

doi:10.6041/j.issn.1000-1298.2020.12.023

复杂大田场景中麦穗检测级联网络优化方法

谢元澄¹ 何超¹ 于增源¹ 沈毅¹ 姜海燕^{1,2} 梁敬东¹

(1. 南京农业大学信息科学技术学院, 南京 210095; 2. 南京农业大学国家信息农业工程技术中心, 南京 210095)

摘要: 单位种植面积的麦穗数量是评估小麦产量的关键农艺指标之一。针对农田复杂场景中存在的大量麦芒、卷曲麦叶、杂草等环境噪声、小尺寸目标和光照不均等导致的麦穗检测准确度下降的问题,提出了一种基于深度学习的麦穗检测方法(FCS R-CNN)。以 Cascade R-CNN 为基本网络模型,通过引入特征金字塔网络(FPN)融合浅层细节特征和高层丰富语义特征,通过采用在线难例挖掘(Online hard example mining, OHEM)技术增加对高损失样本的训练频次,通过 IOU(Intersection over union) 阈值对网络模型进行阶段性融合,最后基于圆形 LBP 纹理特征训练一个 SVM 分类器,对麦穗检测结果进行复验。大田图像测试表明,FCS R-CNN 模型的检测精度达 92.9%,识别单幅图像平均耗时为 0.357 s,平均精度为 81.22%,比 Cascade R-CNN 提高了 21.76 个百分点。

关键词: 麦穗计数; 目标检测; Cascade R-CNN; IOU 级联; 复杂场景

中图分类号: TP391.41 文献标识码: A 文章编号: 1000-1298(2020)12-0212-08

OSID:



Optimization Method for Cascade Network of Wheat Ear Detection in Complex Filed Scene

XIE Yuancheng¹ HE Chao¹ YU Zengyuan¹ SHEN Yi¹ JIANG Haiyan^{1,2} LIANG Jingdong¹

(1. College of Information Sciences and Technology, Nanjing Agricultural University, Nanjing 210095, China

2. National Engineering and Technology Center for Information Agriculture, Nanjing Agricultural University, Nanjing 210095, China)

Abstract: The number of wheat ears per planting area is one of the key agronomic index to evaluate wheat yield. In the field scene, there are usually great differences in the shape, size and posture of wheat ears, and there are serious occlusion between leaves and ears and between ears and ears. At the same time, wheat awn, curly wheat leaves, weeds and uneven illumination introduced a lot of background interference. These complex factors led to a high false detection rate in traditional methods based on color and texture features. The detection method based on deep learning has a high missed detection rate for small-size rice ear images in practical application. To solve these problems, a wheat ear detection method FCS R-CNN based on deep learning was proposed. Taking Cascade R-CNN as the basic network model, a feature pyramid network (FPN) was introduced to fuse shallow detailed features and high-level rich semantic features, and online hard example mining (OHEM) technology was added to increase training frequency for high-loss samples, the network was fused by the IOU threshold. Finally, a SVM classifier was trained based on the circular LBP texture features to carry out the reinspection of wheat ear detection results to further reduce the detection error. In the wheat field image test, the detection accuracy of FCS R-CNN model reached 92.9%, the average precision (AP) was 81.22%, the average time to identify a single image was 0.357 s, and the AP was 21.76 percentage points higher than that of the original Cascade R-CNN. The results showed that the proposed method had better detection results for wheat ear detection in complex scenes, which provided a new idea for wheat yield estimation.

Key words: wheat ear count; object detection; Cascade R-CNN; IOU cascade; complex scene

收稿日期: 2020-03-17 修回日期: 2020-05-19

基金项目: 国家重点研发计划项目(2016YFD0300607)、江苏省重点研发计划(现代农业)重点项目(BE2019383)和中央高校基本科研业务费自主创新重点项目(KYZ201550、KYZ201548)

作者简介: 谢元澄(1976—),男,副教授,主要从事模式识别与生物信息学研究,E-mail: xieych@njau.edu.cn

通信作者: 姜海燕(1967—),女,教授,主要从事作物系统建模与智能计算研究,E-mail: jianghy@njau.edu.cn

0 引言

小麦是我国的主要粮食作物,其产量关系到国家的粮食安全^[1]。对小麦产量的精确估计可为政府在农村土地政策、粮食价格等方面的宏观调控提供支撑^[2]。单位面积麦穗数直接体现小麦种植密度,是决定小麦产量的关键因素,其数据主要在田间获取^[3],快速准确地测量这一指标可以大幅提高监测小麦生长性状的效率,从而为小麦产量及其他表型性状提供早期预测,有利于优化和加速育种进程。近年来,传统图像分割方法广泛应用于农业领域,基于形状^[4]、颜色^[5-9]或者纹理^[5-7]特征对小麦 RGB 图像进行特征提取,并通过机器学习方法实现麦穗检测计数,在一定程度上提高了小麦估产效率。以上方法仅适用于灌浆期^[8]等小麦生长的部分时期,这些时期小麦图像各区域像素之间具有明显的差异性。当麦穗存在粘连或者长势不一时,仅仅依靠基于人工设定并提取特征的方法,其鲁棒性较低。此外,特征提取方法流程也较为复杂,且对图像采集设备具有较高的要求^[10-12]。

计算能力的进步以及大量标记图像的可用性促进了基于卷积神经网络(Convolutional neural network, CNN)的机器学习方法在计算机视觉领域的应用和发展^[13]。在稻穗检测计数研究中^[14-16], CNN 具有良好的检测效果,但存在流程复杂、算法耗时且受噪声干扰影响较大的问题。在将深度学习应用于麦穗检测计数的研究中,张领先等^[17]利用 CNN 结合非极大值抑制(Non-maximum suppression, NMS)的方法进行冬小麦麦穗检测计数,但需要预先对小麦图像进行手动分割,且对麦穗重叠、目标较小等复杂情况下的检测未作说明。另外,使用 Faster R-CNN 等单一模型对麦穗进行检测时^[18-19],由于训练过程中 IOU (Intersection over union) 阈值设定不规范以及训练和预测时的 IOU 阈值不匹配,容易造成模型过拟合,导致对小麦图像细节特征提取不足、检测性能大幅下降。

在复杂大田环境中,小麦麦穗、麦叶与茎秆之间的颜色混叠程度极高,叶片、杂草、麦芒与麦穗之间存在严重的遮挡现象,光照强度不均匀易形成反光和逆光,同时还存在部分尺寸较小的目标隐藏在背景相似的环境中的问题。为有效解决复杂场景下的麦穗检测计数问题,本文提出一种基于深度卷积神经网络的高效麦穗自动计数方法(FCS R-CNN),以 Cascade R-CNN^[20]为基本检测模型,加入特征金字塔网络,通过 IOU 阈值级联的方式进行阶段融合,并在训练过程中对困难样本进行重学习。利用

改进的支持向量机(Support vector machine, SVM)^[21]分类器对所有检出框进行二分类,将 FCS R-CNN 方法与现场人工计数进行比较,以评估该方法在真实场景下的适用性。

1 样本采集

试验所用小麦图像于 2018 年 6 月 4 日在南京市江宁区(118°55'28.9236" E, 31°53'3.3432" N) 4 块正常条播小麦种植园获取,小麦品种为“宁麦 8 号”,小麦处于成熟期。图像采集设备为美国苹果公司生产的 iPhone 7 型智能手机,1 200 万像素。拍摄期间天气晴朗,没有大风,图像采集于 11:00—14:00 和 16:00—19:00 两个时间段,光照强度不同,且拍摄角度随机,大部分图像采集角度小于等于 45°,涵盖了田间麦穗计数作业时所有可能的环境条件。拍摄位置距小麦冠层顶部约 1 m,每块小麦田约获取 300 幅图像,共获取 1 079 幅分辨率为 4 032 像素×3 024 像素的小麦图像,部分小麦图像如图 1 所示。

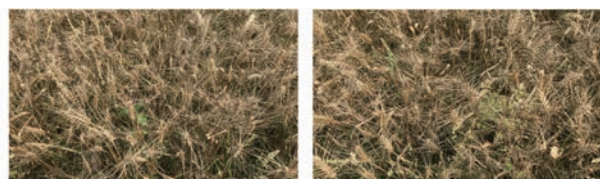


图 1 大田环境下小麦图像

Fig. 1 Wheat images in field environment

将采集到的 1 079 幅小麦图像划分为训练集 863 幅,测试集 216 幅,并对数据集进行 5 折交叉验证。然后使用开源图像标注工具 LabelImg 进行人工图像标注。对标注完成的图像数据进行分析,共标注麦穗边界框(Bounding box, Bbox) 56 182 个,其中训练集 45 011 个,测试集 11 171 个。然后将原始小麦图像尺寸修改为 1 333 像素×800 像素,并输入到 CNN 网络中提取特征,生成特征映射图。

2 级联网络优化

2.1 模型整体框架与改进

第 1 阶段采用改进型 Cascade R-CNN 作为基本检测模型,第 2 阶段对模型中 S0 阶段与 S3 阶段进行基于 IOU 阈值的级联提升检出率,最终接入一个基于圆形 LBP 纹理特征训练的 SVM 分类器,滤除误检框,得到最终的算法框架 FCS R-CNN(图 2)。

Cascade R-CNN 是基于 Faster R-CNN 的网络框架(图 2)。Faster R-CNN 是由 HE 等^[22]提出的目标检测网络,作为 R-CNN^[23]与 Fast R-CNN^[24]网络的改进版本,在结构上抛弃了传统的滑动窗口

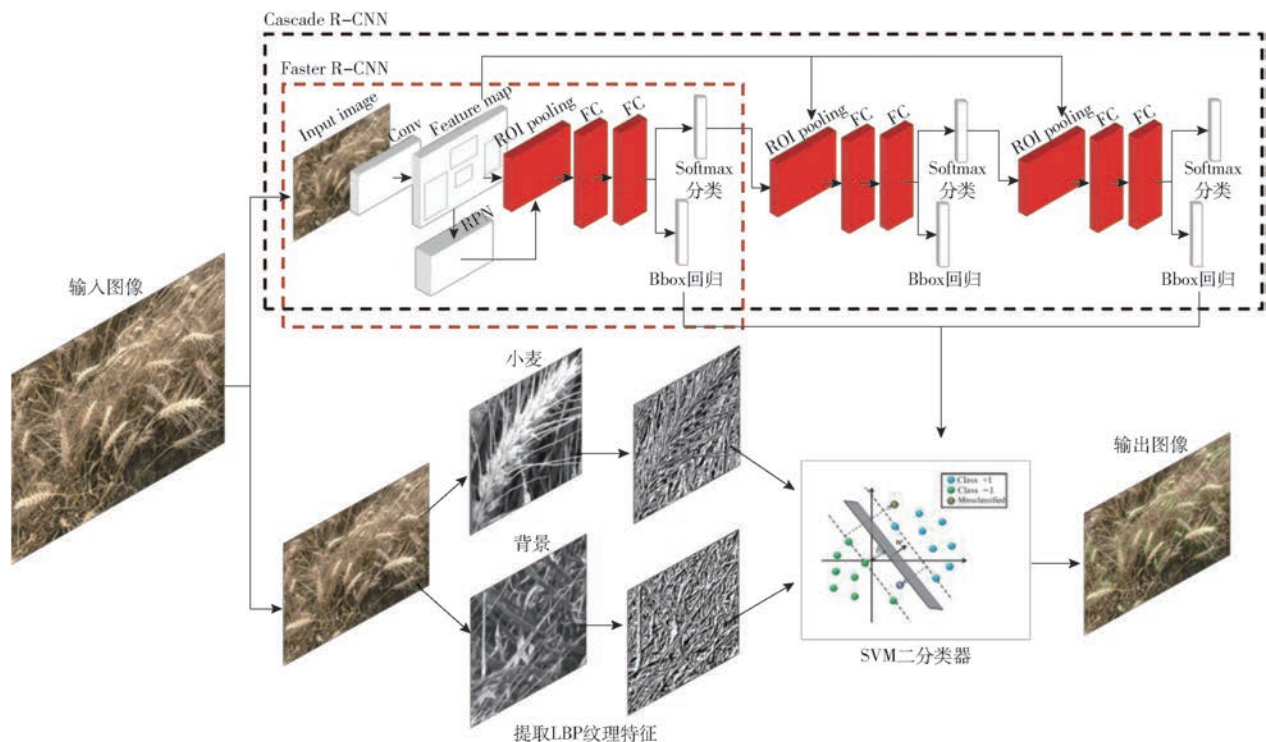


图2 FCS R-CNN网络模型结构图

Fig. 2 Structure diagram of FCS R-CNN

和选择性搜索(Selective search)候选框生成方法,引入区域生成网络(Region proposal network, RPN),并与Fast R-CNN共享特征提取网络,最终在检测速度和平均精度上有较大提高。Cascade R-CNN对Faster R-CNN网络中的RPN网络生成的包围框使用三级级联,IOU阈值分别设定为0.5、0.6、0.7,级联的过程不断改变候选框的分布,并且通过调整阈值的方式重采样,采样后的每个检测器都可以更好地适应当前阈值,提升检测性能。

为使网络能够克服复杂背景信息的干扰,针对Cascade R-CNN中的Faster R-CNN子模块,选取深度残差神经网络(Residual neural network, ResNet)^[25]作为基本的特征提取网络,通过引入特征金字塔网络(Feature pyramid network, FPN)^[26]和在线难例挖掘(Online hard example mining, OHEM)^[27]技术进行改进,改进结果如图3所示。

另外,对Cascade R-CNN模型进行部分优化,针对原始网络中采用的批量归一化(Batch normalization, BN)方法^[28]受到Batch Size影响导致网络性能下降的问题,引入群组归一化(Group normalization, GN),将通道划分成组,然后在每组中计算用于归一化的均值和方差。另外,针对相邻麦穗干扰导致的漏检问题,选取Soft-NMS算法代替传统的非极大值压缩(Non-maximum suppression, NMS)^[29]方法对特征图中提取的候选感兴趣区域(Regions of interest, ROI)进行修剪,获取候选框并

输入ROI pooling层进行池化操作。为了区别于最终的检测模型FCS R-CNN,并方便在试验结果中对FCS R-CNN模型效果进行递进的验证,将基于Cascade R-CNN优化后的第1阶段模型称作CR-CNN。

2.2 特征金字塔网络

引入特征金字塔网络提升模型对视觉场景中尺寸较小目标的检测精度。LIN等^[26]提出的特征金字塔网络作为一种高效的CNN特征提取方法,由自下而上路径、自上而下路径以及横向连接组成,主要解决物体检测时多尺度情形下存在的小目标漏检问题,通过改变网络连接,在不增加原有模型计算量的前提下可大幅度提升小物体检测的性能。如图3所示,FPN使用一个 1×1 的卷积过滤器将C5的通道深度降至256维,得到特征图M5。接着应用一个 3×3 的卷积得到特征图P5, P5正是用于目标预测的第一个特征映射。向上进行最近邻下采样($\times 2$)得到P6,然后沿着自顶向下的路径往下,对之前的特征提取层进行最近邻上采样($\times 2$)。同时,FPN对自底向上通路中的相应特征映射应用 1×1 卷积,然后进行相加。最后同样应用 3×3 卷积得到目标检测的特征映射。这一过滤器减轻了上采样的混叠效应。

2.3 在线难例挖掘

对于复杂场景中的麦穗检测而言,存在大量特征相似的正负样本,这些样本在训练过程中造成了

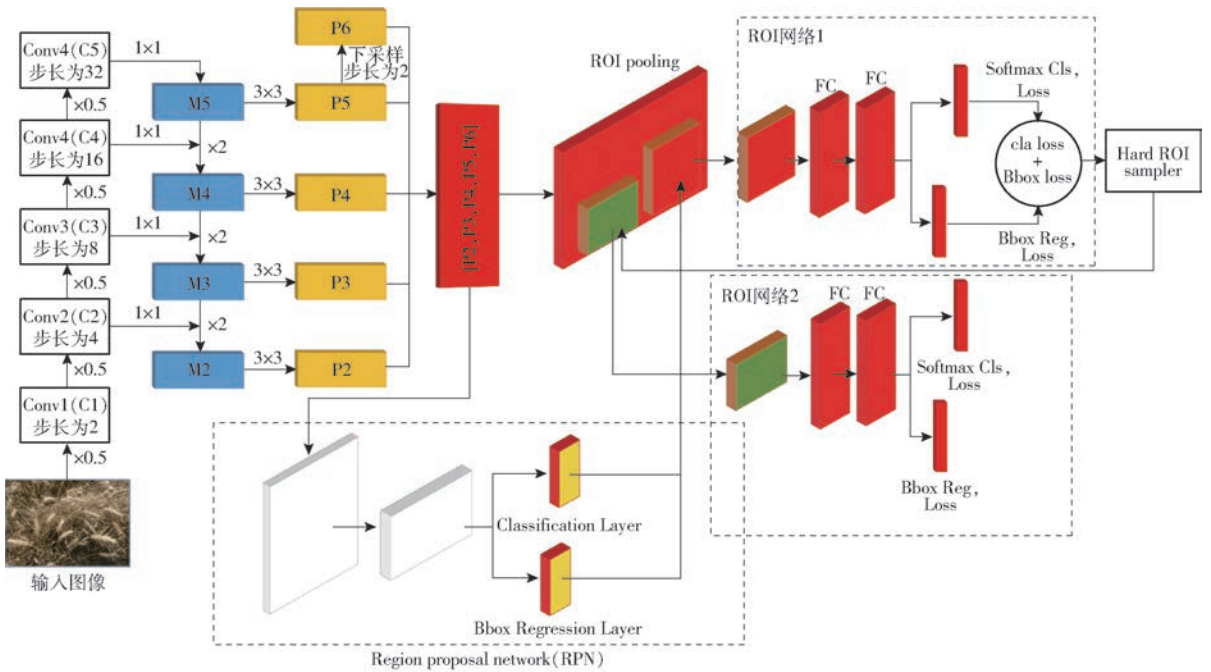


图 3 改进的 Faster R - CNN 子模块

Fig. 3 Improved Faster R - CNN submodule

较大干扰,导致模型预测误差放大,因此引入在线难例挖掘技术,针对性地解决这一问题。OHEM 在训练过程中会自动选择那些困难样本加入训练,这样使得训练更加有效和快速,同时能提升网络的性能。如图 3 所示,将 Faster R - CNN 分为 2 个阶段,第 1 阶段包括通过特征提取网络对输入图像进行特征提取并输入 ROI pooling 层,该阶段对输入图像进行特征提取并收集输入的特征图像,提取目标特征图 (Proposal feature maps)。第 2 阶段包括 ROI pooling 层之后的所有层。将第 2 阶段作为一个单独的 ROI 网络,接着将该网络分成 2 个网络,ROI 网络 1 是只读的,对所有样本进行前向传播得到所有 ROI 的损失值,接着难例取样器 (Hard ROI sampler) 根据损失值排序,利用 NMS 选出 K 个难例 (Hard example),并把这些难例作为 ROI 网络 2 的输入,ROI 网络 2 对 ROI 进行进一步的前向和反向传播,计算 K 个难例的损失值,并将梯度反向传播,从而更新网络参数。通过损失值提高难例的比例,网络因此可以对难例进行多次训练,提升模型准确率。

3 试验与结果分析

3.1 试验设置

对采集的 1 079 幅小麦数据进行训练。软件环境为 pytorch1.0.0、CUDA 9.0、CUDNN7.6.1、Ubuntu 16.04 LTS。训练的硬件环境为 Intel Xeon E5 - 2650 v3 CPU,128 GB 存储器,NVIDIA GTX Titan XP * 2 GPU。利用随机梯度下降法 (Stochastic gradient descent,SGD) 训练网络,初始学习率 (Learning rate)

设置为 0.001,动量因子 (Momentum) 设置为 0.9,权重衰减系数 (Weight decay) 设置为 0.000 1,迭代次数 (Epoch) 设置为 200。

3.2 麦穗检测试验

为验证 FCS R - CNN 模型在复杂场景下的麦穗检测效果,针对 FCS R - CNN 中的 2 个阶段,分别设计两部分试验。试验 1 对比 Cascade R - CNN 和 C R - CNN 两种算法的检测结果,验证级联网络优化方法对 Cascade R - CNN 模型检测性能的提升效果。试验 2 展示 FCS R - CNN 的部分麦穗检测结果,在 C R - CNN 基础上,利用两阶段级联并串联 SVM 分类器的结构改进方法,验证模型的性能提升效果。

试验 1 随机选取两幅测试集图像 (图 A、B) 分别输入到 Cascade R - CNN 与 C R - CNN 网络,进行麦穗检测效果对比,具体的识别结果如图 4 所示。对图 4c、4d 的局部进行放大便于观察。通过图 4 可以明显看出,C R - CNN 检出大量 Cascade R - CNN 漏检的麦穗目标 (图中黄色框标记)。图 4c 中 4 号、图 4d 中 1 号蓝色框标出的是 C R - CNN 误检的目标,该目标与正样本拟合程度较高,但是属于麦芒,因此属于假阳性 (False positive, FP)。利用 Cascade R - CNN 网络检测验证集,共检测 8 523 个麦穗目标,并存在大量假阴性 (False negative, FN) 目标,图 5 给出了部分假阳性和假阴性样例,其中假阴性目标周围大多有大量麦芒、麦叶等遮挡物。对于漏检目标 (假阴性),将 NMS 改为 Soft - NMS 来减少相邻目标重叠导致的冗余框误删,可以提

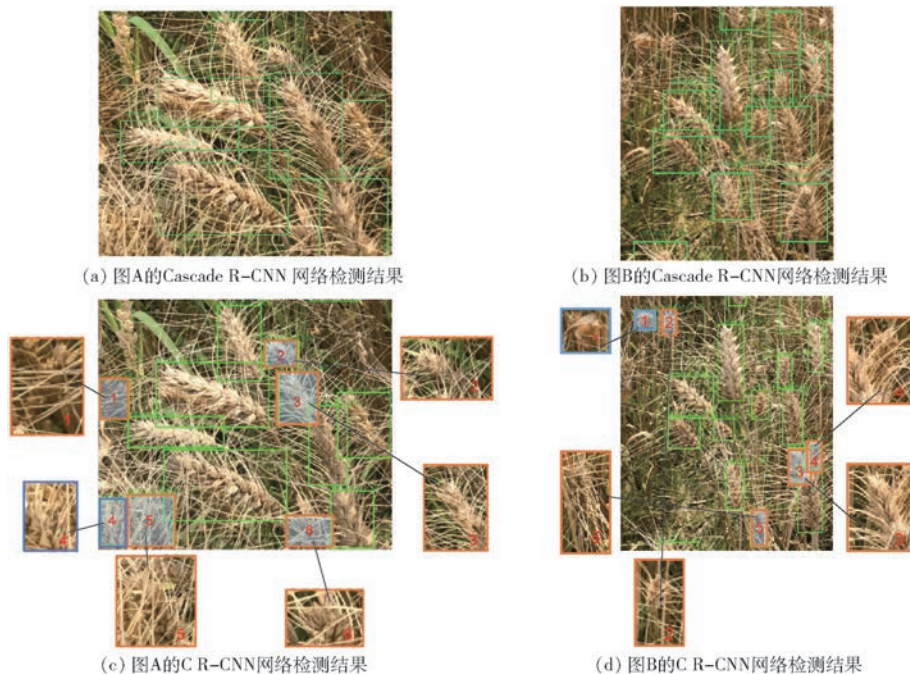


图4 Cascade R - CNN 与 C R - CNN 检测结果对比

Fig. 4 Comparison of Cascade R - CNN and C R - CNN detection results

升相邻较近目标的检测率,两种方法对比结果如图6所示。



图5 部分假阳性和假阴性的样例

Fig. 5 Examples of partial false positives and false negatives

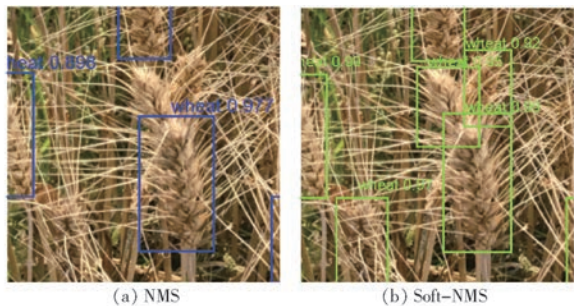


图6 Soft - NMS 与 NMS 算法结果对比

Fig. 6 Comparison of Soft - NMS and NMS algorithm results

试验2针对上述广域复杂环境下生长的小麦,利用FCS R - CNN对麦穗进行检测。图7展示了部分检测结果,图中绿框表示C R - CNN的检测结果,蓝框表示阶段级联后增加的检出框。当麦穗密度较高时,大量麦芒遮挡周围麦穗,也存在部分麦穗重叠现象,可以看出,FCS R - CNN能够实现多种复杂情

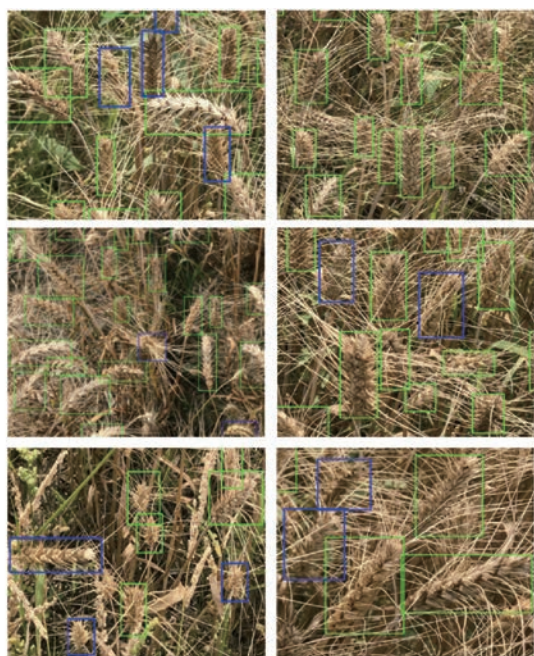


图7 FCS R - CNN 部分检测结果

Fig. 7 Partial detection results of FCS R - CNN

形下的麦穗检测,并且相对于C R - CNN有了更进一步的性能提升。

3.3 结果分析

对图4中两种算法结果的对比分析可知,复杂场景是影响麦穗检测精度、造成假阴性目标产生的主要原因,具体包括:①小麦长势不一,大量麦芒、麦叶的存在严重阻碍了麦穗识别(图4c中1、5、6号框,图4d中4、5号框)。②背景中存在与麦穗颜色相近的杂草等类似物,试验前未对杂草进行单独标注,引起干扰,从而影响麦穗检测。③部分目标由于

在视觉场景中显示尺寸较小且密集,在背景相似的环境中,此类小尺寸目标极易受到环境噪声的干扰(图 4c 中 2 号框,图 4d 中 2、5 号框)。④麦穗之间存在相互重叠遮盖现象,不能正确识别(图 4c 中 3 号框,图 4d 中 3、4 号框)。

对图 7 中 FCS R-CNN 的检测结果进行分析可知,C R-CNN 相对于 FCS R-CNN 的漏检麦穗(图中蓝色框),是由 Cascade R-CNN 本身存在缺陷导致,这可能与 IOU 阈值的提高有关。Cascade R-CNN 由 3 个检测模型组成,每个检测模型都基于不同 IOU 阈值来区分正负样本,前一个检测模型的输出为后一个检测模型的输入,因此,在阈值提高的过程中,阈值较低时或处于两个阈值中间的正样本会出现遗漏,导致网络在训练过程中遗失掉部分候选区域(Region proposal)。

为了验证 FCS R-CNN 模型与单一网络模型的性能差异,选取相同样本数据集分别用 FCS R-CNN、C R-CNN、Cascade R-CNN、Faster R-CNN 4 种模型进行训练并将训练后的模型应用于真实环境下的大田麦穗计数中,对模型训练结果使用准确率(Accuracy)、精度(Precision)、召回率(Recall)和平均精度(Average precision, AP)对试验结果进行综合评价,计算公式为

$$A_{cc} = \frac{T_p + F_N}{T_p + F_N + T_N + F_p} \times 100\% \quad (1)$$

$$P = \frac{T_p}{T_p + F_p} \times 100\% \quad (2)$$

$$R = \frac{T_p}{T_p + F_N} \times 100\% \quad (3)$$

$$A_p = \int_0^1 P(R) dR \quad (4)$$

式中 A_{cc} ——准确率

P ——精度

R ——召回率

A_p ——平均精度

T_p ——真阳性,图像中正确检测出的目标麦穗数量

F_p ——假阳性,图像中将背景误识别为麦穗的数量

T_N ——真阴性,图像中检测出的负样本数量

F_N ——假阴性,图像中漏检的麦穗数量

表 1 给出了 4 种算法在相同测试集下的检测结果。IOU 为 0.5 时 FCS R-CNN 的平均精度取得最大值,比 Cascade R-CNN 高 29.98 个百分点,比 Faster R-CNN 高 48.09 个百分点。图 8 给出了其中 3 种算法精度与召回率的关系曲线,用来权衡准

确率和召回率的关系。由图 8 可以看出,FCS R-CNN 在召回率小于 80% 时精度维持在较高水平,高于另外两种算法,进一步证明 FCS R-CNN 性能上的优越性。FCS R-CNN 的图像检测速度达到 3 f/s,能够满足大田麦穗自动计数的需求。

表 1 4 种算法性能对比

Tab. 1 Performance comparison of four algorithms

算法	IOU 为 0.5 时 平均精度/%	平均精度/%	单幅图像 检测时间/s
Faster R-CNN	42.01	45.23	0.226
Cascade R-CNN	60.12	59.46	0.323
C R-CNN	79.91	70.57	0.331
FCS R-CNN	90.10	81.22	0.357

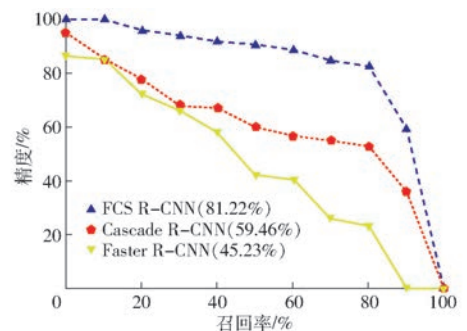


图 8 FCS R-CNN 精度与召回率关系曲线

Fig. 8 Relationship curves of precision and recall of FCS R-CNN

Cascade R-CNN、C R-CNN 与 FCS R-CNN 3 种模型的麦穗计数结果如表 2 所示。从表 2 中可以看出,应用 FCS R-CNN 模型的麦穗统计精度普遍高于 95%,在麦穗较少的图像中能实现全检。216 幅小麦图像中,共标注 11 171 个麦穗,共检出 10 373 个麦穗,检测精度达到 92.9%。

不同算法统计结果表明,单级联模型的识别结果容易受到 IOU 的影响,这是由于 IOU 提升会导致部分正样本遗漏,造成正负样本失衡,从而影响麦穗识别效果。面对复杂场景中麦芒、枯叶、杂草等环境噪声的干扰,会出现大量漏检的麦穗,通过模型阶段级联的方法,可以进一步降低漏检率。

4 结论

(1) 针对复杂场景中麦穗检测精度较低的问题,提出一种基于深度卷积神经网络的麦穗计数方法(FCS R-CNN)。试验表明,在单一网络模型的基础上,对网络进行结构改进可以有效提高对复杂目标的检测性能。FPN 的加入提升了对浅层细节特征的利用率,OHEM 技术可以降低复杂背景信息对目标分割的干扰,并提高对小目标的检测精度;FCS R-CNN 在大田环境下对麦穗的检测精度达到

表2 人工计数与算法计数的麦穗统计对比

Tab.2 Comparison of manual counting and algorithm counting of wheat ear

编号	人工 计数/个	Cascade R-CNN 计数/个	Cascade R-CNN 统计误差/%	C R-CNN 计数/个	C R-CNN 统计误差/%	FCS R-CNN 计数/个	FCS R-CNN 统计误差/%	FCS R-CNN 统计精度/%
1	35	28	20.00	34	2.86	32	8.57	91.43
2	65	61	6.15	59	9.23	63	3.07	96.93
3	98	85	13.27	90	8.16	96	2.04	97.96
4	94	86	8.51	87	7.45	90	4.26	95.74
5	82	77	6.10	82	0	83	-1.21	98.79
6	97	81	16.49	81	16.50	89	8.25	91.75
7	90	72	20.00	80	11.11	90	0	100
8	54	42	22.22	45	16.67	48	11.11	88.89
9	55	51	7.27	54	1.81	59	-7.27	92.73
10	52	48	7.69	45	13.46	51	1.92	98.08
11	50	43	14.00	50	0	52	-4.00	96
12	45	39	13.33	40	11.11	46	-2.22	97.78
13	108	90	16.67	95	12.04	93	13.89	86.11
14	57	49	14.04	49	14.04	54	5.26	94.74
15	12	7	41.67	9	25.00	12	0	100
平均值			15.16		9.96		2.91	95.12

92.9%,单幅麦穗图像的检测时间为0.357 s,具有较强的实时性和鲁棒性。

(2)FCS R-CNN通过两阶段级联并串联SVM分类器对结果进行复验,能够有效解决因IOU阈值

设定不规范而导致模型准确率下降的问题。算法的平均精度达到了81.22%,比Cascade R-CNN模型高21.76个百分点。IOU为0.5时平均精度为90.10%。

参 考 文 献

- [1] LIU H, WANG Z H, YU R, et al. Optimal nitrogen input for higher efficiency and lower environmental impacts of winter wheat production in China[J]. Agriculture Ecosystems & Environment, 2016, 224: 1-11.
- [2] 欧文浩, 苏伟, 薛文振, 等. 基于HJ-1卫星影像的三大农作物估产最佳时相选择[J]. 农业工程学报, 2010, 26(11): 176-182, 385.
OU Wenhao, SU Wei, XUE Wenzhen, et al. Selection of optimum phase for yield estimation of three major crops based on HJ-1 satellite images[J]. Transactions of the CSAE, 2010, 26(11): 176-182, 385. (in Chinese)
- [3] PASK A, PIETRAGALLA J, MULLAN D, et al. Physiological breeding II: a field guide to wheat phenotyping[M]. Mexico, DF: CIMMYT, 2012.
- [4] 李毅念, 杜世伟, 姚敏, 等. 基于小麦群体图像的田间麦穗计数及产量预测方法[J]. 农业工程学报, 2018, 34(21): 185-194.
LI Yinian, DU Shiwei, YAO Min, et al. Method for wheatear counting and yield predicting based on image of wheatear population in field[J]. Transactions of the CSAE, 2018, 34(21): 185-194. (in Chinese)
- [5] 范梦扬, 马钦, 刘峻明, 等. 基于机器视觉的大田环境小麦麦穗计数方法[J/OL]. 农业机械学报, 2015, 46(增刊): 234-239.
FAN Mengyang, MA Qin, LIU Junming, et al. Counting method of wheatear in field based on machine vision technology[J/OL]. Transactions of the Chinese Society for Agricultural Machinery, 2015, 46(Supp.): 234-239. http://www.j-csam.org/jcsam/ch/reader/view_abstract.aspx?flag=1&file_no=2015S038&journal_id=jcsam. DOI: 10.6041/j.issn.1000-1298.2015.S0.038. (in Chinese)
- [6] 刘涛, 孙成明, 王力坚, 等. 基于图像处理技术的大田麦穗计数[J/OL]. 农业机械学报, 2014, 45(2): 282-290.
LIU Tao, SUN Chengming, WANG Lijian, et al. In-field wheat ear counting based on image processing technology[J/OL]. Transactions of the Chinese Society for Agricultural Machinery, 2014, 45(2): 282-290. http://www.j-csam.org/jcsam/ch/reader/view_abstract.aspx?flag=1&file_no=20140247&journal_id=jcsam. DOI: 10.6041/j.issn.1000-1298.2014.02.047. (in Chinese)
- [7] ZHOU Chengquan, DONG Liang, YANG Xiaodong, et al. Wheat ears counting in field conditions based on multi-feature optimization and TWSVM[J]. Frontiers in Plant Science, 2018, 9: 1024.
- [8] 刘哲, 黄文准, 王利平. 基于改进K-means聚类算法的大田麦穗自动计数[J]. 农业工程学报, 2019, 35(3): 174-181.
LIU Zhe, HUANG Wenzhun, WANG Liping. Field wheat ear counting automatically based on improved K-means clustering algorithm[J]. Transactions of the CSAE, 2019, 35(3): 174-181. (in Chinese)
- [9] 杜颖, 蔡义承, 谭昌伟, 等. 基于超像素分割的田间小麦穗数统计方法[J]. 中国农业科学, 2019, 52(1): 21-33.

- DU Ying, CAI Yicheng, TAN Changwei, et al. Field wheat ears counting based on superpixel segmentation method[J]. *Scientia Agricultura Sinica*, 2019, 52(1): 21–33. (in Chinese)
- [10] ZHU Y, CAO Z, LU H, et al. In-field automatic observation of wheat heading stage using computer vision[J]. *Biosystems Engineering*, 2016, 143: 28–41.
- [11] COINTAULT F, GUERIN D, GUILLEMIN J, et al. In-field triticum aestivum ear counting using colour-texture image analysis[J]. *N Z J Crop Hortic Science*, 2008, 21(36): 117–130.
- [12] JOSE A F, MA L B, NIEVES A G, et al. Automatic wheat ear counting using thermal imagery[J]. *Remote Sensing*, 2019, 11(7):751.
- [13] 孙红,李松,李民赞,等. 农业信息成像感知与深度学习应用研究进展[J/OL]. *农业机械学报*,2020,51(5):1–17. SUN Hong, LI Song, LI Minzan, et al. Research progress of image sensing and deep learning in agriculture [J/OL]. *Transactions of the Chinese Society for Agricultural Machinery*, 2020, 51(5): 1–17. http://www.j-csam.org/jc-sam/ch/reader/view_abstract.aspx?file_no=20200501&flag=1. DOI:10.6041/j.issn.1000-1298.2020.05.001. (in Chinese)
- [14] 段凌凤,熊雄,刘谦,等. 基于深度全卷积神经网络的大田稻穗分割[J]. *农业工程学报*,2018,34(12):202–209. DUAN Lingfeng, XIONG Xiong, LIU Qian, et al. Field rice panicles segmentation based on deep full convolutional neural network[J]. *Transactions of the CSAE*, 2018, 34(12): 202–209. (in Chinese)
- [15] XIONG Xiong, DUAN Lingfeng, LIU Lingbo, et al. Panicle-SEG: a robust image segmentation method for rice panicles in the field based on deep learning and superpixel optimization[J]. *Plant Methods*,2017,13: 104.
- [16] ZHOU Chengquan, YE Hongbao, HU Jun, et al. Automated counting of rice panicle by applying deep learning model to images from unmanned aerial vehicle platform[J]. *Sensors*, 2019, 19: 3106.
- [17] 张领先,陈运强,李云霞,等. 基于卷积神经网络的冬小麦麦穗检测计数系统[J/OL]. *农业机械学报*,2019,50(3): 144–150. ZHANG Lingxian, CHEN Yunqiang, LI Yunxia, et al. Detection and counting system for winter wheat ears based on convolutional neural network[J/OL]. *Transactions of the Chinese Society for Agricultural Machinery*, 2019,50(3):144–150. http://www.j-csam.org/jcsam/ch/reader/view_abstract.aspx?flag=1&file_no=20190315&journal_id=jcsam. DOI: 10.6041/j.issn.1000-1298.2019.03.015. (in Chinese)
- [18] SIMON M, JIN Xiuliang, LU Hao, et al. Ear density estimation from high resolution RGB imagery using deeplearning technique[J]. *Agricultural and Forest Meteorology*,2019,264:225–234.
- [19] POURIA S T, NICOLAS V, EVA M, et al. Deepcount: in-field automatic quantification of wheat spikes using simple linear iterative clustering and deep convolutional neural networks[J]. *Frontiers in Plant Science*, 2019:1176. DOI:10.3389/fpls.2019.01176.
- [20] CAI Z, NUNO V. Cascade R – CNN: delving into high quality object detection[C]//*IEEE Conference on Computer Vision and Pattern Recognition*. Salt Lake City: IEEE, 2018: 6154–6162.
- [21] JOHN P. Sequential minimal optimization: a fast algorithm for training support vector machines[R]. Microsoft Research, Technical Report MSR – TR – 98 – 14,1998.
- [22] HE K, REN S, GIRSHICK R, et al. Faster R – CNN: towards real-time object detection with regionproposal networks[C]//*In CVPR*,2016:1506–1611.
- [23] GIRSHICK R, DONAHUE J, DARRELL T, et al. Rich feature hierarchies for accurate object detection and semanticsegmentation[C]//*IEEE Conference on Computer Vision and Pattern Recognition*,2014: 580–587.
- [24] GIRSHICK R. Fast R – CNN[C]//*IEEE International Conference on Computer Vision*,2015: 1440–1448.
- [25] HE K, ZHANG X, REN S, et al. Deep residual learning for image recognition[C]//*2016 IEEE Conference on Computer Vision and Pattern Recognition*. Las Vegas, NV, USA: IEEE, 2016:770–780.
- [26] LIN T, DOLLAR P, GIRSHICK R, et al. Feature pyramid networks for object detection[C]//*IEEE Conference on Computer Vision and Pattern Recognition*. New York: IEEE, 2017: 2117–2125.
- [27] ABHINAV S, ABHINAV G, ROSS G. Training region-based object detectors with online hard example mining[C]//*2016 IEEE Conference on Computer Vision and Pattern Recognition*. Las Vegas, NV, USA: IEEE, 2016:761–769.
- [28] SERGEY L, CHRISTIAN S. Batch normalization: accelerating deep network training by reducing internal covariate shift[C]//*Proceedings of the 32nd International Conference on Machine Learning(ICML)*. Lille, France:IEEE, 2015:448–456.
- [29] ALEXANDER N, LUC V G. Efficient non-maximum suppression[C]//*In ICPR*, 2006: 850–855.