doi:10.6041/j.issn.1000-1298.2021.10.029

基于轻量二阶段检测模型的自然环境多类蔬菜幼苗识别

孟庆宽[1,2]　张　漫[3]　叶剑华[1]　都泽鑫[1]　宋名果[1]　张志鹏[1]

(1. 天津职业技术师范大学自动化与电气工程学院, 天津 300222; 2. 天津市信息传感与智能控制重点实验室, 天津 300222;
3. 中国农业大学现代精细农业系统集成研究教育部重点实验室, 北京 100083)

摘要: 为实现自然环境下蔬菜幼苗精准快速识别,本文以豆角、花菜、白菜、茄子、辣椒、黄瓜等形态差异大、具有代表性的蔬菜幼苗为研究对象,提出一种基于轻量化二阶段检测模型的多类蔬菜幼苗检测方法。模型采用混合深度分离卷积作为前置基础网络对输入图像进行运算,以提高图像特征提取速度与效率;在此基础上,引入特征金字塔网络(Feature pyramid networks, FPN)单元融合特征提取网络不同层级特征信息,用于增强深度学习检测模型对多尺度目标的检测精度;然后,通过压缩检测头网络通道维数和全连接层数量,降低模型参数规模与计算复杂度;最后,将距离交并比(Distance - IoU, DIoU)损失作为目标边框回归损失函数,使预测框位置回归更加准确。试验结果表明,本文提出的深度学习推理模型对多类蔬菜幼苗的平均精度均值为 97.47%,识别速度为 19.07 f/s,模型占用存储空间为 60 MB,对小目标作物以及叶片遮挡作物的平均精度均值达到 88.55%,相比于 Faster R - CNN、R - FCN 模型具有良好的泛化性能和鲁棒性,可以为蔬菜田间农业智能装备精准作业所涉及的蔬菜幼苗检测识别问题提供新方案。

关键词: 蔬菜幼苗; 深度学习; 作物识别; 轻量卷积; 二阶段检测模型

中图分类号: TP391.41　　**文献标识码**: A　　**文章编号**: 1000-1298(2021)10-0282-09　　**OSID**:

Identification of Multiple Vegetable Seedlings Based on Two-stage Lightweight Detection Model

MENG Qingkuan[1,2]　ZHANG Man[3]　YE Jianhua[1]　DU Zexin[1]　SONG Mingguo[1]　ZHANG Zhipeng[1]

(1. *College of Automation and Electrical Engineering*, *Tianjin University of Technology and Education*, *Tianjin* 300222, *China*
2. *Tianjin Key Laboratory of Information Sensing and Intelligent Control*, *Tianjin* 300222, *China*
3. *Key Laboratory of Modern Precision Agriculture System Integration Research*, *Ministry of Education*, *China Agricultural University*, *Beijing* 100083, *China*)

Abstract: The identification of vegetable seedlings offers many useful applications in precision agriculture, such as automated weeding, variable rate fertilization and precise spraying of diseased plants. Aiming to recognize vegetable seedlings accurately and rapidly in natural environment, multiple kinds of vegetable seedlings were taken as study object, and a two-stage based lightweight detection model was proposed. In order to improve the speed and efficiency of image feature extraction, a mixed depth wise convolution that naturally mixed up multiple kernel sizes in a single convolution was applied as backbone network to process input images. Moreover, the feature pyramid networks (FPN) was introduced to integrate different feature maps of backbone network with the aim of improving the identification accuracy of deep learning detection model for multi-scale targets. By minimizing network channel dimensions and decreasing the number of full connection layers in detection head, the two-stage based detection model parameters and computational complexity were greatly reduced. In addition, a distance-IoU (DIoU) loss was proposed for bounding box regression to make the predicted box match with the target box perfectly. Experimental results showed that the mean average precision and recognition speed of multiple kinds of vegetable seedlings based on the proposed model were 97.47% and 19.07 f/s, respectively, and model

收稿日期: 2020-10-30　修回日期: 2020-12-20

基金项目: 国家自然科学基金项目(31971786)、天津市自然科学基金项目(18JCQNJC04500)、天津市教委科研计划项目(JWK1613、JWK1604)和天津职业技术师范大学校级预研项目(KJ2009、KYQD1706)

作者简介: 孟庆宽(1983—),男,讲师,博士,主要从事农业自动化图像信息处理研究,E-mail: 373414672@qq.com

通信作者: 张漫(1975—),女,教授,博士生导师,主要从事农业电气化与自动化研究,E-mail: cauzm@cau.edu.cn

size was 60 MB. The average accuracy of detection model can obtain 88.55%, when a crop size was less than 32 pixels × 32 pixels or leaves occlusion occurred. It was demonstrated that the two-stage based lightweight detection model havd good generalization and robustness performance compared with that by other models, such as Faster R - CNN and R - FCN. The approach presented obtained high accurate rate and fast inference speed in the recognition of vegetable seedlings, which opened a new journey for the research of vegetable detection in precision agriculture.

Key words: vegetable seedlings; deep learning; crop identification; lightweight convolution; two-stage detection model

0　引言

蔬菜包含丰富的维生素、矿物质和膳食纤维，是维持人体营养平衡，保持身体健康的重要食物之一[1-2]。近些年，我国蔬菜种植面积稳定在 2 000 万 hm^2 左右，年产量达 7 亿 t，已经超过粮食产量成为第一大农产品[3]。蔬菜产业的迅速发展满足了人们日常生活需要，但在种植过程中也存在施肥过量、农药使用超标等问题，对生态环境和人体健康产生不利影响。随着电子技术和计算机技术的发展，自动化智能农业装备逐渐应用到农业生产中，通过靶向喷药、变量施肥、机械除草等一系列手段提高了蔬菜作物的产量与安全品质[4-6]。

目前，大部分智能农业装备基于机器视觉获取作物图像，快速准确地识别作物种类并进行定位是实现智能农业装备高效精准作业的先决条件[7-8]。在作物识别定位方面，国内外学者开展了一系列研究[9-14]。这些方法主要基于作物颜色、形状、纹理、光谱和位置等特征中的一种或多种的组合实现作物识别定位，但在实际应用过程中只能针对特定环境下作物进行检测，并且容易受到自然光照、背景噪声和枝叶遮挡等因素影响导致识别准确率降低。

相比于常规方法，卷积神经网络在近些年得到迅速发展，其通过卷积层、池化层、全连接层对输入图像由浅入深地提取不同层次特征，经过信息分类和位置回归，实现对目标的准确检测[15-19]。这些基于深度学习技术的目标检测方法按照检测所需要的步骤可以分为一阶段目标检测法和二阶段目标检测法。一阶段检测法以 SSD 和 YOLO 系列算法为代表，主要思想是对图像不同位置进行均匀密集采样产生不同尺度和长宽比的候选框，通过卷积神经网络提取特征后直接对候选框进行分类回归，此类方法的特点是检测速度快，但检测精度有限。二阶段检测法先利用区域建议网络（Region proposal network，RPN）生成一系列稀疏的预测框，然后将特征图与预测框综合输入到全连接层完成分类回归，代表算法包括 Faster R - CNN 与 F - RCN，此类方法的特点是精度高，但消耗时间较长。

复杂农业环境中作物识别的精准性与实时性是影响农业智能机械高效作业的重要因素之一。本文以自然环境下多类蔬菜幼苗为研究对象，提出一种轻量化二阶段目标检测模型，以期在具有较高识别精度的基础上，提高模型的推理速度。目标检测模型的设计思路为：先采用混合深度分离卷积作为前置基础网络对输入图像进行运算，以提高图像特征提取速度与效率；在此基础上融入特征金字塔网络，实现浅层特征与深层特征的融合，加强对多尺度目标的检测精度；最后将检测头单元进行轻量化设计，减少冗余参数，降低模型计算量，并引入距离交并比（Distance - IoU，DIoU）损失作为目标边框回归损失函数，使位置回归更加准确快速。

1　轻量化蔬菜幼苗检测模型设计

二阶段目标检测模型主要由前置基础卷积网络、区域生成网络和检测头单元组成。前置基础网络用于提取图像不同层次特征信息；区域生成网络的功能是生成候选区域，预测候选区域样本属性（正样本或负样本），对正样本候选区域初步进行边框回归；检测头单元由感兴趣区域池化层（Region of interest pooling，ROI pooling）和区域卷积神经网络子网（Region based convolutional neural network subnet，R - CNN subnet）构成，主要负责预测池化后候选区域的目标类别并再次进行边框回归以得到预测框精确位置。二阶段目标检测模型主要计算量集中在前置基础网络和检测头单元，通过对这两部分网络单元进行轻量化设计，能够有效降低模型参数规模和计算复杂度。

1.1　轻量化前置基础网络设计

轻量深度卷积神经网络采用深度可分离卷积代替标准卷积运算单元从而获得更高的计算效率和运算速度。常规深度卷积网络模型利用单一尺寸卷积核获取图像信息，特征像素对应的图像分辨率固定，无法获取不同尺度特征，导致模型在准确率和效率方面难以达到均衡。卷积运算中卷积核负责图像局部特征提取，大尺寸卷积核包含较大的感受野，能捕获更多的特征信息，提高模型识别准确率；小尺寸卷

积核包含参数少,具有较高的计算效率,可以降低模型运算量和存储空间。基于这个特点,将多个不同尺寸卷积核融合到一个单独的深度可分离卷积运算中,构成混合深度分离卷积神经网络,使计算得到的特征图像素具有不同感受野,能对应多种图像分辨率模式,获得多尺度特征信息,从而有效提高特征提取精度和效率[20]。混合深度分离卷积神经网络的设计思路为:设输入特征图为 $X_{(h,w,c)}$,h 表示特征图高度,w 表示宽度,c 表示通道数,沿通道方向将特征图划分为 g 组子特征图($X_{(h,w,c_1)}$,$X_{(h,w,c_2)}$,…,$X_{(h,w,c_g)}$),$c_s(s=1,2,\cdots,g)$ 表示第 s 组子特征图通道数,$c_1+c_2+\cdots+c_g=c$。建立 g 组不同尺寸深度卷积核($W_{(k_1,k_1,c_1,m)}$,$W_{(k_2,k_2,c_2,m)}$,…,$W_{(k_g,k_g,c_g,m)}$),m 表示通道乘数,$k_t\times k_t(t=1,2,\cdots,g)$ 表示第 t 组卷积核尺寸。第 t 组输入子特征图与对应深度卷积核进行运算得到第 t 组输出子特征图,具体定义为

$$Y^t_{x,y,z_t}=\sum_{i=0}^{h-1}\sum_{j=0}^{w-1}X^t_{x+i,y+j,c_t}W^t_{i,j,c_t,m} \tag{1}$$

式中 x——特征图像素行号

y——特征图像素列号

c_t——第 t 组输入子特征图通道数

z_t——第 t 组输出子特征图通道数

i——卷积核元素行号

j——卷积核元素列号

Y^t_{x,y,z_t}——第 t 组输出子特征图

X^t_{x,y,c_t}——第 t 组输入子特征图

$W^t_{i,j,c_t,m}$——第 t 组深度卷积核

根据式(1)计算结果,采用 Contact 运算将所有子特征图在通道维度上以加法形式进行拼接得到最终输出特征图,计算公式为

$$Y_{x,y,z}=\text{Contact}(Y^1_{x,y,z_1},Y^2_{x,y,z_1},\cdots,Y^g_{x,y,z_g}) \tag{2}$$

其中 $z=z_1+z_2+\cdots+z_g$

式中 Contact——特征图通道维数拼接运算函数

z——输出特征图通道数

$Y_{x,y,z}$——拼接后输出特征图

输出特征图尺寸与输入子特征图相同,通道维数等于参与运算子特征图通道维数之和。

本文混合深度分离卷积神经网络示意图如图 1 所示,特征图最大分组数 $g=5$,每组具有相同的通道数,对应的深度卷积核尺寸为{3×3,5×5,7×7,9×9,11×11},特征图分组后与不同尺寸卷积核进行运算,然后对结果进行拼接得到输出。图 1 卷积神经网络运算过程中根据特征图尺寸划分为 5 个阶段(Stage),相同尺寸特征图属于同一阶段,相邻阶段特征图尺寸比为 2。

1.2 特征金字塔网络融合设计

Faster R-CNN、R-FCN 检测模型将特征提取网络最后一阶段的高层次特征图输入到 RPN 网络,提高了模型训练与推理速度,不过高层特征图分辨率低,无法有效表征不同尺度物体,尤其对小尺寸物体的检测存在一定局限性。文献[21]利用卷积神经网络中所包含的多尺度层级结构设计了特征金字塔网络,通过融合浅层特征图的高分辨率特征和深层特征图的丰富语义信息,提高了深度推理模型对多尺度目标的检测精度。为实现多类别蔬菜幼苗的准确分类和精确定位,本文在混合深度分离卷积神经网络的基础上融入特征金字塔网络,如图 2 所示。

图 2 中混合深度分离卷积按照自底向上的顺序依次产生不同阶段的特征图,阶段 $x/y(x=1,2,\cdots,$

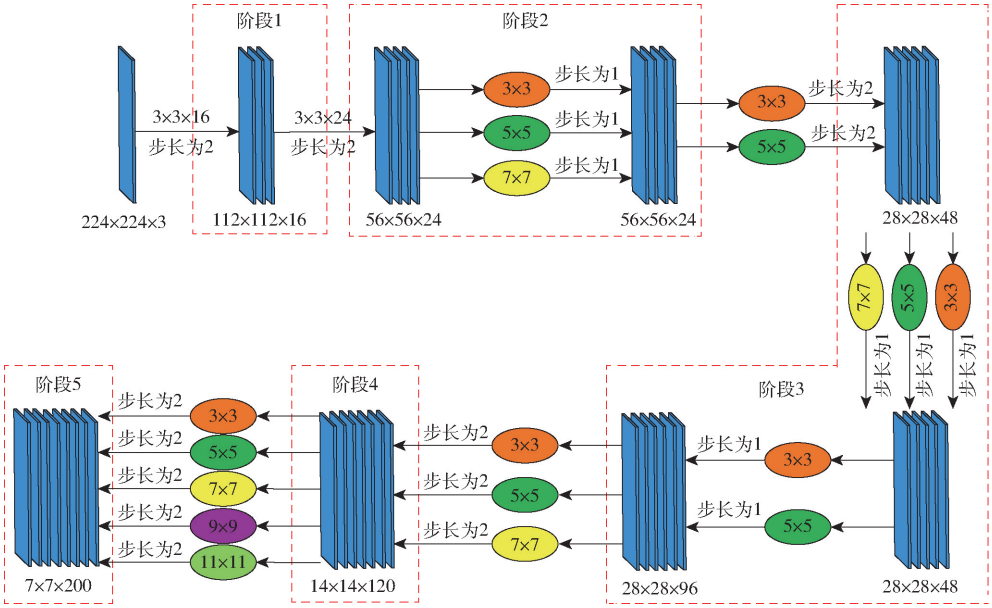


图 1 混合深度分离卷积神经网络结构示意图

Fig. 1 Structure schematic of mixed depthwise convolutional neural network

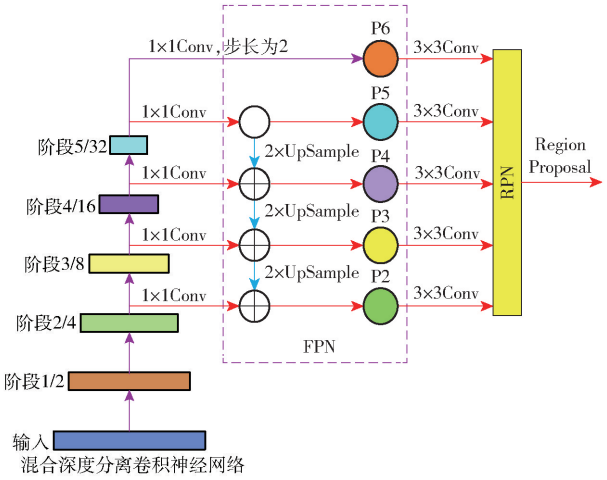


图 2　混合深度分离卷积融合特征金字塔网络图

Fig. 2　Diagram of mixed depthwise convolutional fusion with feature pyramid networks

5；$y=2,4,\cdots,32$）中 x 表示特征图所处的阶段数，y 表示本阶段特征图尺寸相对于输入图像缩小的倍数。阶段 2～5 经过 1×1 卷积运算后分别输入到 FPN 单元，其中 1×1 卷积的作用是使各阶段特征图输入到 FPN 的通道数保持一致。FPN 单元按照自顶向下的顺序对输入的高层次特征图进行上采样扩大分辨率，然后通过相加的方式与相邻低层次特征进行融合。融合后的特征图一方面输入到后续网络进行预测推理，另一方面继续通过上采样与下层特征图进行融合。混合深度分离卷积阶段 2～5 对应 FPN 的 P2～P5 层级，P6 为阶段 5 下采样得到，用于 RPN 网络中产生预测框，不参与融合操作。{P2，P3，P4，P5，P6} 中每个层级负责单一尺度的信息处理，分别对应 $\{16^2,32^2,64^2,128^2,256^2\}$ 5 种尺度预测框，每种预测框具有{1∶1，1∶2，2∶1} 3 种长宽比例，共 15 种预测框对目标对象与背景进行预测。

1.3　检测头网络轻量化设计

在检测头部分，Faster R－CNN 在 ROI pooling 之后连接 2 个包含大量参数的全连接层，候选区域在全连接层上的参数计算不共享，当候选区域数量较多且存在大量重叠时导致计算速度降低；R－FCN 去掉了全连接层以减少冗余参数，在前置基础卷积网络之后生成位置敏感分数图，候选区域基于位置敏感分数图进行池化运算实现参数共享，不过位置敏感分数图的通道数多达 3 696 个，计算量仍然较大。综上所述，二阶段目标检测模型含有一个比较重的检测头，大量计算集中在检测头部分，严重影响了模型运行速度[22]。

为克服二阶段目标检测框架检测头较重的问题，本文在 Faster R－CNN 和 R－FCN 模型的基础上，通过压缩网络通道维数和参数规模构建轻量化检测头单元，具体设计方法为：针对混合深度分离卷积网络输出特征图采用大尺寸非对称卷积生成 $\alpha\times p\times p$ 通道的特征图，α 是与类别无关且数值较小的数，取值为 10，$p\times p$ 等于候选区域池化后的网格数量，取值为 49，经过计算得到 490 通道的特征图，相比于 R－FCN 特征图通道维数显著降低；然后，引入 ROI Align 运算将不同尺寸候选区域所对应的特征信息池化生成固定尺寸特征图，ROI Align 运算利用双线性差值法获取坐标为浮点数的像素点数值，将整个特征聚集过程转化为一个连续的操作，相比于 ROI pooling 操作，消除了 2 次整型量化带来的计算误差，可以提高目标检测精度；最后，接入 1 个全连接层获得检测目标全局特征，并基于 2 个并行分支完成目标分类和位置预测。本文使用的大尺寸非对称卷积的卷积核为 1×15 和 15×1，以降低模型参数计算量，同时大尺寸卷积核具有较大感受野，能增强模型特征提取能力。图 3 为轻量化二阶段目标检测模型原理框图。

1.4　损失函数

二阶段目标检测框架损失函数为多任务损失，由分类损失与位置回归损失构成。分类损失采用交叉熵函数计算预测类别与真实类别的误差；位置回归损失采用 Smooth L1 函数计算预测框与真实框之间位置坐标误差，通过多次迭代缩小误差范围，使预测框与真实框具有较大程度的重叠。交并比（Intersection over union，IoU）表示预测框与真实框的交叠程度，是目标检测领域衡量预测框准确度的

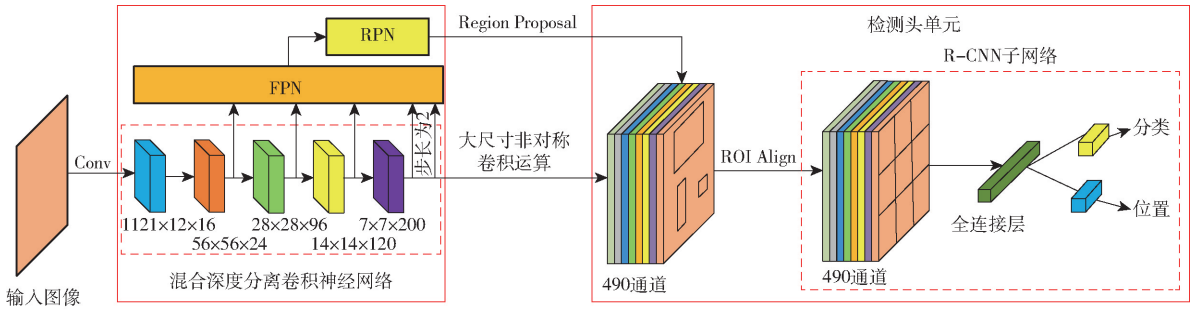


图 3　轻量化二阶段目标检测模型原理框图

Fig. 3　Diagram of two-stage based objection detection model with lightweight

指标,可以进行反向传播,能够作为目标函数进行优化。一些检测框架结合 IoU 优化损失函数取得良好效果,但当预测框与真实框没有交叠时,IoU 为 0,损失函数也为 0,没有梯度回传,无法进行学习[23-24]。针对 IoU 损失函数存在的不足,文献[25]定义了 DIoU 损失函数,DIoU 对预测框与真实框的距离、重叠率和尺度等因素建立关联,直接最小化二者的中心点距离,即使预测框与真实框不存在交叠,也可以为边框提供移动方向,使位置回归更加快速准确。DIoU 损失函数定义为

$$L_{diou}(A,B)=1-I_{iou}(A,B)+\frac{\rho^2(A_{ctr},B_{ctr})}{d^2} \tag{3}$$

其中

$$I_{iou}(A,B)=\frac{A\cap B}{A\cup B} \tag{4}$$

式中 L_{diou}——DIoU 损失函数

A——预测框 B——真实框

d——A、B 最小包围框对角线长度

$\rho(\cdot)$——欧氏距离函数

A_{ctr}——预测框中心点坐标

B_{ctr}——真实框中心点坐标

$I_{iou}(A,B)$——预测框与真实框交并比

本文引入 DIoU 作为位置回归损失函数,构建检测模型损失函数。损失函数包括 RPN 损失函数与检测头损失函数两部分,每部分又由分类损失函数与位置回归损失函数构成。损失函数计算公式为

$$L_{total}=L_{rpn}(p_l,a_l)+L_{head}(p,u,o,s) \tag{5}$$

其中

$$L_{rpn}(p_l,a_l)=\frac{1}{N_{cls}}\sum_l L_{cls}(p_l,p_l^*)+\lambda\frac{1}{N_{reg}}\sum_l p_l^* L_{diou}(a_l,s_l^*) \tag{6}$$

$$L_{head}(p,u,o,s)=L_{cls}(p,u)+\lambda'[u\geqslant 1]L_{diou}(o,s) \tag{7}$$

式中 L_{total}——目标检测模型损失函数

L_{rpn}——RPN 网络损失函数

L_{head}——检测头网络损失函数

l——锚框索引 $[\cdot]$——回归计算

p_l——第 l 个锚框二分类预测概率

p_l^*——第 l 个锚框判别值

a_l——第 l 个锚框对应的预测框

s_l^*——第 l 个锚框对应的真实框

p——预测类别概率

u——真实类别标签值

λ、λ'——权重参数

L_{cls}——分类损失函数

N_{cls}——采样的锚框数量

N_{reg}——采样正负样本数

o——RPN 网络输出的预测框

s——与预测框对应的真实框

2 蔬菜幼苗检测试验

2.1 试验平台

试验运行平台为台式计算机,硬件配置和软件环境为:采用 AMD Ryzen 73700CPU,主频 2.4 GHz,运行内存 16 GB,GPU 型号为 GeForec RTX2070,显存为 8 GB,统一计算设备架构 CUDA 版本为 10.0,深度神经网络加速库 CUDNN 版本为 7.4,深度学习框架选择 TensorFlow,在 Windows 10 操作系统上基于 Python 语言进行编程设计。

2.2 数据采集与处理

试验图像采集于天津市宁河农场,获取时间为 2020 年 4—6 月,蔬菜幼苗包括白菜、花菜、茄子、黄瓜、辣椒、豆角,这 6 种蔬菜形态差异大,能够充分反映控制模型泛化能力与通用性。为保证数据样本的多样性,在不同天气条件、不同光照方向、不同土壤背景情况下进行拍摄,共获得 3 600 幅图像,图像尺寸为 480 像素×720 像素,存储格式为 JPG。采集的图像通过几何变换与颜色变换进行数据增强,以提高训练模型的泛化性能。数据增强后的图像共 7 920 幅,包含白菜图像 1 356 幅、花菜图像 1 394 幅、茄子图像 1 362 幅、黄瓜图像 1 264 幅、辣椒图像 1 232 幅、豆角图像 1 312 幅。按照 8∶1∶1将增强数据集划分为训练集、验证集和测试集,3 种数据集独立互斥,分别用于目标检测模型的训练、参数优化和性能评估。

2.3 模型训练

为提高目标检测模型训练速度和收敛性能,将混合深度分离卷积模型在本文硬件平台上基于 ImageNet 数据集进行预训练。首先,对 ImageNet 数据集进行预处理,将训练集和验证集数据转换为 TFRecord 格式;然后,基于 TFRecord 数据文件开始训练,训练时批尺寸为 64,训练图像 1 280 000 幅,图像尺寸缩放为 224 像素×224 像素,训练周期为 80,每个周期迭代次数为 20 000,总迭代次数 1 600 000,动量因子为 0.9,权值衰减系数为 1×10^{-4},初始学习率为 0.01,利用分段常数对学习率进行衰减,最终学习率衰减到 0.000 01;最后,经过多次迭代训练得到混合深度分离卷积预训练模型。

在此基础上,采用随机梯度下降法将目标检测模型部署在显卡上以端对端的联合方式进行训练,前置特征提取网络使用预训练好的混合深度分离卷积权重初始化,其余层用均值为 0、标准差为 0.01 的高斯分布随机初始化。为提高模型的训练效果,通过在线困难样本挖掘技术(Online hard example

mining,OHEM)强化对难分样本的训练[26]。网络训练过程中,训练集图像为 5 544 幅,批尺寸为 8,训练周期为 100,1 个周期迭代次数为 693,动量因子为 0.9,权值衰减系数为 5×10^{-4},初始学习率为 0.002,衰减率为 0.9,每迭代 1 个周期保存一次模型,并在验证集上测试模型的准确率。

2.4　评价指标

为客观评价多类蔬菜幼苗识别模型的泛化能力,选择平均精度(Average precision, AP)、平均精度均值(Mean average precision, MAP)、检测速度、网络规模作为模型性能度量指标。本文应用场景中,设蔬菜幼苗为正样本,其他类物体和背景为负样本,将模型正确检测出的某类蔬菜幼苗数量和预测为此类蔬菜幼苗数量的比值定义为精确率(Precision,P),用于衡量对正样本的识别能力;将测试集数据中某类蔬菜幼苗被正确预测数量与此类蔬菜幼苗总数的比值定义为召回率(Recall,R),用于衡量对正样本的覆盖能力;平均精度与精确率、召回率有关,为精确率-召回率曲线与坐标轴的积分,用于衡量模型的识别效果,数值越大表示对蔬菜幼苗识别效果越好。平均精度均值是多类别蔬菜幼苗识别平均精度的平均值,数值越高说明卷积神经网络的识别精确度越高。

3　试验结果分析

3.1　训练损失对比

利用 Faster R-CNN、R-FCN、本文模型对训练集图像进行训练。Faster R-CNN 模型中批尺寸为 1,训练周期为 100,1 个周期迭代次数为 5 544,初始学习率为 0.000 5,衰减率为 0.9,每迭代 1 个周期保存一次模型;R-FCN 模型中批尺寸为 2,训练周期数为 100,1 个周期迭代次数为 2 772,初始学习率为 0.001,衰减率为 0.9,每迭代 1 个周期保存一次模型,3 种模型的训练损失与周期数量关系曲线如图 4 所示。

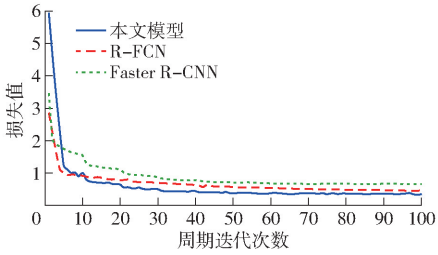


图 4　不同网络模型训练损失曲线对比

Fig. 4　Training loss curves of different networks

由图 4 可以看出,3 个模型的损失率随着周期迭代次数的增加逐渐降低,表明预测值与标注值的偏离程度不断缩小,经过 10 个迭代周期训练后本文模型训练损失小于另外两种模型,此时 Faster R-CNN、R-FCN 与本文模型迭代次数分别为 55 440、27 720、6 930 次,因此本文模型具有更快的收敛速度;此外,本文模型损失曲线收敛达到的稳定值低于 Faster R-CNN 与 R-FCN,表明本文模型具有更好的收敛性能。

3.2　模型有效性验证

为验证深度学习检测模型设计的有效性,本文对模型进行消融试验,首先建立基础检测模型(由混合深度分离卷积前置网络、轻量检测头单元、ROI pooling 运算,基于坐标回归的损失函数构成),然后在基础模型上增加 FPN 单元、进行 ROI Align 运算替换与 DIoU 损失替换构成扩展模型,将测试集样本输入到各版本模型进行预测,IoU 取值 0.5,结果如表 1 所示。由表 1 可以看出,随着功能模块的增加与替代,检测模型的 AP 和 MAP 逐渐提高,最终模型的 MAP 可以达到 97.47%。

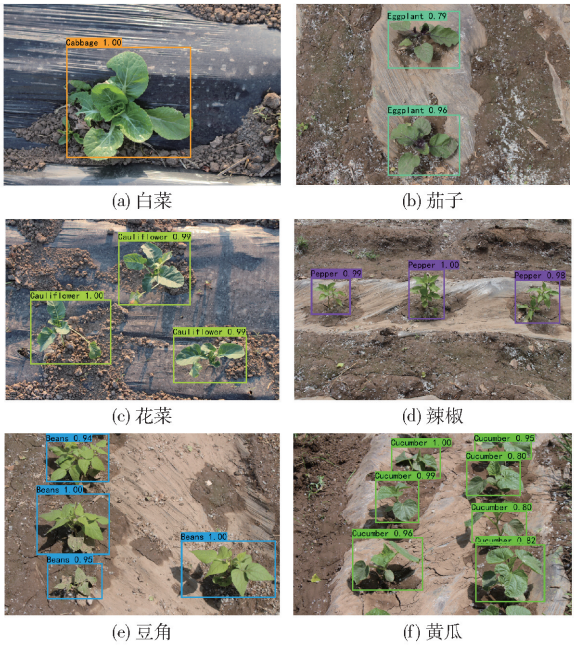
表 1　检测模型消融测试结果

Tab. 1　Ablation experimental results of detection model　　%

检测模型	AP						MAP
	白菜	花菜	黄瓜	辣椒	茄子	豆角	
基础模型	85.56	96.72	96.17	90.69	88.95	87.34	90.91
基础模型 + FPN	91.21	98.23	98.18	95.04	91.99	93.33	94.67
基础模型 + FPN + ROI Align	92.20	99.06	98.62	97.00	94.30	95.85	96.16
基础模型 + FPN + ROI Align + DIoU	93.98	99.17	98.85	98.22	96.89	97.69	97.47

图 5 为本文模型对多类蔬菜幼苗的检测效果图,可以看出本文模型除了对于叶片完整清晰的蔬菜幼苗能够准确进行位置回归与类别预测外,在作物叶片受到病虫害腐蚀出现残缺以及灌溉导致叶片附着泥土的情况下依然可以准确识别(图 5a、5e)。这是因为本文模型采用 FPN 网络将浅层高分辨率特征与深层语义特征融合,使不同层级特征信息丰富程度增加,能够更加有效地表达目标对象细节信息和抽象信息,即使目标对象表型发生变化或受到外界因素影响造成特征模糊,检测模型仍然具有较高的识别精度和较强的鲁棒性。此外,预测框能够准确覆盖蔬菜幼苗投影区域,没有出现边框过大包

含较多背景或者过小导致部分目标区域没有被包括的情况,说明引入 DIoU 作为位置回归损失能够使预测边框具有较高的准确性。



(a) 白菜 (b) 茄子

(c) 花菜 (d) 辣椒

(e) 豆角 (f) 黄瓜

图 5 不同种类蔬菜幼苗检测结果

Fig. 5 Detection results of different vegetable seedlings

3.3 模型性能对比测试

采用相同训练集样本对 Faster R - CNN 与 R - FCN 模型进行训练,将训练好的模型与本文模型进行性能对比,其中 Faster R - CNN 与 R - FCN 的基础网络采用 ResNet - 101。

图 6 为不同网络模型检测蔬菜幼苗得到的 $P-R$ 曲线。从图 6 中可以看出,召回率在[0,50%]区间时,3 种模型的精确率接近,维持在较高水平;召回率在[50%,100%]区间时,随着召回率的增加 3 个模型精确率呈下降趋势,在召回率相同的条件下本文模型精确率高于另外两种模型。

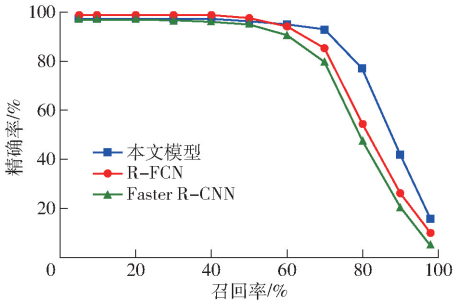


图 6 不同网络模型 $P-R$ 关系曲线对比

Fig. 6 Comparison of $P-R$ curves for various kinds of network

表 2 为 3 种检测模型的性能对比。在准确性方面,本文模型对蔬菜幼苗的平均精度均值比 Faster R - CNN 和 R - FCN 分别高 1.19、0.83 个百分点,具有较强的特征提取与精准的位置回归能力;在实时性方面,采用每秒处理的图像帧数作为速度衡量指标,由于本文模型对特征提取网络和检测头进行轻量化设计,降低了网络复杂度,推理速度分别为 Faster R - CNN、R - FCN 的 10.71、4.12 倍。在网络规模方面,采用模型存储空间作为衡量指标,Faster R - CNN 为 108 MB,R - FCN 为 96 MB,本文模型为 60 MB,小于前 2 种模型。试验结果表明,本文提出的轻量化二阶段深度学习检测模型,在平均精度均值、检测速度、网络存储空间等方面相比于 Faster R - CNN和 R - FCN 具有显著优势。

表 2 不同检测模型性能对比

Tab. 2 Comparison of different networks performance

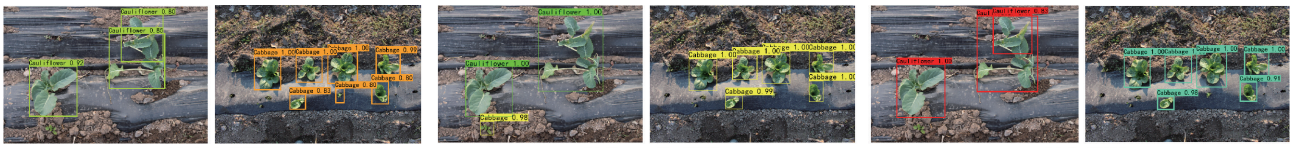
模型	MAP/%	检测速度/(f·s^{-1})	存储空间/MB
本文模型	97.47	19.07	60
Faster R - CNN	96.28	1.78	108
R - FCN	96.64	4.62	96

自然环境下,部分作物由于自身发育不均衡以及气候条件影响造成外形尺寸偏小;此外,株间距较近的作物进入展叶期后叶片间存在一定程度遮挡,以上情况给作物识别检测带来困难。为判别小目标作物(面积小于 32 像素 × 32 像素)和叶片交叠遮挡条件下的模型识别精度,本试验对每种情况随机抽取 100 幅图像,分别输入到 Faster R -CNN、R - FCN 和本文模型进行测试,结果如表 3 所示。本文模型对小目标作物和遮挡作物的平均精度均值分别为 87.73%、89.36%,2 种情况下平均精度均值为 88.55%,3 项指标均高于 Faster R - CNN 、R - FCN 模型。这是因为本文模型的前置基础网络利用不同尺寸卷积核进行运算具有较强的多尺度特征提取能力,FPN 单元将不同尺度特征信息融合强化了目标特征表达能力,小目标或叶片遮挡作物即使在特征图上语义信息有所损失依然可以得到有效识别。图 7 为 3 种模型对小目标和遮挡作物的检测效果图,左图为遮挡作物,右图为小目标作物。本文模型能够准确检测不同环境下的作物,Faster R - CNN 存在作物漏检、错检的情况,R - FCN 可以检测出遮挡作物,但是没

表 3 不同检测模型对小目标与遮挡作物检测精度对比

Tab. 3 Detection results of small and occluded crops with different detection models %

模型	MAP		平均值
	小目标作物	遮挡作物	
本文模型	87.73	89.36	88.55
Faster R - CNN	81.63	85.54	83.59
R - FCN	82.56	86.21	84.38



(a) 本文模型　(b) Faster R-CNN　(c) R-FCN

图 7　不同模型的检测效果

Fig. 7　Detection results of different network models

有正确检测出小目标作物。

4　结论

(1)采用轻量混合深度分离卷积作为前置特征提取网络,提高了特征信息提取效率和速度;通过压缩检测头网络特征通道规模与全连接层数量,降低了检测模型计算复杂度。本文模型存储空间为 60 MB,检测速度为 19.07 f/s,明显小于 Faster R－CNN 和 R－FCN 模型的存储空间,检测速度显著提高。

(2)基于 FPN 单元对多尺度特征层级信息进行融合,增加不同特征层的信息丰富程度,采用 ROI Align 运算代替 ROI pooling 运算并引入 DIoU 损失优化模型损失函数,实现目标预测框精准回归。试验结果表明,本文模型的 MAP 为 97.47%,分别比 Faster R－CNN 和 R－FCN 高 1.19、0.83 个百分点。同时,对小目标作物和叶片遮挡作物的识别具有良好的泛化性能。

参　考　文　献

[1] DARMON N, DARMON M, MAILLOT M, et al. A nutrient density standard for vegetables and fruits: nutrients per calorie and nutrients per unit cost[J]. Journal of the American Dietetic Association, 2005, 105(12): 1881－1887.

[2] 刘彦君,刘哲,孟祥红,等. 基于多样度、匹配度和平衡度的常见蔬菜营养价值评价[J]. 中国农业科学,2019,52(18):3177－3191.
LIU Yanjun, LIU Zhe, MENG Xianghong, et al. Evaluation of the vegetables nutritive values based on 'three degree' method[J]. Scientia Agricultura Sinica, 2019, 52(18): 3177－3191. (in Chinese)

[3] 中华人民共和国国家统计局. 中国统计年鉴(2018)[M]. 北京:中国统计出版社,2018.

[4] CHATTHA H S, ZAMAN Q U, CHANG Y K, et al. Variable rate spreader for real-time spot-application of granular fertilizer in wild blueberry[J]. Computers and Electronics in Agriculture, 2014, 100: 70－78.

[5] TANG J, CHEN X, MIAO R, et al. Weed detection using image processing under different illumination for site-specific areas spraying[J]. Computers and Electronics in Agriculture, 2016, 122: 103－111.

[6] NØRREMARK M, GRIEPENTROG H W, NIELSEN J, et al. The development and assessment of the accuracy of an autonomous GPS-based system for intra-row mechanical weed control in row crops[J]. Biosystems Engineering, 2008, 101(4): 396－410.

[7] ASTRAND B, BAERVELDT A. A vision based row-following system for agricultural field machinery[J]. Mechatronics, 2005, 15: 251－269.

[8] MEYER G E, NETO J C. Verification of color vegetation indices for automated crop imaging applications[J]. Computer and Electronics in Agriculture, 2008, 63(2): 282－293.

[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] NIEUWENHUIZEN A T, HOFSTEE J W, HENTEN E J. Adaptive detection of volunteer potato plants in sugar beet fields[J]. Precision Agriculture, 2010, 11(5): 433－447.

[11] PEREZ-RUIZ M, SLAUGHTER D C, FATHALLAH F A, et al. Co-robotic intra-row weed control system[J]. Biosystems Engineering, 2014, 126: 45－55.

[12] 张春龙,黄小龙,刘卫东,等. 苗间锄草机器人信息获取方法的研究[J]. 农业工程学报,2012,28(9):142－146.
ZHANG Chunlong, HUANG Xiaolong, LIU Weidong, et al. Information acquisition method for mechanical intra-row weeding robot[J]. Transactions of the CSAE, 2012, 28(9): 142－146. (in Chinese)

[13] 胡炼,罗锡文,曾山,等. 基于机器视觉的株间机械除草装置的作物识别与定位方法[J]. 农业工程学报,2013,29(10):12－18.
HU Lian, LUO Xiwen, ZENG Shan, et al. Plant recognition and localization for intra-row mechanical weeding device based on machine vision[J]. Transactions of the CSAE, 2013, 29(10): 12－18. (in Chinese)

[14] 祖琴,赵春江,邓巍,等. 基于可见-近红外光谱分析的圆白菜与杂草识别研究[J]. 光谱学与光谱分析,2013, 33(5):1202－1205.
ZU Qin, ZHAO Chunjiang, DENG Wei, et al. Research on discrimination of cabbage and weeds based on visible and near-infrared spectrum analysis[J]. Spectroscopy and Spectral Analysis, 2013, 33(5): 1202－1205. (in Chinese)

[15] 孙哲,张春龙,葛鲁镇,等. 基于 Faster R－CNN 的田间西兰花幼苗图像检测方法[J/OL]. 农业机械学报,2019,50(7):216－221.
SUN Zhe, ZHANG Chunlong, GE Luzhen, et al. Image detection method for broccoli seedlings in field based on Faster R－CNN [J/OL]. Transactions of the Chinese Society for Agricultural Machinery, 2019, 50(7): 216－221. http://www.j-csam.org/jcsam/ch/reader/view_abstract.aspx? file_no=20190723&flag=1& journal_id=jcsam. DOI: 10.6041/j.issn.1000-1298.

2019.07.023. (in Chinese)
[16] 王丹丹,何东健. 基于 R-FCN 深度卷积神经网络的机器人疏果前苹果目标的识别[J]. 农业工程学报,2019,35(3):156-163.
WANG Dandan, HE Dongjian. Recognition of apple targets before fruits thinning by robot based on R-FCN deep convolution neural network[J]. Transactions of the CSAE, 2019,35(3): 156-163. (in Chinese)
[17] 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)
[18] 魏宏彬,张端金,杜广明,等.基于改进型 YOLO v3 的蔬菜识别算法[J].郑州大学学报(工学版),2020,41(2):7-12,31.
WEI Hongbin,ZHANG Duanjin, DU Guangming, et al. Vegetable recognition algorithm based on improved YOLO v3[J]. Journal of Zhengzhou University (Engineering Science), 2020, 41(2): 7-12,31. (in Chinese)
[19] QUI R C, YANG C, MOGHIMI A,et al. Detection of fusarium head blight in wheat using a deep neural network and color imaging[J]. Remote Sens., 2019,11(22):2658.
[20] TAN M, LE Q. MixConv: mixed depthwise convolutional kernels[EB/OL]. [2019-11-01]. https://arxiv.org/ abs/1907.09595.
[21] LIN T, DOLLAR P, GIRSHICK R, et al. Feature pyramid networks for object detection[C]// Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2017: 936-944.
[22] LI Z, PENG C, YU G, et al. Light-Head R-CNN: in defense of two-stage object detector[EB/OL]. [2019-08-15]. https://arxiv.org/abs/1711.07264.pdf.
[23] YU J, JIANG Y, WANG Z, et al. Unitbox: an advanced object detection network[C]// Proceedings of the 24th ACM on Multimedia Conference, 2016: 516-520.
[24] REZATOFIGHI H, TSOI N, GWAK J, et al. Generalized intersection over union: a metric and a loss for bounding box regression[C]// Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2019:658-666.
[25] ZHENG Z, WANG P, LIU W, et al. Distance-IoU loss: faster and better learning for bounding box regression[C]// Proceedings of the AAAI Conference on Artificial Intelligence, 2020, 34(7):12993-13000.
[26] SHRIVASTAVA A,GUPTA A,GIRSHICK R. Training region-based object detectors with online hard example mining[C]// Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2016: 761-769.

(上接第 202 页)

[16] 丛宏斌,赵立欣,孟海波, 等. 立式环模秸秆压块机成型过程建模与参数优化[J/OL]. 农业机械学报, 2014,45(10):187-193.
CONG Hongbin, ZHAO Lixin, MENG Haibo, et al. Process modeling and parameter optimization of straw briquetting machine with vertical circular mould[J/OL]. Transactions of the Chinese Society for Agricultural Machinery, 2014,45(10):187-193. http://www.j-csam.org/jcsam/ch/reader/view_abstract.aspx? flag=1&file_no=20141029&journal_id=jcsam. DOI: 10.6041/j.issn.1000-1298.2014.10.029. (in Chinese)
[17] 宁廷州,俞国胜,皮森淼,等. 对辊柱塞式成型机成型参数优化[J/OL]. 农业机械学报, 2016,47(7):208-215.
NING Tingzhou, YU Guosheng, PI Senmiao, et al. Optimization of forming parameters on roller briquetting machine with plungers[J/OL]. Transactions of the Chinese Society for Agricultural Machinery, 2016,47(7):208-215. http://www.j-csam.org/jcsam/ch/reader/view_abstract.aspx? flag=1&file_no=20160729&journal_id=jcsam. DOI:10.6041/j.issn.1000-1298.2016.07.029. (in Chinese)
[18] 中国农业大学. 多级辊压秸秆成型装置: 201510173725.5[P]. 2015-07-22.
[19] 丁宁. 秸秆压缩成型技术与装置研究[D]. 北京:中国农业大学,2018.
DING Ning. Research of compression molding technology and equipment for straw[D]. Beijing: China Agricultural University, 2018. (in Chinese)
[20] 刘平义,骆龙敏,李海涛,等. 秸秆多级辊压成型原理与装置设计[J]. 农业机械学报,2016,47(增刊):317-323.
LIU Pingyi, LUO Longmin, LI Haitao, et al. Theory and device design of multi-stage roller forming for straw[J]. Transactions of the Chinese Society for Agricultural Machinery, 2016, 47(Supp.): 317-323. (in Chinese)
[21] 丁宁,李海涛,闫安,等. 秸秆多级连续冷辊压成型试验研究[J/OL]. 农业机械学报,2021,52(5):279-285.
DING Ning, LI Haitao, YAN An, et al. Experiment on multi-stage continuous cold roll forming of straw[J/OL]. Transactions of the Chinese Society for Agricultural Machinery, 2021,52(5):279-285. http://www.j-csam.org/jcsam/ch/reader/view_abstract.aspx? flag=1&file_no=20210530&journal_id=jcsam. DOI:10.6041/j.issn.1000-1298.2021.05.030. (in Chinese)
[22] 吴娟娟,霍丽丽,赵立欣, 等. 玉米秸秆安全存储平衡水分试验研究[J]. 中国农机化学报, 2016,37(11):73-77.
WU Juanjuan, HUO Lili, ZHAO Lixin, et al. Experiment research on maize straw equilibrium moisture for secure storage[J]. Journal of Chinese Agricultural Mechanization, 2016,37(11):73-77. (in Chinese)
[23] 郭磊. 秸秆压块成型因素与压模腔体的优化研究[D]. 北京:中国农业大学,2016.
GUO Lei. Effects of straw characteristics and research of compression die on briquetting[D]. Beijing: China Agricultural University, 2016. (in Chinese)
[24] 何为,薛卫东,唐斌. 优化试验设计方法与数据处理[M]. 北京:化学工业出版社,2012.
[25] 雷廷宙,何晓峰,王志伟,等. 生物质固体成型燃料生产技术[M]. 北京:化学工业出版社,2020.