38(1): 142 - 158.
- [15] GIRSHICK R, DONAHUE J, DARRELL 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.
- [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 and Machine Intelligence, 2015, 39(6): 1137 - 1149.
- [17] HE K, GKIOXARI G, DOLLAR P, et al. Mask R - CNN [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2017(99): 2961 - 2969.
- [18] HE K M, ZHANG X Y, REN S Q, et al. Spatial pyramid pooling in deep convolutional networks for visual recognition[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2015, 37(9): 1904 - 1916.
- [19] REDMON J, DIVVALA S, GIRSHICK R, et al. You only look once: unified, real-time object detection[C] // Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2016: 779 - 788.
- [20] LIU W, ANGUELOV D, ERHAN D, et al. SSD: single shot multiBox detector[C] // European Conference on Computer Vision. Cham: Springer, 2016: 21 - 37.
- [21] PEI D, JING M, LIU H. A fast RetinaNet fusion framework for multi-spectral pedestrian detection[J]. Infrared Physics & Technology, 2020, 105: 103178.
- [22] DUAN K, BAI S, XIE L. Centernet: keypoint triplets for object detection[C] // Proceedings of the IEEE International Conference on Computer Vision, 2019: 6569 - 6578.
- [23] 樊湘鹏, 周建平, 许燕, 等. 基于优化 Faster R - CNN 的棉花苗期杂草识别与定位[J]. 农业机械学报, 2021, 52(5): 26 - 34.
FAN Xiangpeng, ZHOU Jianping, XU Yan, et al. Identification and localization of weeds based on optimized Faster R - CNN in cotton seedling stage[J]. Transactions of the Chinese Society for Agricultural Machinery, 2021, 52(5): 26 - 34. (in Chinese)
- [24] 姜红花, 张传银, 张昭, 等. 基于 Mask R - CNN 的玉米田间杂草检测方法[J]. 农业机械学报, 2020, 51(6): 220 - 228, 247.
JIANG Honghua, ZHANG Chuanyin, ZHANG Zhao, et al. Detection method of corn weed based on Mask R - CNN [J]. Transactions of the Chinese Society for Agricultural Machinery, 2020, 51(6): 220 - 228, 247. (in Chinese)