

基于改进 YOLO v4 网络的马铃薯自动育苗叶芽检测方法

修春波^{1,2} 孙乐乐¹

(1. 天津工业大学控制科学与工程学院, 天津 300387;
2. 天津工业大学天津市电气装备智能控制重点实验室, 天津 300387)

摘要: 为提高马铃薯幼苗叶芽检测识别的准确率,提高自动育苗生产系统的工作效率,提出了基于 YOLO v4 网络的改进识别网络。将 YOLO v4 特征提取部分 CSPDarknet53 中的残差块(Residual Block)替换为 Res2Net,并采用深度可分离卷积操作减小计算量。由此,在增大卷积神经网络感受野的同时,能够获得叶芽更加细小的特征信息,减少马铃薯叶芽的漏检率。设计了基于扩张卷积的空间特征金字塔(D-SPP 模块),并嵌入和替换到特征提取部分的 3 个特征层输出中,用于提高马铃薯叶芽目标识别定位的准确性。采用消融实验对改进策略的有效性进行了验证分析。实验结果表明,改进的识别网络对马铃薯叶芽检测的精确率为 95.72%,召回率为 94.91%,综合评价指标 F1 值为 95%,平均精确率为 96.03%。与 Faster R-CNN、YOLO v3、YOLO v4 网络相比,改进的识别网络具有更好的识别性能,从而可有效提高马铃薯自动育苗生产系统的工作效率。

关键词: 深度学习; 马铃薯; 叶芽检测; 扩张卷积; 感受野; YOLO v4

中图分类号: S24 文献标识码: A 文章编号: 1000-1298(2022)06-0265-09 OSID: 

Potato Leaf Bud Detection Method Based on Improved YOLO v4 Network

XIU Chunbo^{1,2} SUN Lele¹

(1. School of Control Science and Engineering, Tiangong University, Tianjin 300387, China
2. Tianjin Key Laboratory of Intelligent Control of Electrical Equipment, Tiangong University, Tianjin 300387, China)

Abstract: In order to improve the precision of the detection and recognition of the potato seedling leaf bud and improve the efficiency of the automatic seedling production system, an improved recognition network based on the YOLO v4 network was proposed. The residual block in the feature extraction part CSPDarknet53 was replaced with Res2Net, and the depthwise separable convolution was used to reduce the computation. In this way, the receptive field of the convolutional neural network can be enlarged, the finer feature information of leaf bud can be got, and the missed detection rate of potato leaf bud can be reduced. Furthermore, a spatial feature pyramid (D-SPP module) based on dilated convolution was designed and embedded in the output of the three feature layers of the feature extraction part to improve the recognition and localization precision of potato leaf bud target. The ablation experiment was used to verify the effectiveness of the improved strategies. The experiment results showed that the recognition precision, the recall rate, the comprehensive evaluation index F1 value and the average precision of the improved network were 95.72%, 94.91%, 95% and 96.03% respectively. Comparing with the common networks such as Faster R-CNN, YOLO v3 and YOLO v4, the improved network had the better recognition performances, thus the production efficiency automatic seedling production system can be enhanced.

Key words: deep learning; potato; leaf bud detection; dilated convolution; receptive field; YOLO v4

0 引言

农业生产智能化技术的应用可以有效提高生产

效率、节约人工成本,在农业生产中的多个领域得到广泛应用^[1-3]。例如,采用深度学习技术可实现对马铃薯叶片病害情况的智能检测^[4]、病变的准确分

收稿日期: 2021-07-15 修回日期: 2021-09-19
基金项目: 天津市自然科学基金项目(18JCYBJC88300、18JCYBJC88400)
作者简介: 修春波(1978—),男,教授,博士,主要从事人工智能及应用研究,E-mail: xiuchunbo@tiangong.edu.cn

割^[5]以及病害有效诊断^[6],以此提高马铃薯种植生产效率。而马铃薯种植产业的首要环节是马铃薯的幼苗培育,马铃薯幼苗培育过程中的幼苗剪切操作由于工艺相对复杂,目前仍以人工操作为主,在幼苗剪切过程中容易出现细菌感染的情况,从而导致生产成本增加,且生产效率不高。如果采用自动育苗技术就能避免操作污染,提高育苗成活率,提高育苗栽培效率。马铃薯自动化育苗系统中的核心机构为幼苗剪切装置,可利用计算机视觉技术对马铃薯幼苗的叶芽进行定位,从而确定剪切位置。由于叶芽在自然生长过程中大小、姿态具有随机性,因此,相机所采集的图像会出现叶芽与茎秆之间的遮挡情况。传统的目标识别与定位方法很难满足实际工况场景的要求^[7-9],目前鲜有技术成熟马铃薯幼苗培育自动化系统应用于实际生产中。

目前深度学习的快速发展为求解目标识别与定位问题提供了新技术^[10]。深度学习网络能够实现特征的自动选择与提取,有效提高了特征提取的质量和效率。基于目标识别与定位的原理,深度学习网络大致可分为 Two-stage 和 One-stage 两类检测方法。Two-stage 类检测方法将目标检测过程分为 2 个阶段,第 1 个阶段在待检测图像中生成一系列样本候选区域,第 2 个阶段利用深度学习网络对候选区域进行分类识别,从而完成目标检测任务。Two-stage 检测方法的典型网络有 R-CNN 网络^[11-12]、Fast R-CNN^[13]、Faster R-CNN^[14-16]等,这类方法通常具有较高的检测精度,但也存在训练时间较长、检测速度较慢等缺点。相比较而言,One-stage 类检测方法在目标检测过程中不需要产生候选框,而是将目标边框定位问题转化为回归问题进行求解,典型的网络有 SSD^[17-18]、YOLO 系列^[19-22]等网络,由于这类网络能够直接定位出目标位置,因此通常具有较快的检测速度,但目标定位精度会受到一定的影响。在农业生产中,基于深度学习的目标检测技术通常只能应用于较大的水果目标检测识别中^[23-24],而在马铃薯幼苗叶芽这类容易出现遮挡的复杂小目标识别中应用效果并不理想。

为了有效解决马铃薯幼苗的叶芽识别与定位问题,尤其提高对茎叶遮挡以及小叶芽目标的识别准确性,在 YOLO v4 网络的基础上,构建改进的目标识别网络。改进的识别网络通过使用基于深度可分离卷积的 Res2Net 模块优化特征提取部分的结构,利用扩张卷积构建空间特征金字塔,实现各特征层输出中特征图的不同尺度感受野特征信息的有机融合,以丰富特征融合部分信息,从而保证网络既具有一定的检测速度,又具有较高的识别准确性。

1 马铃薯幼苗自动剪切样机

马铃薯自动育苗系统需要实现试管苗的传送、试管苗裁剪等功能,在无尘、无菌、恒温、恒湿的自动化生产车间中,能够有效降低试管苗被细菌感染的风险,提高育苗成活率。整个育苗生产工艺中,试管苗的自动剪切是核心技术,为此,设计了马铃薯幼苗自动剪切样机,如图 1 所示。

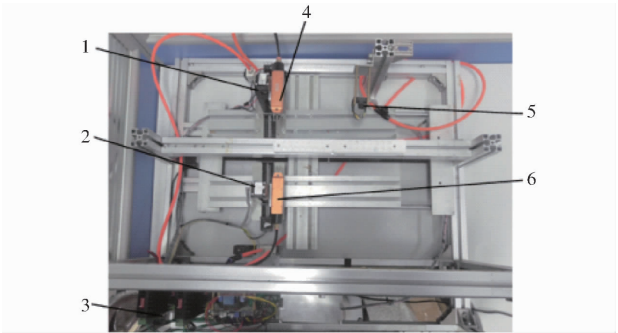


图 1 马铃薯幼苗自动剪切样机

Fig. 1 Potato seedling automatic shearing equipment

1、2. 导轨 3. 控制系统 4、6. 剪切刀具 5. 摄像头

马铃薯幼苗自动剪切样机主要由剪切刀具、摄像头、导轨、控制系统组成。其中摄像头用于获取马铃薯幼苗图像,剪切刀具用于剪切马铃薯叶芽,导轨用于搭载刀具运动,控制系统用于控制导轨以及刀具运动。剪切平台搭载具有剪切功能的刀具,待剪切试管苗由传送机构送至裁剪工位,利用摄像机采集试管苗图像,采用基于深度学习的识别网络实现叶芽识别与定位,机械运动平台搭载刀具运动至剪切位置,利用气动机构驱动刀具实现幼苗剪切,待所有叶芽均完成剪切后,传送机构将其送出,完成剪切过程。

2 识别网络设计

2.1 数据预处理

马铃薯自动剪切平台利用培养液培育带有叶芽的马铃薯单节茎段,生根发芽后再生成单株幼苗。在马铃薯育苗生产工艺中,要求各裁剪段至少保留一片叶芽,这样才能确保试管苗成功繁殖。因此,叶芽位置的准确识别与定位直接决定了育苗系统的生产效率及育苗成本。

为了实现叶芽的自动检测,需要建立样本数据集。图像样本采集于培养液中自主培育的冀张薯 12 号马铃薯幼苗。采用 OV5640 型摄像头采集马铃薯幼苗图像,图像尺寸为 2 592 像素 × 1 944 像素。将采集到的图像分割为 400 幅单株幼苗样本图像,其中部分样本如图 2 所示。

由图 2 可见,马铃薯幼苗由于自然生长导致叶

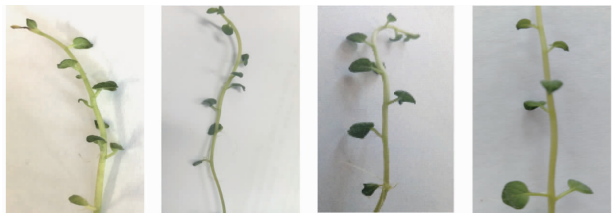


图 2 马铃薯幼苗样本图像
Fig. 2 Potato seedling sample images

芽形态各异、大小不一，且叶芽与秆茎之间存在遮挡，这对准确识别叶芽位置造成了极大的困难。

为了提高识别的准确性，采用图像旋转、亮度变化等数据增强技术将样本图像数量扩充至 1 400 幅，以此提高样本的多样性，改善识别网络的泛化能力。部分数据增强样本图像如图 3 所示。

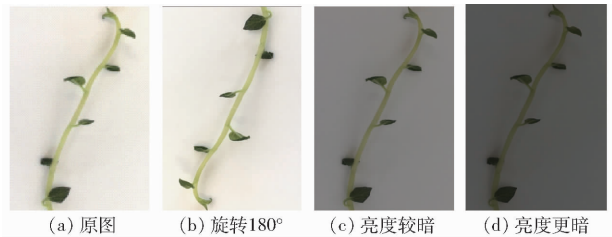


图 3 数据增强样本图像
Fig. 3 Data enhancement sample images

使用 LabelImg 标注软件对样本图像中马铃薯叶芽位置进行数据标注，获得样本类别和坐标信息，将数据样本按照比例 8: 1: 1 划分为训练集、验证集和测试集。

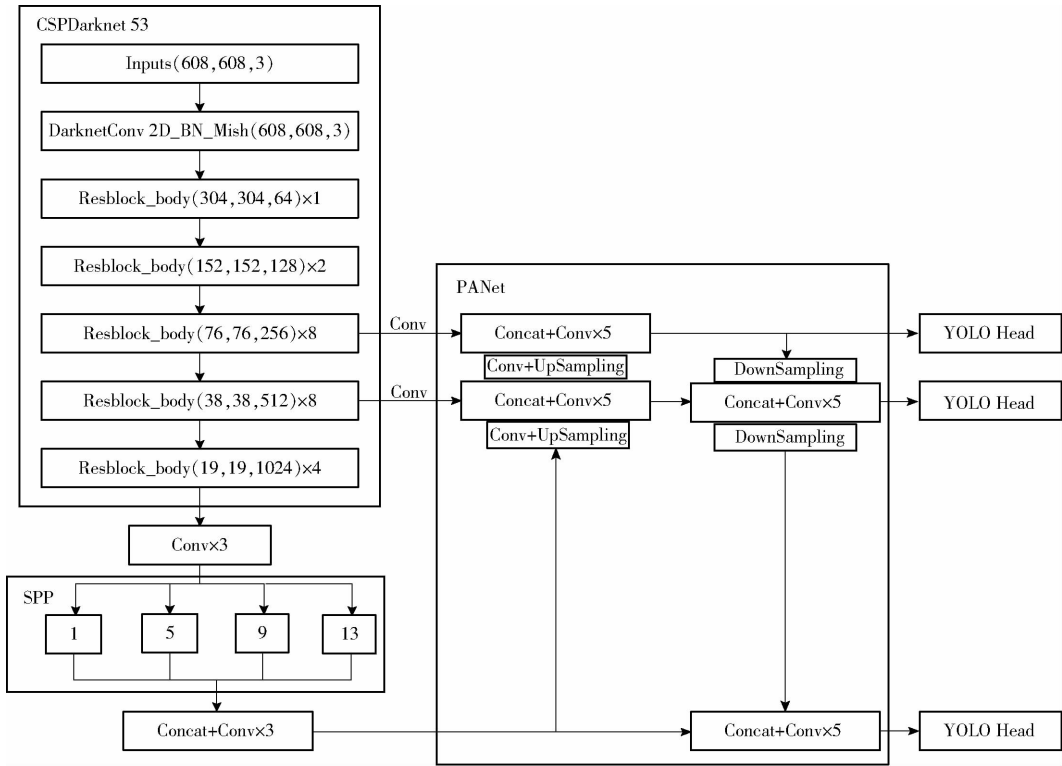


图 4 YOLO v4 网络结构图
Fig. 4 Structure diagram of YOLO v4 network

2.2 叶芽识别网络

YOLO v4 网络是 YOLO 系列网络中性能较好的目标检测与识别网络，其主要结构如图 4 所示。

图 4 中，YOLO v4 网络结构主要包括特征提取 (Backbone)、颈部特征融合 (Neck) 和检测头 (Head) 3 部分。特征提取部分的 CSPDarknet53 由 5 个 Resblock_body 模块组成，各 Resblock_body 模块中包含不同数量的 Residual Block 结构，并且特征提取部分的最后一层特征层输出中加入了 SPP 模块，用来融合多尺度的感受野信息，从而分离出显著的上下文特征；颈部的 PANet^[25] 用于实现特征提取部分中的特征信息从深层到浅层、再从浅层到深层的双向融合。检测头部分用于获取 3 个尺度的输出特征，并对目标进行检测，获得目标的类别信息以及位置信息。

YOLO v4 网络在对大中型显著分离的目标识别过程中具有良好的识别性能，但由于马铃薯幼苗叶芽存在目标小、遮挡严重等情况，因此直接利用 YOLO v4 网络进行马铃薯幼苗叶芽识别的性能并不理想，为此对 YOLO v4 网络在获得细粒度特征信息和准确的位置信息两方面进行改进，使其适应马铃薯幼苗叶芽的检测识别。

2.2.1 特征提取部分改进

YOLO v4 网络中的 Residual Block 结构借鉴了 ResNet 结构的设计思想。ResNet 结构能够有效避免因网络深度增加而导致的梯度消失或退化问题。

在利用 YOLO v4 网络对马铃薯幼苗的叶芽进行检测识别的过程中,会出现一定数量的漏检、错检等问题。在 ResNet 的基础上提出的 Res2Net 结构^[26-27]利用分层的残差连接,在增加每个网络层的感受视野同时,能够在细粒度层次上表示特征信息,从而能够提取叶芽更多的细小特征信息,减少漏检、错检的情况。ResNet 及 Res2Net 结构如图 5 所示。

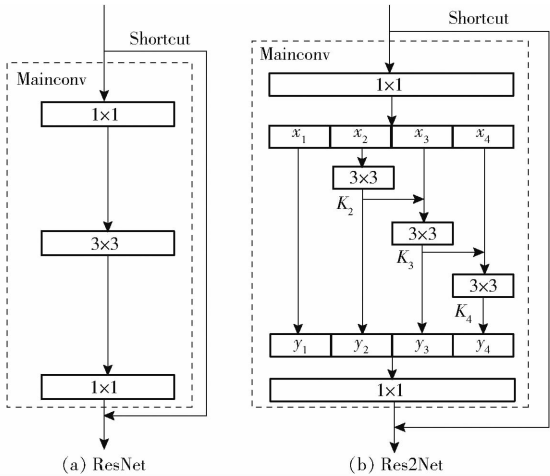


图 5 ResNet 及 Res2Net 的结构图

Fig. 5 Structure diagrams of ResNet and Res2Net

ResNet 主要由 Shortcut 和 Mainconv 两部分组成。Shortcut 部分利用恒等映射将输入特征图生成残差边,Mainconv 作为特征提取部分由 2 个 1×1 卷积层和 1 个 3×3 卷积层组成,利用 1×1 卷积层对输入特征图通道进行降维,利用 3×3 卷积层对降维后的特征图进行特征提取,再利用 1×1 卷积层将提取特征后的特征图通道恢复到输入特征图的维度,以实现提取的特征信息与残差边特征信息的融合。

图 5b 为 Res2Net 结构,其保留了 ResNet 中的 2 个 1×1 卷积层,对 3×3 卷积层进行了改造。输入特征图经过 1×1 卷积层降维后分成 4 个子特征图,每个子特征图表示为 $x_i (i = 1, 2, 3, 4)$,其中子特征图 x_1 不进行 3×3 卷积操作,直接进行子特征图输出,将输出特征图记为 y_1 ,其他子特征图 $x_i (i = 2, 3, 4)$ 均进行 3×3 卷积操作,记作 $K_i (i = 2, 3, 4)$,将其输出特征图记作 $y_i (i = 2, 3, 4)$,子特征图 x_i 和前一个输出特征图 y_{i-1} 相加后进行 3×3 卷积操作,通过“Concat”操作对输出的各部分特征图 y_i 进行拼接操作。除 y_1 外,每个输出特征图 y_i 均融合其前一个输出特征图的特征信息,每个输出特征图的感受野得到增大的同时,各输出特征图利用“Concat”操作进行拼接,也融合了叶芽多尺度的细粒度特征信息。

为了减少网络的计算量,构建了基于深度可分离卷积的 Res2Net 模块,即将 Res2Net 模块中的 3×3 卷积替换为深度可分离卷积。深度可分离卷

积^[28]由逐通道卷积和逐点卷积两部分组成。图 6 为输入输出通道数均为 3 的深度可分离卷积操作示意图。

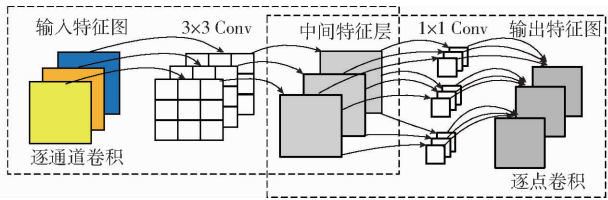


图 6 深度可分离卷积

Fig. 6 Depthwise separable convolution

图 6 中,逐通道卷积对每个输入通道的特征图进行 3×3 卷积操作,生成与输入通道数相同的特征图。逐点卷积对逐通道卷积生成的特征图进行 1×1 卷积操作,生成输出特征图。

对于传统卷积操作,如果输入通道数为 N ,输出通道数为 C ,卷积核尺寸为 K ,则其计算量 O_1 为

$$O_1 = K \times K \times N \times C \tag{1}$$

而对于深度可分离卷积操作,其计算量 O_2 为

$$O_2 = K \times K \times N + 1 \times 1 \times N \times C \tag{2}$$

由此可见,当 K 与 C 均不为 1 时,深度可分离卷积可以有效减少传统卷积的计算量。

因此,为了降低叶芽检测识别过程中的漏检率和错检率,并减少网络模型的计算参数,采用深度可分离卷积替换 Res2Net 模块中的 3×3 卷积,同时采用改进后的 Res2Net 模块替换 CSPDarknet 53 中的 Residual Block 结构,在细粒度层次上表示特征信息,获得叶芽细小特征信息,从而提高叶芽识别的准确性。

2.2.2 D-SPP 模块嵌入设计

YOLO v4 网络在特征提取部分的最后一个特征层输出中使用了 SPP 模块^[29],SPP 模块的结构如图 7 所示。

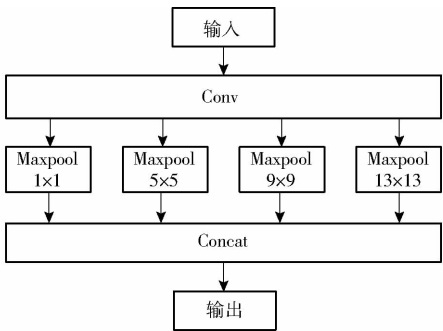


图 7 SPP 模块结构图

Fig. 7 Structure diagram of SPP module

图 7 中,特征图经过卷积操作后分别进行核尺寸为 $1 \times 1, 5 \times 5, 9 \times 9, 13 \times 13$ 的最大池化操作,然后通过“Concat”操作将池化操作后的特征图进行拼接,输出通道数变为原来通道数的 4 倍。

SPP 模块可以获得局部感受野以及近似全局感受野的特征信息,融合不同尺度感受野的特征信息,丰富特征图的表达能力。但输入的特征图经过 SPP 模块最大池化操作会丢失目标的位置信息,为此,采用扩张卷积^[30]操作替换 SPP 模块中的最大池化操作,在保持原感受野不变的情况下,确保目标位置信息不丢失。

扩张卷积是在标准卷积核中注入空洞,增加了一个参数用来定义卷积的扩张率,即卷积核之间注入的空洞数量,在不增加卷积核数量的情况下可以获得局部或者近似全局的感受野信息,以获得更多的目标位置信息。空洞卷积核尺寸 k^* 为

$$k^* = (k - 1) \times r + 1 \quad (3)$$

式中 k ——扩张前卷积核尺寸

r ——扩张率

为了能够实现扩张卷积操作, k 的最小值为 3。此外,根据特征图尺寸选择合适的扩张率能够获得近似全局的感受野。

所设计的基于扩张卷积的 SPP 模块命名为 D-SPP 模块,其结构图如图 8 所示。

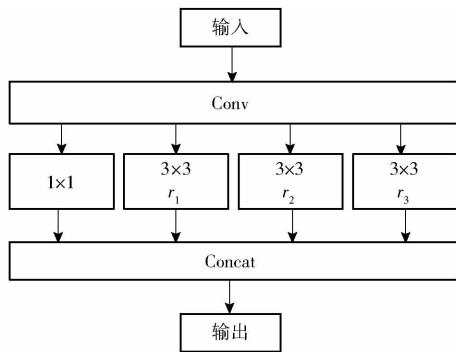


图 8 D-SPP 模块结构图

Fig. 8 Structure diagram of D-SPP module

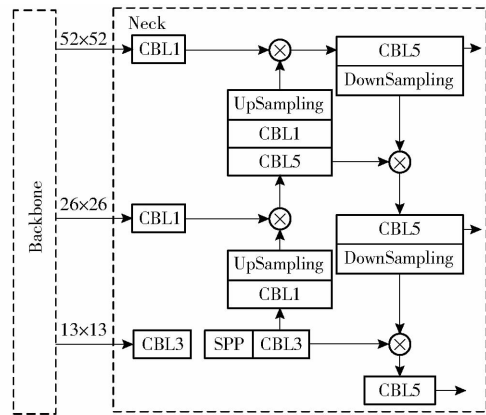
图 8 中,D-SPP 模块采用扩张率分别为 $r_1 = 2$ 、 $r_2 = 4$ 、 $r_3 = 6$ 的 3×3 扩张卷积操作替换了 SPP 模块中的最大池化操作,可获得 5×5 、 9×9 的局部感受野以及 13×13 的全局感受野。这样,D-SPP 模块的感受野与原 SPP 模块的感受野相同,并且避免了目标位置信息的丢失。

由于 YOLO v4 特征提取部分输出的 3 个不同深度的特征层中,浅层特征层的特征图具有较少的语义信息,但目标位置准确,而深层特征层的特征图具有较多的语义信息,但目标位置不够准确。为此,将 D-SPP 模块嵌入到 52×52 和 26×26 的特征图所在特征层输出中,融合不同感受野的目标位置信息,以提高目标定位的准确性。

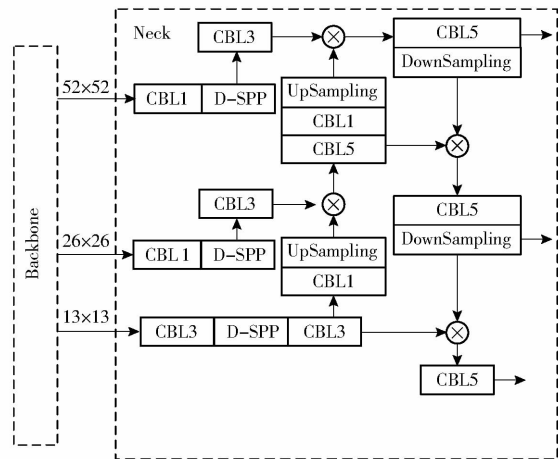
将扩张率分别为 $r_1 = 8$ 、 $r_2 = 16$ 、 $r_3 = 24$ 的 D-SPP 模块嵌入到 52×52 的特征图所在特征层的输

出中,可获得 17×17 、 33×33 的局部感受野和 49×49 的近似全局感受野;将扩张率分别为 $r_1 = 4$ 、 $r_2 = 8$ 、 $r_3 = 12$ 的 D-SPP 模块嵌入到 26×26 的特征图所在特征层输出中,可获得 9×9 、 17×17 的局部感受野以及 25×25 的近似全局感受野。

D-SPP 模块在 3 个特征层输出中的具体位置如图 9 所示。



(a) 原特征层输出部分



(b) D-SPP 嵌入位置

图 9 D-SPP 模块嵌入到网络中的位置

Fig. 9 Location of D-SPP module in network

改进网络在特征提取部分 3 个不同深度的特征层输出中嵌入 D-SPP 模块,可以获得各特征层中特征图的局部以及近似全局的感受野信息,颈部特征融合部分在融合网络浅层特征层的特征图中的目标位置信息的同时,融合了深层特征层中特征图的高语义信息,从而提高网络对目标识别定位的准确性。

3 实验与结果分析

识别检测网络在服务器上完成离线训练,服务器处理器为 Intel(R) Core(TM) i7-7820x CPU,显卡型号为 GeForce RTX 2080Ti 11GB,识别检测网络在线工作计算机处理器为 Intel(R) Core(TM) i5-8300H CPU,显卡型号为 GeForce GTX 1050Ti 4GB,采用 Python 3.6、Pytorch 1.4.0 深度学习框架完成

算法的设计。

3.1 目标检测结果

采用本文方法对马铃薯幼苗叶芽进行检测识别,并与 Faster R - CNN 网络、YOLO v3 网络以及 YOLO v4 网络进行对比分析。图 10 给出了叶芽检测与识别结果对比图。

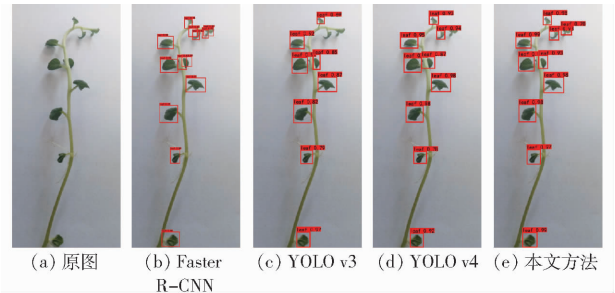


图 10 叶芽大小差异显著情况下不同网络的识别与定位结果

Fig. 10 Recognition and localization results of different networks under significant differences in leaf bud size

图 10 中,待检测的叶芽大小差异显著。Faster R - CNN 网络可以识别与定位出所有的叶芽,但存在较多重复检测的叶芽目标;YOLO v3 网络和 YOLO v4 网络可以检测出大部分叶芽位置,但右上侧的小叶芽存在漏检情况。本文所设计的识别网络能够准确识别和定位出所有叶芽位置,同时无重复检测,避免了对小叶芽漏检情况的发生。

图 11 为秆茎遮挡情况下的叶芽识别定位结果。

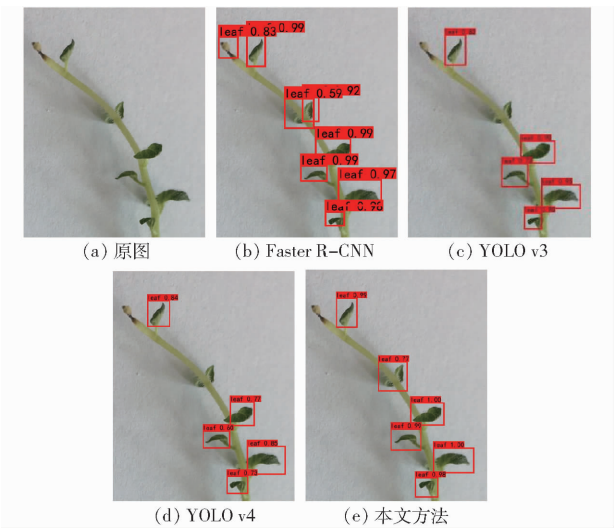


图 11 叶芽秆茎遮挡情况下不同网络的叶芽识别定位结果

Fig. 11 Recognition and localization results of different networks under leaf buds stem shading

图 11 中,部分待检测的叶芽存在严重的遮挡情况。Faster R - CNN 网络可以检测识别出所有叶芽,但却存在叶芽错检以及对遮挡叶芽重复检测的情况;YOLO v3 网络与 YOLO v4 网络对遮挡叶芽检测均存在漏检情况;本文所设计的识别网络可以正确

识别定位出所有叶芽位置,且对遮挡的叶芽检测时无重复检测和漏检情况发生,有效提高了叶芽检测的准确率。

目标检测与识别中常用的性能评价指标包括精确率 (Precision, P)、召回率 (Recall, R)、综合评价指标 F1 值以及目标类别的平均精确率 (Average precision, AP)。各网络的性能指标对比结果如表 1 所示。

表 1 检测网络的性能指标对比						
Tab. 1 Comparison of performance metrics for detection networks						
网络模型	$P/\%$	$R/\%$	F1 值/ $\%$	AP/ $\%$	单次训练 时间/s	检测时 间/s
Faster R - CNN	59.57	89.59	72	82.72	220.8	1.739
YOLO v3	87.87	80.67	84	87.22	97.2	0.429
YOLO v4	90.64	90.42	91	92.75	152.8	0.519
本文方法	95.72	94.91	95	96.03	201.2	0.655

由表 1 可见,在马铃薯幼苗叶芽识别检测中, YOLO v3、YOLO v4 网络在叶芽检测中的性能与 Faster R - CNN 网络相比有了较大的改善,但精确率和平均精确率仍较低。相比而言,本文方法的各项识别性能均为最优,其中,精确率比 Faster R - CNN 网络提高了 60.68%,比 YOLO v3 网络提高了 8.93%,比 YOLO v4 网络提高了 5.60%;召回率比 Faster R - CNN 网络提高了 5.94%,比 YOLO v3 网络提高了 17.65%,比 YOLO v4 网络提高了 4.97%;F1 值比 Faster R - CNN 网络提高了 31.94%,比 YOLO v3 网络提高了 13.10%,比 YOLO v4 网络提高了 4.40%;平均精确率比 Faster R - CNN 网络提高了 16.09%,比 YOLO v3 网络提高了 10.10%,比 YOLO v4 网络提高了 3.54%。

本文方法的单次训练时间和单帧图像检测时间略高于 YOLO v4、YOLO v3 网络。由于识别网络采用离线训练、在线检测的方式,因此训练时间的长短不影响网络的应用性能,而本文方法单帧图像的检测时间小于 1 s,幼苗剪切系统允许的识别时间为 1 s,因此,网络的检测速度仍能满足幼苗在线剪切工艺的要求。

图 12 为不同识别网络的损失值变化曲线。图 12 中, YOLO v3 和 YOLO v4 网络在前 20 个迭代周期损失值下降迅速,而在 20 个迭代周期之后损失值逐渐趋于收敛,本文方法在前 15 个迭代周期损失值下降迅速, Faster R - CNN 网络的损失值在 0 ~ 6 之间变化,其放大图如图 13 所示。

由图 13 可知, Faster R - CNN 网络的损失值在训练过程的前 50 个迭代周期内下降迅速, 50 个迭

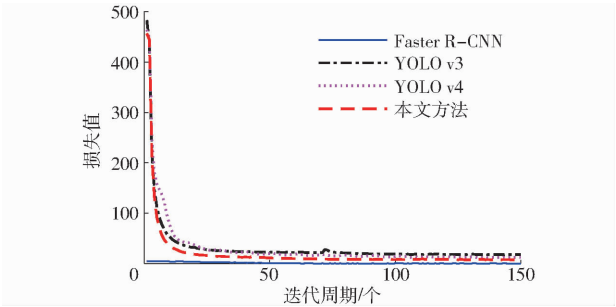


图 12 4 种网络的损失值变化曲线

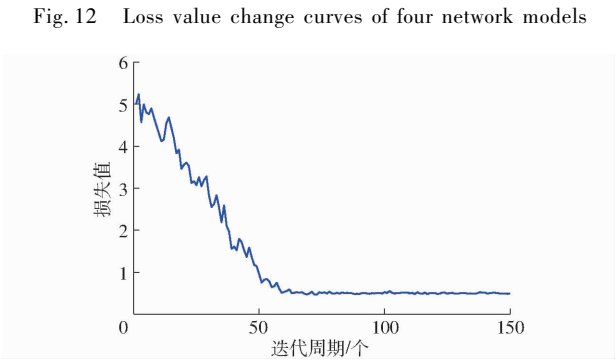


图 13 Faster R-CNN 网络的损失值变化曲线

Fig. 13 Loss value change curve of Faster R-CNN network

代周期后损失值趋于收敛。

由上述对比可见,本文识别网络的损失值可在更少的训练周期内趋于收敛。图 14 为各检测方法的精确率-召回率曲线。

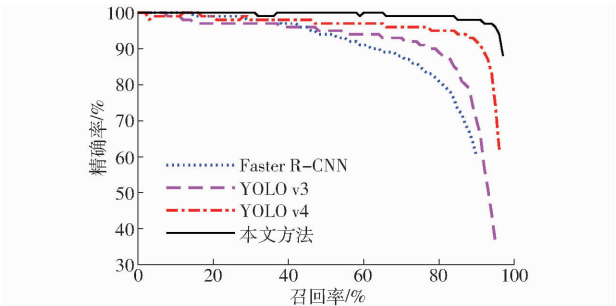


图 14 不同网络的精确率-召回率曲线

图 14 中曲线与坐标轴所围面积为平均精确率。平均精确率越高,网络对目标类别的检测性能越好,本文网络的平均精确率大于其它 3 种网络,表明本文网络对马铃薯叶芽检测的性能优于其它 3 种网络。

由此可见,本文网络利用基于深度可分离卷积的 Res2Net 模块替换特征提取部分中的 Residual Block 结构,同时在特征提取部分的 3 个特征层输出中嵌入 D-SPP 模块,能够有效提升网络识别与定位精度。

3.2 消融实验

由于本文方法在 YOLO v4 的基础上采用了多个改进策略,为了验证所采用的各项改进策略的有

效性,设计了消融实验进行对比研究。

模型 1 为 YOLO v4 网络,模型 2 为将 YOLO v4 特征提取部分的 Residual Block 结构替换为基于深度可分离卷积的 Res2Net 模块后得到的网络。模型 3 为将 YOLO v4 的 13×13 特征图所在特征层输出中的 SPP 模块替换为 D-SPP 模块,同时在 26×26 和 52×52 特征图所在特征层输出中嵌入 D-SPP 模块后得到的网络。模型 4 为本文网络。各网络模型均训练 150 个周期,所得到的损失值变化曲线如图 15 所示。

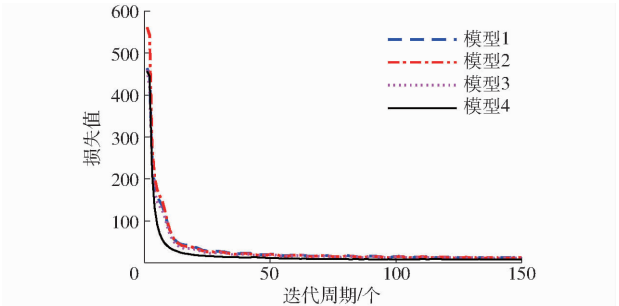


图 15 消融实验损失值变化曲线

Fig. 15 Loss value change curve of ablation experiment

图 15 中,消融实验各网络模型在训练过程的前 20 个迭代周期内损失值快速下降,之后逐渐趋于收敛。消融实验的性能对比结果如表 2 所示,表中“√”表示在 YOLO v4 网络中添加该模块。

表 2 消融实验结果						
Tab. 2 Results of ablation experiment						
网络模型	Res2Net	D-SPP	P/%	R/%	F1 值/%	AP/%
模型 1			90.64	90.42	91	92.75
模型 2	√		92.04	92.40	92	94.30
模型 3		√	92.38	93.54	93	94.25
模型 4	√	√	95.72	94.91	95	96.03

由表 2 可见,将特征提取部分中的 Residual Block 结构替换为基于深度可分离卷积的 Res2Net 模块,在细粒度层次上表示特征信息,能更好地提取叶芽的细小特征,精确率较 YOLO v4 网络提高了 1.54%,召回率较 YOLO v4 网络提高了 2.19%,F1 值较 YOLO v4 网络提高了 1.10%,平均精确率较 YOLO v4 网络提高了 1.67%,有效改善了网络对于小目标的检测性能。在 3 个特征层输出中嵌入和替换 D-SPP 模块,融合了局部以及近似全局的感受野的目标位置信息,对叶芽的识别定位更加准确,精确率较 YOLO v4 网络提高了 1.92%,召回率较 YOLO v4 网络提高了 3.45%,F1 值较 YOLO v4 网络提高了 2.20%,平均精确率较 YOLO v4 网络提高了 1.62%,增加了网络对马铃薯叶芽定位的准确性。将上述两种改进方法同时融合到基准网络中,

能够进一步提升马铃薯叶芽的识别与定位结果的准确性,精确率较 YOLO v4 网络提高了 5.60%,召回率较 YOLO v4 网络提高了 4.97%,F1 值较 YOLO v4 网络提高了 4.40%,平均精确率较 YOLO v4 网络提升了 3.54%,由此验证了本文改进策略的有效性。

4 结 论

(1)针对马铃薯幼苗叶芽形态各异、大小不一,存在叶茎遮挡与覆盖等目标检测难题,结合基于深度可分离卷积的 Res2Net 模块以及基于扩张卷积的 D-SPP 模块,构建了以 YOLO v4 网络为基础网络

的马铃薯幼苗叶芽检测模型,用于实现马铃薯幼苗叶芽的有效识别,并采用消融实验完成了改进策略的有效性验证分析。

(2)实验结果表明,本文设计的识别网络对马铃薯叶芽检测的精确率为 95.72%,召回率为 94.91%,F1 值为 95%,平均精确率为 96.03%。与 Faster R-CNN、YOLO v3、YOLO v4 网络相比,本文方法对马铃薯叶芽检测具有更好的识别和定位性能。

(3)本文所设计的网络采用离线训练、在线检测的工作方式,检测识别速度满足马铃薯自动化育苗剪切系统的工艺要求,能够有效提高马铃薯育苗生产的工作效率。

参 考 文 献

[1] 刘慧力,贾洪雷,王刚,等. 基于深度学习与图像处理的玉米茎秆识别方法与试验[J]. 农业机械学报, 2020, 51(4): 207-215.
LIU Huili, JIA Honglei, WANG Gang, et al. Method and experiment of maize(*Zea Mays* L.) stems recognition based on deep learning and image processing[J]. Transactions of the Chinese Society for Agricultural Machinery, 2020, 51(4):207-215. (in Chinese)

[2] JIN Y. Recognition technology of agricultural picking robot based on image detection technology[J]. INMA Bucharest, 2020, 62(3):191-200.

[3] DENG X, ZHOU B, HOU Y. A reliability test method for agricultural paddy field intelligent robot[J]. INMA Bucharest, 2021, 63(1):271-280.

[4] 徐岩,李晓振,吴作宏,等. 基于残差注意力网络的马铃薯叶部病害识别[J]. 山东科技大学学报(自然科学版), 2021, 40(2):76-83.
XU Yan, LI Xiaozhen, WU Zuohong, et al. Potato leaf disease recognition via residual attention network[J]. Journal of Shandong University of Science and Technology(Natural Science), 2021, 40(2):76-83. (in Chinese)

[5] GAO J, WESTERGAARD J C, SUNDMARK E, et al. Automatic late blight lesion recognition and severity quantification based on field imagery of diverse potato genotypes by deep learning[J]. Knowledge-Based Systems, 2021, 214:106723.

[6] TIWARI D, AAHIAH M, GANGWAR N, et al. Potato leaf diseases detection using deep learning[C]//2020 4th International Conference on Intelligent Computing and Control Systems (ICICCS), 2020:461-466.

[7] GAO C, WU Y, HAO X. Hierarchical suppression based matched filter for hyperspectral imagery target detection[J]. Sensors, 2020, 21(1):144.

[8] GU M, LIU H, WANG Y, et al. PolSAR target detection via reflection symmetry and a wishart classifier[J]. IEEE Access, 2020, 8:103317-103326.

[9] ZHOU C, YANG G, LIANG D, et al. An integrated skeleton extraction and pruning method for spatial recognition of maize seedlings in MGVI and UAV remote images[J]. IEEE Transactions on Geoscience and Remote Sensing, 2018, 56(8): 4618-4632.

[10] 崔家华,张云洲,王争,等. 面向嵌入式平台的轻量级目标检测网络[J]. 光学学报, 2019, 39(4):307-313.
CUI Jiahua, ZHANG Yunzhou, WANG Zheng, et al. Light-weight object detection networks for embedded platform[J]. Acta Optica Sinica, 2019, 39(4):307-313. (in Chinese)

[11] LI W Y, LI J Y, SARMA K V, et al. Path R-CNN for prostate cancer diagnosis and gleason grading of histological images [J]. IEEE Transactions on Medical Imaging, 2019, 38(4):945-954.

[12] LEE H, EUM S, KWON H. ME R-CNN: multi-expert R-CNN for object detection[J]. IEEE Transactions on Image Processing, 2020, 29:1030-1044.

[13] LI J, LIANG X, SHEN S, et al. Scale-aware Fast R-CNN for pedestrian detection[J]. IEEE Transactions on Multimedia, 2018, 20(4):985-996.

[14] SHIH K H, CHIU C T, LIN J, et al. Real-time object detection with reduced region proposal network via multi-feature concatenation[J]. IEEE Transactions on Neural Networks and Learning Systems, 2020, 31(6):2164-2173.

[15] CAO C, WANG B, ZHANG W, et al. An improved Faster R-CNN for small object detection[J]. IEEE Access, 2019, 7: 106838-106846.

[16] 刘召,张黎明,耿美晓,等. 基于改进的 Faster R-CNN 高压线缆目标检测方法[J]. 智能系统学报, 2019, 14(4):627-634.
LIU Zhao, ZHANG Liming, GENG Meixiao, et al. Object detection of high-voltage cable based on improved Faster R-CNN

[J]. CAAI Transactions on Intelligent Systems, 2019, 14(4):627 – 634. (in Chinese)

[17] ZHAI S, SHANG D, WANG S, et al. DF – SSD: an improved SSD object detection algorithm based on denseNet and feature fusion[J]. IEEE Access, 2020, 8:24344 – 24357.

[18] YU Y, HAN X, DU L. Target part detection based on improved SSD algorithm[J]. Journal of Physics: Conference Series, 2020, 1486(3):32024 – 32029.

[19] 马王鹏, 王晓东, 刘冬. 探地雷达图像异常识别中 YOLO 目标检测算法的研究[J]. 测绘通报, 2019, 4(增刊 1):72 – 76.
MA Wangpeng, WANG Xiaodong, LIU Dong. YOLO target detection algorithm research in GPR anomaly image recognition [J]. Bulletin of Surveying and Mapping, 2019, 4(Supp. 1):72 – 76. (in Chinese)

[20] BARREIROS M, DANTAS D, SILVA L, et al. Zebrafish tracking using YOLO v2 and Kalman filter[J]. Scientific Reports, 2021, 11(1):3219.

[21] 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)

[22] DU S, ZHANG P, ZHANG B, et al. Weak and occluded vehicle detection in complex infrared environment based on improved YOLO v4[J]. IEEE Access, 2021, 9:25671 – 25680.

[23] 彭红星, 黄博, 邵园园, 等. 自然环境下多类水果采摘目标识别的通用改进 SSD 模型[J]. 农业工程学报, 2018, 34(16):155 – 162.
PENG Hongxing, HUANG Bo, SHAO Yuanyuan, et al. General improved SSD model for picking object recognition of multiple fruits in natural environment[J]. Transactions of the CSAE, 2018, 34(16): 155 – 162. (in Chinese)

[24] TIAN Y, YANG G, WANG Z, et al. Apple detection during different growth stages in orchards using the improved YOLO v3 model[J]. Computers & Electronics in Agriculture, 2019, 157: 417 – 426.

[25] LIU T, PANG B, AI S, et al. Study on visual detection algorithm of sea surface targets based on improved YOLO v3[J]. Sensors, 2020, 20(24):7263.

[26] 方鹏, 郝宏运, 王红英. 基于知识蒸馏的叠层笼养蛋鸡行为识别模型研究[J]. 农业机械学报, 2021, 52(10):300 – 306.
FANG Peng, HAO Hongyun, WANG Hongying. Behavior recognition model of stacked-cage layers based on knowledge distillation[J]. Transactions of the Chinese Society for Agricultural Machinery, 2021, 52(10):300 – 306. (in Chinese)

[27] GAO S, CHENG M M, ZHAO K, et al. Res2Net: a new multi-scale backbone architecture[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2021, 43(2):652 – 662.

[28] HUNG K W, ZHANG Z, JIANG J. Real-time image Super-Resolution using recursive depthwise separable convolution network [J]. IEEE Access, 2019, 7:99804 – 99816.

[29] HUANG Z, WANG J, FU X, et al. DC – SPP – YOLO: dense connection and spatial pyramid pooling based YOLO for object detection[J]. Information Sciences, 2020, 522:241 – 258.

[30] 许腾, 唐贵进, 刘清萍, 等. 基于空洞卷积和 Focal Loss 的改进 YOLO v3 算法[J]. 南京邮电大学学报(自然科学版), 2020, 40(6):100 – 108.
XU Teng, TANG Guijin, LIU Qingping, et al. Improved YOLO v3 based on dilated convolution and Focal Loss[J]. Journal of Nanjing University of Posts and Telecommunications (Natural Science Edition), 2020, 40(6):100 – 108. (in Chinese)