

基于快速卷积神经网络的果园果实检测试验研究^{*}

张磊¹, 姜军生², 李昕昱³, 宋健², 解福祥²

(1. 山东科技大学机械电子工程学院, 山东青岛, 266590; 2. 潍坊学院机电与车辆工程学院, 山东潍坊, 261061;
3. 山东工业技师学院, 山东潍坊, 261061)

摘要:近年来,基于数字图像处理和机器学习算法的果实自动识别检测研究已经越来越成熟。针对传统检测方法检测过程中难以满足实时性要求的缺点,采用了基于 Faster-RCNN 的果实快速检测模型。模型由卷积神经网络(CNN)和区域提议网络(RPN)组成,首先由 CNN 进行卷积和池化操作提取特征,然后由 RPN 选取候选区域,通过网络全连接层参数共享,由目标识别分类器和边界框预测回归器得到多个可能包含目标的预测框,最后通过非极大值抑制挑选出精度最高的预测框完成目标检测。分别对桃子、苹果和橙子的三种果实进行检测,采用迁移学习方法,使用已经预训练好的两种深度神经网络模型 ZFnet 和 VGG16,通过数据集的训练对 Dropout 及候选区域数量进行参数调整完成网络调优。检测并分析果实不同布局形态下模型的检测效果。试验结果表明,当 Dropout 取值为 0.5 或 0.6,候选区域数量为 300 时网络模型最佳,ZFnet 网络中,苹果平均精确度为 92.70%,桃子为 90.00%,而橙子为 89.72%。VGG16 网络中,苹果平均精度为 94.17%,桃子为 91.46%,橙子为 90.22%。且 ZFnet 和 VGG16 的图像处理速度分别达到 17 fps 和 7 fps,能够达到果实实时检测的目的。

关键词:果实检测;深度学习;农业机器人;神经网络;迁移学习

中图分类号:S24; TP24 **文献标识码:**A **文章编号:**2095-5553 (2020) 10-0183-09

张磊, 姜军生, 李昕昱, 宋健, 解福祥. 基于快速卷积神经网络的果园果实检测试验研究[J]. 中国农机化学报, 2020, 41(10): 183—190, 210

Zhang Lei, Jiang Junsheng, Li Xinyu, Song Jian, Xie Fuxiang. Experimental research on orchard fruit detection based on fast convolutional neural network [J]. Journal of Chinese Agricultural Mechanization, 2020, 41(10): 183—190, 210

0 引言

随着机器视觉和人工智能技术的发展,果实自动检测一直是农业工程领域研究的热点。实现快速且准确的果实检测,不仅能为果实采摘机器人提供可靠的行动依据,还能为果园产量图获取以及数字化果园管理等应用提供科学的技术指导。近年来,基于数字图像处理和机器学习算法的果实自动识别检测研究已经越来越成熟。传统思路是首先人工提取检测目标的特征,然后选择合适的分类器进行训练,最终将训练好的分类器用于果实识别和边界框回归,从而完成果实检测任务。虽然该思路在特定条件下针对特定目标能够达到较高的准确率^[1],但在自然光照下的户外果园中,由于光照条件、果实生长密集度、果实的视距以及相机视点等因素的不确定性,增加了特征设计的难度和不确定性,进而很大程度上影响了算法的鲁棒性和泛化

性能,且在检测过程中难以满足实时性要求^[2-3]。

深度学习的概念由 Hinton 等在 2006 年的提出,该方法逐渐被广泛应用于处理目标检测、语义图像分割、姿态估计、自然语言处理、语音识别等问题^[4]。其核心优势在于通过构建包含多个隐含层的神经网络模型和海量训练数据,使系统自主学习并提取从低层级到高层级的特征函数,实现原始数据特征的自学习,避免了人工计算特征带来的干扰并优化了特征空间。在农业应用领域,Tan 等^[5]通过红外视频传感器实时检测果园果实信息,利用卷积神经网络(CNN)对苹果病变情况进行判断。首先对苹果病变图像数据进行姿态和光照仿真,判断病变图像的普遍发生状态,创建能够代表大多病变情况的数据集。然后输入到 CNN 中进行卷积和池化并训练模型,该方法有较高的准确率和召回率。Bargoti 等^[6]利用卷积神经网络(CNN)对苹果树图像进行卷积特征提取,然后将特征输入到支持

收稿日期:2020 年 2 月 14 日 修回日期:2020 年 5 月 28 日

^{*} 基金项目:国家自然科学基金项目(51505337);山东省重点研发计划项目(2019GNC106144);山东省农机装备研发创新技术项目(2018YF005—05);山东省高等学校科技计划项目(J17KA150)

第一作者:张磊,男,1993 年生,山东潍坊人,硕士研究生;研究方向为机器人技术。E-mail: 1562631752@qq.com

通讯作者:姜军生,男,1963 年生,山东潍坊人,硕士,教授,硕导;研究方向为机器人技术。E-mail: 163jsjiang@163.cn

向量机(SVM)分类器中,通过池化操作降低特征空间维度,最终输出分割后图像实现苹果的像素级分割。并且结合传统二维图像分割识别算法,将深度学习输出的结果通过分水岭分割和 Hough 圆估计,完成了果实的准确识别检测。但因为深度学习和图像处理方法的融合造成了算法容量较大,在定位准确率和实时性方面仍需要改进。为了能将卷积神经网络直接运用到物体检测研究中,Girshick 等^[7]提出了区域卷积神经网络算法(RCNN),为基于深度学习的目标检测算法提供了新的思路,相比传统算法大大提高了目标检测效率。并且通过对算法容量、实时性和鲁棒性等方面的改进,学者们相继提出了 SPP-Net^[8]、Fast-RCNN^[9]、Faster-RCNN^[10]等算法对各类问题进行优化,使得深度神经网络在物体检测研究中有更出色的表现。

本文以实现自然光照条件下,果园环境中果实的自动检测为目的,将深度学习算法应用到果实的识别检测任务中,通过试验结果分析和算法性能对比,验证了深度学习算法在果实检测研究中的可行性,为该领域相关研究提供了依据。

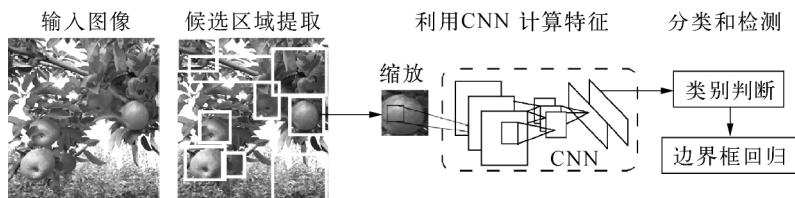


图1 RCNN 算法流程

Fig. 1 Flowchart of RCNN

虽然 RCNN 一方面通过前期候选区域的选择以及后期利用重叠度进行候选框框回归能够得到精度较高的目标检测边界框,另一方面得益于 CNN 在图像特征提取方面的优异表现最终能够获得较高的精度和召回率。但是其缺点也显而易见,首先是训练模型时需要经过多个数据处理阶段,使得训练时间和空间的开销非常大,训练网络时间长达 84 h,测试一副图像需要 47 s。

为了能使 RCNN 模型在鲁棒性和实时性得到进一步提升,微软研究院的学者们从特征提取简化、CNN 模型改进以及分类和回归联合训练等方面入手相继提出了 SPP-Net 和 Fast-RCNN 网络,不仅在检测准确度方面有很大提升,模型运算速度也越来越接近实时性目标检测的要求。然而候选区域提取步骤的独立进行,仍然是算法时间开销方面的瓶颈问题。

1.2 Faster-RCNN 目标检测算法概述

Shaoqing 等在 RCNN 的基础上,融合了 SPP-Net 和 Fast-RCNN 的特点,构建了 Faster-RCNN 模型。该模型创造性的提出了共享的卷积网络组成区域提议网络 RPN,能自主预测出质量高且数量少的候选框。

1 基于深度学习的目标检测模型概述

1.1 RCNN 目标检测模型概述

传统的目标检测方法一般步骤,首先在给定的图像上选择候选区域,然后对这些区域提取特征,最后训练分类器进行分类。而这一过程存在两个问题:一是基于滑动窗口的区域选择策略容易造成窗口冗余且计算复杂度高;二是人工设计特征步骤繁多且不能具有很强的鲁棒性。针对上述问题,RCNN 首先提出了选择性搜索方法,即预先分割并提取一系列可能是物体的候选区域作为输入,随后引入了 CNN 模型进行特征自学习。该算法主要分 4 个步骤:(1)采用选择性搜索方法将一张图片分割成 2 000 个左右的候选区域,由于候选区域尺寸不一,通过缩放形成统一的输入尺寸。(2)把统一尺寸的候选区域输入 CNN 中,进行特征提取。(3)将通过 CNN 自学习得到的特征输入到 SVM 分类器中进行类别判断。(4)使用回归器对候选区域的边界框(候选框)位置进行修正。算法流程如图 1 所示。

同时再由基于 Fast-RCNN 的特征提取、分类和回归框架完成目标检测。Faster-RCNN 模型训练阶段的整体框架如下,如图 2 所示。

1) 将已经进行对象标注的图片输入到 CNN 中,通过卷积层和池化层进行特征提取。

2) 基于区域提议网络(RPN)生成 300 个候选区域,然后将所有候选区域映射到 CNN 的最后一层卷积特征图上,最后通过兴趣区域池化(RoI pooling)使每个候选区域生成固定尺寸的特征向量。

3) 将得到的特征向量输入两个独立平行的全连接层:box 分类层(box-classification layer)和 box 回归层(box-regression layer),分别进行分类预测和边界框回归预测,通过 CNN 与 RPN 的参数共享得到映射到原图的预测框包含目标的概率值,即得分;以及预测框位置参数。

4) 经过分类和回归运算后,通常会得到多个可能是被检测目标的预测框,通过计算 n 个预测框的得分然后进行非极大值抑制,最终选取精度最高的边界框,并输出预测窗口坐标。

从本质上来讲,Faster-RCNN 可以简单的看作由 PRN 和 Fast-RCNN 组合而成,用 RPN 作为选择性搜索方法是它的核心思想。其优势不仅在于通过 RPN 直接预测出质量高、数量少的候选区域,而且大部分工

作都在 GPU 中完成,实现了端对端的训练过程,大大降低了时间和空间的开销,从而大幅度提升了目标检测的速度,让业界看到了深度学习算法在目标检测任务上进行实时处理的希望。

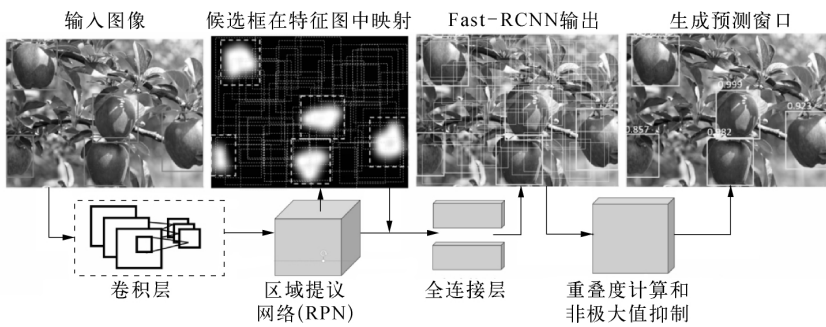


图 2 Faster-RCNN 算法流程
Fig. 2 Flowchart of Faster-RCNN

2 目标检测模型训练

2.1 数据预处理

1) 用于训练模型和测试的图像尺寸根据图像获取设备不同会有所差异,本文主要采用两种相机获取不同果树图像:Kinect V2 和佳能 100D,采集到的彩色图像像素尺寸分别为 1 920 像素×1 080 像素和 3 456 像素×5 184 像素。虽然 Faster-RCNN 能够接受任意尺寸的图片,但分辨率过高的图片不仅会加大 GPU 运算的开销,还会给后期迁移学习做网络参数微调带来不便,所以首先对所有采集到的图片进行下采样,统一缩放成 300 像素×300 像素尺寸。

2) 基于深度学习的模型训练精度和训练集数据量成正比,但往往由于各种客观因素在用特定目标训练网络时数据量很小,从而容易造成过拟合^[11]。为了增加模型精度和扩充样本数据,让模型具有更高的精确率,通常采用数据增强方法人工增加训练数据的多样性。本文主要采用水平翻转、0°~180°随机旋转以及重缩放处理操作。

3) 训练模型时,需要对输入的图像人工进行目标标注,目的是让系统得到正例信息和目标的真实位置信息,以便于后期的分类预测和边界框回归。本文基于 Python 语言编写了图像标注工具,如图 3 及表 1 所示。

表 1 图样标注样例输出坐标
Tab. 1 Drawing output coordinates of sample

类别/坐标	x_1	y_1	x_2	y_2
Peach1	86	8	146	88
Peach2	127	62	176	142
Peach3	69	100	122	176
Peach4	20	216	75	287
Peach5	165	190	224	266
Peach6	215	155	269	232

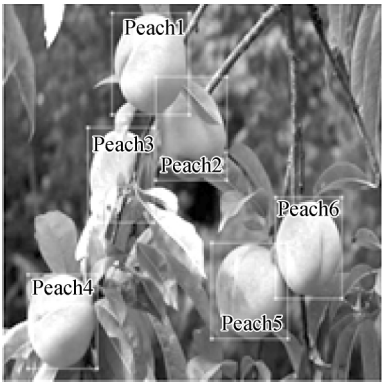


图 3 图像标注样例
Fig. 3 Example of image annotation

本文能够快速标注目标然后输出目标坐标值, (x_1, y_1) 为目标左上角坐标点, (x_2, y_2) 为右下角坐标点。并将标注信息保存在.txt 文档和.xml 文件中,用于后期处理操作。

2.2 CNN 模型训练

Faster-RCNN 基本结构仍然是卷积神经网络,重新训练一个深层级的 CNN 不仅需要海量的样本数据,还要花费数周时间完成训练和参数配置。本文采用迁移学习方法^[12-14],即使用已经基于其他数据集训练好的 CNN 模型,得到网络关键参数如权重的初始化设置,然后通过对新的标签数据的特征学习,进行网络参数微调,从而能够让模型面对新的目标检测任务时,在保证检测精度的基础上大幅度提高工作效率^[15]。

本文采用了基于 ImageNet 数据库预训练好的两种 CNN 模型:ZFnet 网络和 VGG16 网络,前者包含 5 个卷积层 2 个池化层以及最后通过 3 个全连接层输出特征;后者包含 16 个卷积层 5 个池化层以及 3 个全连接层。对于数据量较小的检测任务,ZFnet 已经能得到不错的检测精度。VGG16 因为具有更深的网络结构和较小的滤波器尺寸,不仅能够胜任大规模数据集的检测任

务,同时能够得到更优秀平均精确度均值(mAP)。

针对 CNN 的模型监督优化和参数更新的方法有很多,通过协调正向与反向传播的参数更新来控制参数优化,利用负梯度 $\nabla L(W)$ 和上一次权重的更新值 V_t 的线性组合来更新权重 W , 如式(1)所示。

$$\begin{aligned} V_{t+1} &= \mu V_t - \alpha \nabla L(W_t) \\ W_{t+1} &= W_t + V_{t+1} \end{aligned} \quad (1)$$

式中: α ——学习率;

$\nabla L(W_t)$ ——负梯度的权重;

μ ——一次更新值的权重。

2.3 区域提议网络(RPN)训练

Faster-RCNN 的核心在于通过建立 RPN 来输出可能包含目标的候选区域,RPN 结构如图 4 所示。

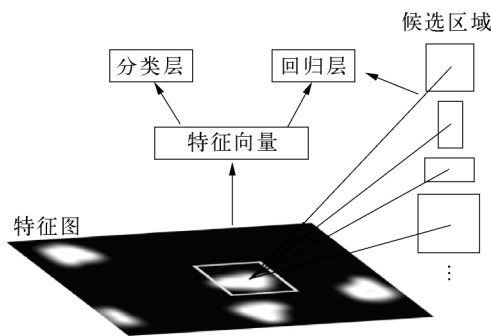


图 4 PRN 结构图

Fig. 4 Region Proposal Network

通过 CNN 网络的最后一层卷积层生成特征图后,RPN 以特征图上每个点为中心使用滑框进行滑动。为了能够应对不同尺寸的物体,RPN 分别采用窗口面积 128 像素×128 像素、256 像素×256 像素、512 像素×512 像素对应的长宽比 1:1、1:2、2:1 的 9 种类型滑窗。这些窗口经过卷积形成 256 维(ZFnet)或 512 维(VGG16)的特征向量,该向量反映了该点对应窗口的深层特征,然后将这些特征向量分别输入分类层和回归层进行预测:分类层挑选概率得分较高的候选区域作为目标样本,回归层对这些目标样本的候选框的位置和大小进行修正。分类层和回归层损失函数

$$\begin{aligned} L(\{p_i\}, \{t_i\}) &= \frac{1}{N_{cls}} \sum_i L_{cls}(p_i, p_i^*) + \\ &\quad \gamma \frac{1}{N_{reg}} \sum_i p_i^* L_{reg}(t_i, t_i^*) \end{aligned} \quad (2)$$

式中: p_i ——预测得分;

p_i^* ——标注样本得分,正例样本为 1,负例样本为 0;

t_i ——目标坐标参数;

t_i^* ——标注样本坐标参数,坐标参数有四个,分别是中心点 x 和 y 值以及候选框的宽度和高度;

N_{cls} 、 N_{reg} ——mini-batch size、Anchor Location 的数量;

γ ——均衡参数,默认值取 10。

3 试验与评价

3.1 试验数据

本文试验数据中,70%的果实图片在果园场景实地采集,地点为美国佐治亚州 Butts 郡和佛罗里达州 Baker 市的商业果园基地,分别于 2016 年 6 月、9 月以及 11 月获取了桃子、苹果和橙子的成熟期果树图像,并以同一棵树的图片数量不能大于 5 张的原则进行数据采集。同时,为了扩充数据量增加模型训练精度,从 Google 图像库和百度图像库获取了 30%的相应果实的图像数据。然后用 2.1 节介绍的数据处理方法对训练集数据进行处理,其中训练集图片占 80%,测试集图片占 20%,试验数据分布如表 2 所示。

表 2 试验数据分布
Tab. 2 Distribution of dataset

果实	采样尺寸 /(像素×像素)	果实大小 /mm	平均果实 数量/图	训练集 图片数
桃子	300×300	50±15	5.5	253
苹果	300×300	43±20	6.7	556
橙子	300×300	30±15	7.0	152

3.2 果实检测

在实际应用时,场景中一般只出现一个种类的果实,只针对一种果实的分类问题能转化为最简单的二分类问题,即背景和果实,且类别数量的限制能有效提高分类效率,所以本文将每个果实种类的数据进行分开训练,得到只针对该水果的深度学习模型^[16-18]。本文试验硬件配置:Core i7 处理器、CPU 内存 16 GB、GPU 为 GTX1080,显存 8 GB、硬盘 1 TB。深度神经网络模型均在 Caffe 深度学习框架下完成。完成数据准备工作后,将训练集数据输入已经预训练好的 ZFnet 和 VGG16 网络进行网络微调。设置最大迭代次数为 10 000,初始学习率 α 为 0.01,学习率变化因子 g 为 0.1, μ 为 0.9,同时按照 2.3 节方法训练 RPN。

在测试阶段,考虑到果树图像中果实发生遮挡、重叠的比例较大,且果实目标相对较小,设置重叠率 $IoU=0.4$ 。当对目标的预测框与真实标注框的重叠率大于该值时,则认为是判断正确的正例样本。对于果实检测结果和模型性能的评价,主要采用精确度 P (Precision)、召回率 R (Recall)以及度量函数 F_1 (F_1 —Score),定义如下^[19-22]。

$$P = \frac{T_P}{T_P + F_P} \quad (3)$$

$$R = \frac{T_P}{T_P + F_N} \tag{4}$$

$$F_1 = \frac{2PR}{P + R} \tag{5}$$

式中： T_P ——被正确检测到的正例样本个数，即正确检测；
 F_P ——错误的判断为正例样本的个数，即错误检测；
 F_N ——错误的判断为负例样本的个数，即丢失。

3.2.1 试验结果

完成准备工作并训练好检测模型后，将测试集数据输入模型进行果实检测，结果如表 3 所示。首先，三种水果在基于 VGG16 和 ZFnet 的检测结果都表现出较理想的预测效果，精确度均在 90% 以上，其中苹果类精确度最高，达到 94.17%，得益于苹果在三类数据中拥有最多的训练集数据，模型能够得到更多机会更新权重等性能参数从而提高性能。其次，在基础网络模型对比中，VGG16 的性能要整体高于 ZFnet，三类水果中的召回率均高于 ZFnet 网络，在桃子检测中，ZFnet 相比较 VGG16 召回率相差高达 4.83%，主要得益于前者在系统硬件要求以及训练时间开销方面要大大高于后者。在检测速度方面，ZFnet 的图像处理速度达到 17 fps，VGG16 处理速度为 7 fps，所以对于数据规模较小，精度要求不高的应用，ZFnet 已经完全能够胜任目标实时检测任务。而 VGG16 不仅模型性能要高于前者，还能够处理大规模数据，且在处理速度方面也能满足实时检测的要求。

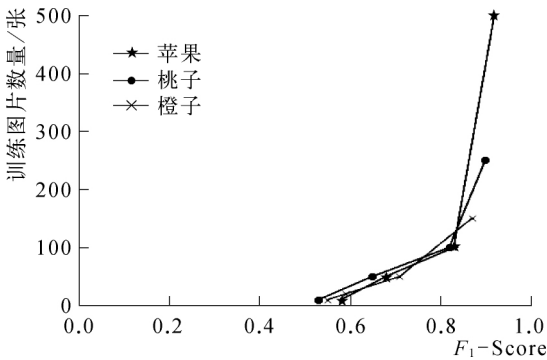
表 3 检测结果
Tab. 3 Result of detection

果实	CNN	精确度/%	召回率/%	F_1 -Score/%
桃子	VGG16	91.46	94.31	92.86
	ZFnet	90.00	89.48	89.74
苹果	VGG16	94.17	91.94	93.04
	ZFnet	92.70	88.61	90.60
橙子	VGG16	90.22	87.91	89.05
	ZFnet	89.72	83.15	86.31

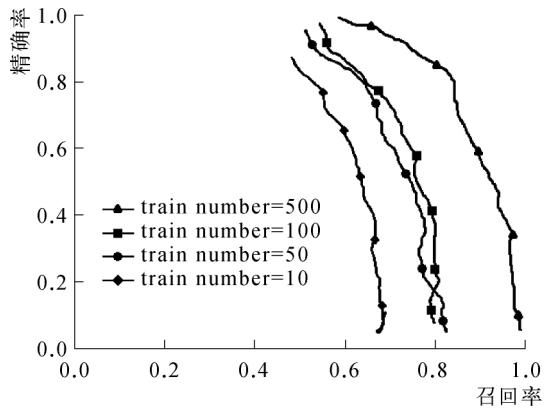
3.2.2 训练数据量对模型性能影响

通过表 3 中数据看出，苹果类的检测在三类数据中拥有最高的精确度，侧面反映了训练集的大小对模型性能的影响。为了验证不同训练集数量对 Faster-RCNN 模型性能的影响，本文用不同数量的训练集数据训练模型然后测试结果，试验均使用 VGG16 网络模型，结果如图 5 所示。图 5(a)显示了三种果实的 F_1 -Score 值随着训练图片数量的增加而上升，随着

训练集的增大，模型拥有更好的性能。在相同训练集下， F_1 -Score 值也受到平均每张图片包含果实数量的影响。当训练图片在 250 张时， F_1 -Score 值的提高就需要训练图片的大幅度增加。另一方面，图 5(b)显示了对于苹果类，不同大小训练集下的精确—召回曲线，该曲线反映了模型对样本预测的综合性能，从图中可以看出训练集越大，模型的综合性能越好，当训练集达到 500 时，效果最好，这也可以从左图中看出。



(a) 三种果实 F_1 -Score 值随训练图片数量的变化



(b) 苹果在不同训练图片数量下的精确—召回曲线

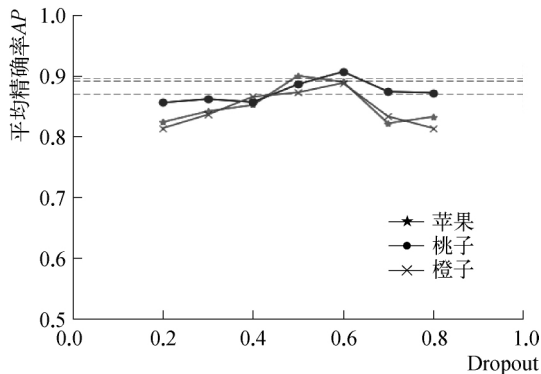
图 5 训练数据对模型性能影响
Fig. 5 Impact of the number of training images for performance of model

3.3 模型调优

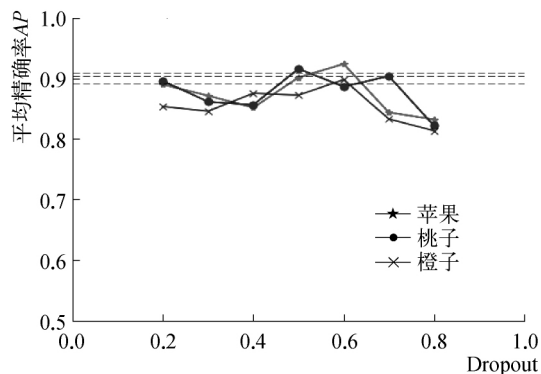
3.2 节试验已经得出训练样本的数量对模型性能的影响。但在现实应用中，往往由于各种客观因素不能得到理想数量的训练样本，而训练集过小容易造成模型过拟合。这就需要采用一些其他方法对模型进行优化进而保证最优的模型性能。本文主要通过两方面讨论模型优化方法。

1) Dropout 是指在模型训练时随机让网络中的某些隐含层节点的权重在该轮迭代中不更新，这样权重值的更新就不再依赖具有固定关系隐含节点的共同作用，防止了某些特征只在特定场景下才作用的情况，增加了特征的泛用性^[23-26]。Dropout 取值为[0.0, 1.0]，为停止更新节点的比例。试验结果如图 6 所示，颜色

虚线代表了不进行 Dropout 时模型的平均精确率。从结果中可以看出,进行 Dropout 后的最高输出精确率要高于未设置 Dropout,两种网络下,三种果实均在 Dropout 取值为 0.5 或 0.6 时达到最高点,即在该值时,模型的性能最优。



(a) ZFnet 网络



(b) VGG16 网络

图 6 不同 Dropout 值下的平均精确率

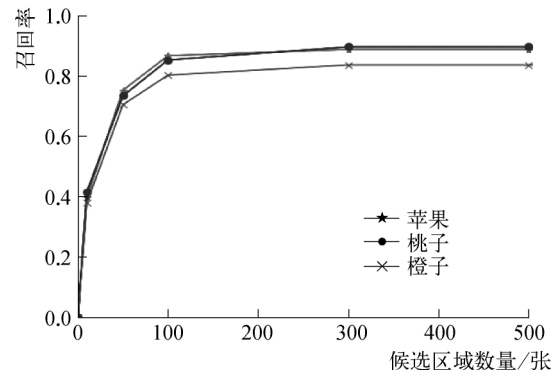
Fig. 6 Average precision with different dropout

2) 候选区域数量,虽然本文对基于 RPN 生成的候选区域数量定为 300,但观察不同候选区域数量下模型的检测率(召回率)的变化,有助于了解 RPN 对整个模型性能的贡献以及 Faster-RCNN 的核心优势。结果如图 7 所示。召回率随候选区域数量的增大而增大,当区域数量为 300 时,召回率值最高。随后,当区域数量增大到 500 时,召回率基本保持不变。另一反面,两种网络中橙子的召回率均低于苹果与桃子,源于橙子训练图片数量远低于其他两类,比桃子和苹果分别低 40%和 73%,桃子为 252 张,苹果为 556 张,所以当训练集达到 250 左右这个数量时,候选区域数量的增加对于召回率的提高意义不大。

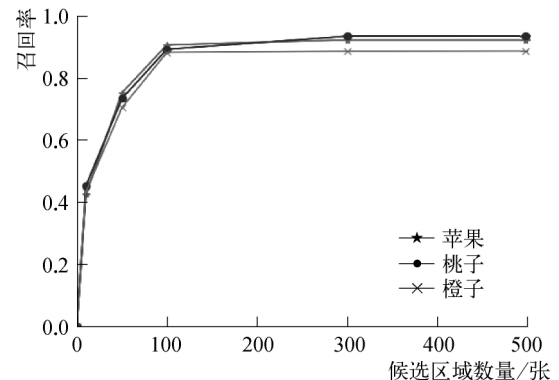
4 讨论

在非结构化果园环境中,果实的生长布局拥有很强的随机性,并且在图像获取过程中,不可避免的会发生部分果实被遮挡以及果实重叠的情况。本文从基于

Faster-RCNN 的检测模型对三类水果的测试集数据的检测结果中,挑选出四类比较有代表性的果实布局形态:稀疏型、密集型、远景拍摄以及遮挡和重叠情况较多的图像。



(a) ZFnet 网络



(b) VGG16 网络

图 7 召回率随候选区域数量变化曲线

Fig. 7 Recall with different number of proposal region

检测结果如图 8 所示,从第一列图像中可以看出,对于生长较为稀疏且遮挡和重叠概率不高的图像,本文模型能够基本全部准确检测出果实,并且边界框的位置也十分准确,对检测结果进行统计,该类图像苹果、桃子、橙子检测精确度分别为 95.94%,94.12%,94.07%。对于生长较为密集的第二列图像,在场景深处的果实容易被检测丢失,且一些重叠的果实系统会检测为一个果实,该类图像检测精确度分别为 89.56%,89.33%,89.71%。在第三列图像中,因为相机视角较远会造成果实相对较小,但检测结果仍然具有较高的准确率,该类检测结果精确度分别为 92.93%,92.85%,91.77%。对于遮挡和重叠情况,从第四列图像中可以看出,系统将重叠果实判定为一个果实的概率是比较高的,且不容易检测出遮挡率较高的果实,该类检测结果精确度分别为 89.36%,88.79%,88.17%,这是由于一方面因为被严重遮挡的果实在颜色反射以及形状表现方面和背景没有很强的可分性,另一方面即使被选为候选目标区域,也因未到达设置的 IoU 值而被过滤掉。

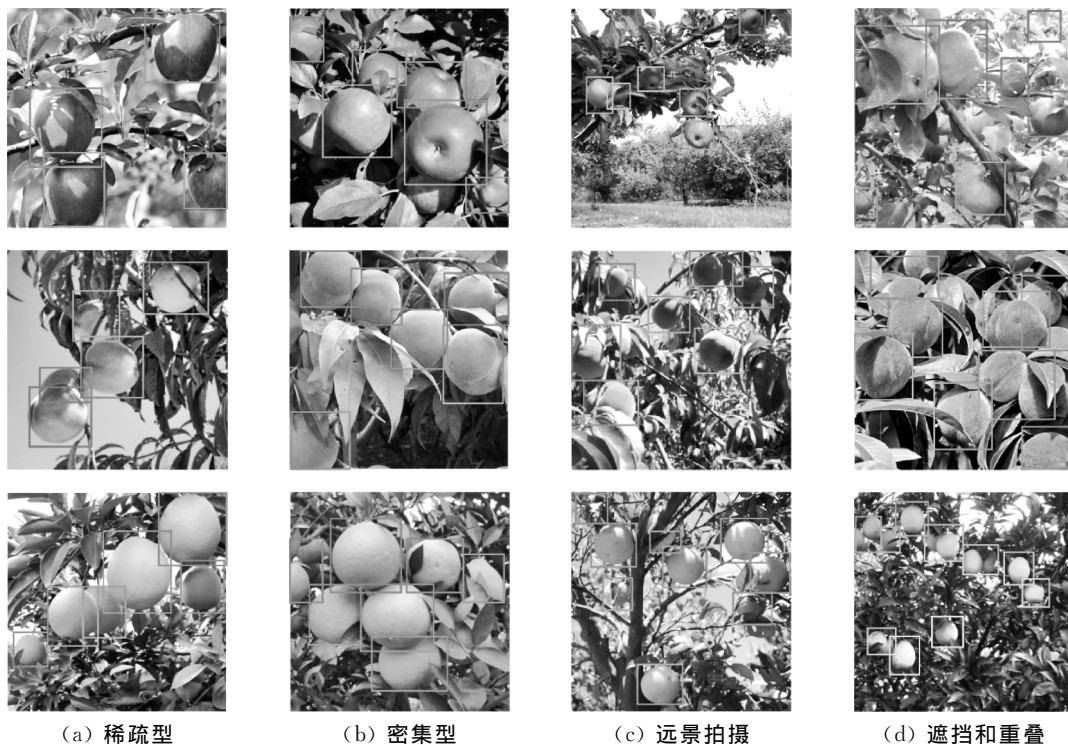


图 8 检测结果

Fig. 8 Results of detection

但从整体检测结果来看,基于深度学习的果实检测模型在测试数据集中表现出较高的精确率和召回率,反映了该模型具有较好的泛化性能。同时,在自然光条件下的果园环境中对于果实遮挡和重叠等情况,也具有较强的鲁棒性。并且,相比于传统机器学习算法耗时的缺点,本文模型检测每副图像的平均耗时在 59 ms 左右,在检测速度方面具有较大优势。

5 结论

本文研究了基于深度神经网络的目标检测算法在果实目标检测中的应用,使用 Faster-RCNN 模型分别检测了桃子、苹果和橙子的图像数据,并分析了果实不同布局形态下模型的检测效果。

1) 对于果实检测,我们设计了稀疏型,密集型,远景拍摄以及遮挡重叠四种情景进行试验,试验结果表明,该模型不仅在果实检测方面具有较高的准确率,并且在检测速度方面相比传统的机器学习算法有很大的提高,ZFnet 和 VGG16 的图像处理速度分别达到 17 fps 和 7 fps,能够满足实时检测的基本要求。另一方面,该模型在检测果实遮挡和重叠率较高的场景时仍需要提高其鲁棒性,在本文进一步研究中,会考虑从多源图像融合以及算法改进等方面提高模型性能。

2) 对网络模型的 Dropout 取值进行试验,确定取

值在 0.5 或者 0.6 时网络模型最佳。ZFnet 网络中,苹果平均精确度为 92.70%,桃子为 90.00%,而橙子为 89.72%。VGG16 网络中,苹果平均精度为 94.17%,桃子为 91.46%,橙子为 90.22%。VGG16 网络整体性能优于 ZFnet 网络。

3) 对候选区域数量的设定进行试验,当候选区域在 300 时,训练数据为 250 张图片时,网络模型的召回率最高。此时提高网络性能只能通过增加训练数据集不断更新权重参数进行,召回率数值始终稳定,不随训练数据集的增加而增加。

参 考 文 献

- [1] 宋健,王凯,张晓琛. 茄子采摘机器人目标识别与测距方法[J]. 实验室研究与探索, 2015, 34(9): 54—57.
Song Jian, Wang Kai, Zhang Xiaochen. A target recognition and ranging method for eggplant picking robot [J]. Research and Exploration in Laboratory, 2015, 34(9): 54—57.
- [2] 赵金英,张铁中,杨丽. 西红柿采摘机器人视觉系统的目标提取[J]. 农业机械学报, 2006, 37(10): 200—203.
- [3] 谢万里,李宏志,周辉,等. 基于迁移学习与卷积神经网络的鱼濒死预警系统研究[J]. 中国农机化学报, 2019, 40(2): 186—192.
Xie Wanli, Li Hongzhi, Zhou Hui, et al. Fish dying on the basis of convolution neural network and transfer learning [J]. Journal of Chinese Agricultural Mechanization, 2019, 40(2): 186—192.

- [4] 赵杰文, 刘木华, 杨国彬. 基于 HIS 颜色特征的田间成熟番茄识别技术[J]. 农业机械学报, 2004, 35(5): 122—124. Zhao Jiewen, Liu Muhua, Yang Guobin. Discrimination of mature tomato based on HIS color space in natural outdoor scenes [J]. Transactions of the Chinese Society for Agricultural Machinery, 2004, 35(5): 122—124.
- [5] Tan W, Zhao C, Wu H. Intelligent alerting for fruit-melon lesion image based on momentum deep learning [J]. Multimedia Tools and Applications, 2016, 75 (24): 16741—16761.
- [6] Bargoti S, Underwood J. Deep fruit detection in orchards [C]. IEEE International Conference on Robotics & Automation. IEEE, 2017.
- [7] 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.
- [8] Girshick R. Fast R — CNN [C]. IEEE International Conference on Computer Vision, 2015.
- [9] 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 & Machine Intelligence, 2015, 39(6): 1137—1149.
- [10] 宋健. 基于神经网络的茄子采摘机器人视觉识别方法[J]. 潍坊学院学报, 2011, 11(6): 90—93. Song Jian. Target identification method of eggplants picking robot based on improved BP Neural Network [J]. Journal of Weifang University, 2011, 11(6): 90—93.
- [11] 初广丽, 张伟, 王延杰, 等. 基于机器视觉的水果采摘机器人目标识别方法[J]. 中国农机化学报, 2018, 39(2): 83—88. Chu Guangli, Zhang Wei, Wang Yanjie, et al. A method of fruit picking robot target identification based on machine vision [J]. Journal of Chinese Agricultural Mechanization, 2018, 39(2): 83—88.
- [12] 周云成, 许童羽, 邓寒冰, 等. 基于双卷积链 Fast R — CNN 的番茄关键器官识别方法[J]. 沈阳农业大学学报, 2018, 49(1): 65—74. Zhou Yuncheng, Xu Tongyu, Deng Hanbing, et al. Recognition method of tomato key organs based on dual convolution Fast R — CNN [J]. Journal of Shenyang Agricultural University, 2018, 49(1): 65—74.
- [13] 吕小莲, 吕小荣, 倪受春. 田间采摘西红柿计算机识别方法的设计研究[J]. 传感技术学报, 2010, 23(1): 82—86. Lv Xiaolian, Lv Xiaorong, Ni Shouchun. Design and research of the computer identification method of picking tomatoes in field [J]. Chinese Journal of Sensors and Actuators, 2010, 23(1): 82—86.
- [14] He K, Gkioxari G, Dollar P, et al. Mask-RCNN [C]. 2017 IEEE International Conference on Computer Vision (ICCV). IEEE Computer Society, 2017.
- [15] Yu Y, Zhang K, Yang L, et al. Fruit detection for strawberry harvesting robot in non-structural environment based on Mask-RCNN [J]. Computers and Electronics in Agriculture, 2019, 163: 104846.
- [16] 欧攀, 路奎, 张正, 等. 基于 Mask RCNN 的目标识别与空间定位[J]. 计算机测量与控制, 2019, 27(6): 172—176. Ou Pan, Lu Kui, Zhang Zheng, et al. Target recognition and spatial localization based on mask RCNN [J]. Computer Measurement & Control, 2019, 27(6): 172—176.
- [17] 吴金亮, 王港, 梁硕, 等. 基于 Mask-RCNN 的舰船目标检测研究[J]. 无线电工程, 2018, 48(11): 947—952. Wu Jinliang, Wang Gang, Liang Shuo, et al. Ship object detection based on mask R-CNN [J]. Radio Engineering, 2018, 48(11): 947—952.
- [18] Lin T Y, Maire M, Belongie S, et al. Microsoft COCO: Common Objects in Context [J]. 2014.
- [19] Kamilaris A, Prenafeta-Boldu F X. Deep learning in agriculture: A survey [J]. Computers and Electronics in Agriculture, 2018, 147: 70—90.
- [20] 张云, 许江淳, 王志伟, 等. 基于机器视觉的棉花异性纤维检测技术优化研究[J]. 中国农机化学报, 2018, 39(9): 61—65. Zhang Yun, Xu Jiangchun, Wang Zhiwei, et al. Optimization of cotton heterosexual detection technology based on machine vision [J]. Journal of Chinese Agricultural Mechanization, 2018, 39(9): 61—65.
- [21] 李林升, 曾平平. 改进深度学习框架 Faster-RCNN 的苹果目标检测[J]. 机械设计与研究, 2019, 35(5): 24—27. Li Linsheng, Zeng Pingping. Apple target detection based on improved Faster-RCNN framework of deep learning [J]. Machine Design & Research, 2019, 35(5): 24—27.
- [22] 邓立苗, 唐俊, 马文杰. 基于图像处理的玉米叶片特征提取与识别系统[J]. 中国农机化学报, 2014, 35(6): 72—75, 79. Deng Limiao, Tang Jun, Ma Wenjie. Feature extraction and recognition system of maize leaf based on image processing [J]. Journal of Chinese Agricultural Mechanization, 2014, 35(6): 72—75, 79.
- [23] 闫建伟, 赵源, 张乐伟, 等. 改进 Faster-RCNN 自然环境下识别刺梨果实[J]. 农业工程学报, 2019, 35(18): 143—150. Yan Jianwei, Zhao Yuan, Zhang Lewei, et al. Recognition of Rosa roxbunghii in natural environment based on improved Faster RCNN [J]. Transactions of the Chinese Society of Agricultural Engineering, 2019, 35 (18): 143—150.
- [24] 陈礼鹏, 穆龙涛, 刘浩洲, 等. 基于猕猴桃果萼图像的多目标果实识别方法[J]. 计算机工程与设计, 2018, 39(6): 1738—1744. Chen Lipeng, Mu Longtao, Liu Haozhou, et al. Multi-target fruit recognition method based on kiwifruit calyx image [J]. Computer Engineering and Design, 2018, 39 (6): 1738—1744.

(下转第 210 页)

Inversion of soil moisture in corn field based on thermal infrared remote sensing image

Guo Hui, Bu Xiaodong, Huang Kejing, Liu Yingchao

(Beijing Vocational College of Agriculture, Beijing 102208, China)

Abstract: Soil moisture monitoring is essential for tracking crop growth. The purpose of this study is to effectively retrieve the field soil water content during the main growth period of the corn. It was used thermal infrared remote sensing images acquired by the drone platform as a data source, and was based on the thermal inertia method to retrieve soil moisture at the field scale. By establishing a linear regression model between soil thermal inertia and soil water content, it performed model accuracy verification in test fields. The results show that when the soil moisture content is retrieved based on the thermal inertia method in an actual farmland environment, as the water treatment level increases, its inversion accuracy increases first and then decreases. The accuracy verification results of the model inversion of soil water content under different water treatments are: $R^2 = 0.71$, $RMSE = 3.09\%$. The research results show that the thermal inertia method has higher accuracy of soil moisture retrieval, which provides a reference for monitoring soil moisture in the field based on UAV thermal infrared remote sensing.

Keywords: soil moisture; corn; thermal infrared remote sensing; inversion

(上接第 190 页)

[25] 吴雪梅, 唐仙, 张富贵, 等. 基于 K-means 聚类法的茶叶嫩芽识别研究[J]. 中国农机化学报, 2015, 36(5): 161—164, 179.

Wu Xuemei, Tang Xian, Zhang Fugui, et al. Tea buds image identification based on lab color model and K-means clustering [J]. Journal of Chinese Agricultural Mechanization, 2015, 36(5): 161—164, 179.

[26] 王宜全, 吴扬东, 祁兵. 基于 Faster RCNN 的间作类农田无人机喷雾方法研究[J]. 中国农机化学报, 2019, 40(6): 76—81.

Wang Yiquan, Wu Yangdong, Qi Bing. Study on UAV spray method of intercropping farmland based on Faster RCNN [J]. Journal of Chinese Agricultural Mechanization, 2019, 40(6): 76—81.

Experimental research on orchard fruit detection based on fast convolutional neural network

Zhang Lei¹, Jiang Junsheng², Li Xinyu³, Song Jian², Xie Fuxiang²

(1. College of Mechanical and Electronic Engineering, Shandong University of Science and Technology, Qingdao, 266590, China;

2. College of Mechanical and Vehicle Engineering, Weifang University, Weifang, 261061, China;

3. Shandong Institute of Industrial Technicians, Weifang, 261061, China)

Abstract: In recent years, research on automatic fruit recognition and detection based on digital image processing and machine learning algorithms has become more and more mature. Aiming at the shortcomings of the traditional detection method that it is difficult to meet the real-time requirements, a fast fruit detection model based on Faster-RCNN is adopted. The model is composed of Convolutional Neural Network (CNN) and Region Proposal Network (RPN). First, CNN performs convolution and pooling operations to extract features, then RPN selects candidate regions, shares the parameters through the fully connected layer of the network, and classifies by target recognition. The predictor and the bounding box prediction regress or obtain multiple prediction frames that may contain the target, and finally select the prediction frame with the highest accuracy by non-maximum suppression to complete the target detection. The three fruits of peaches, apples and oranges were tested respectively, using the transfer learning method, using two pre-trained deep neural network models ZFnet and VGG16, and adjusting the parameters of Dropout and the number of candidate regions through the training of the data set. Network tuning. Detect and analyze the detection effect of the model under different layout patterns of fruits. The test results show that when the value of Dropout is 0.5 or 0.6, and the number of candidate regions is 300, the result has the best accuracy and recall rate. In ZFnet network, the average accuracy of apple is 92.70%, that of peach is 90.00%, and that of orange is 89.72%. In VGG16 network, the average precision of apple, peach and orange is 94.17%, 91.46% and 90.22%. The image processing speed of ZFnet and VGG16 reaches 17 fps and 7 fps respectively, which can achieve real-time detection of fruits purpose.

Keywords: fruit detection; deep learning; agricultural robot; neural networks; transfer learning