

基于改进 Faster R – CNN 的水稻稻穗检测方法

张远琴 肖德琴 陈焕坤 刘又夫
(华南农业大学数学与信息学院, 广州 510642)

摘要: 为了快速而准确地统计视频监控区域内的水稻穗数,提出了一种基于改进 Faster R – CNN 的稻穗检测方法。针对稻穗目标较小的问题,在 Inception_ResNet – v2 的基础上引入空洞卷积进行优化;对于不同生长期稻穗差别大的问题,设计了针对标注框尺度的 K-means 聚类,为候选区域生成网络提供先验知识,从而提高了检测精度。鉴于小尺寸稻穗目标的特殊性,用 ROIAAlign 替代 ROI Pooling,提高了感兴趣区域的提取精度。试验测试时,根据水稻不同发育期稻穗的表型特征差异自制了 3 类数据集,并选取最佳聚类数为 10。模型对比试验表明,本文方法的稻穗检测平均精度均值达到 80.3%,较 Faster R – CNN 模型提升了 2.4 个百分点,且比 SSD 和 YOLO 系列模型有较大幅度的提升。

关键词: 稻穗检测;改进 Faster R – CNN; Inception_ResNet – v2; K-means

中图分类号: TP391.4; S511.3 **文献标识码:** A **文章编号:** 1000-1298(2021)08-0231-10 **OSID:** 

Rice Panicle Detection Method Based on Improved Faster R – CNN

ZHANG Yuanqin XIAO Deqin CHEN Huankun LIU Youfu
(College of Mathematics and Informatics, South China Agricultural University, Guangzhou 510642, China)

Abstract: Rice panicle detection is the core research basis of automatic rice panicle counting and rice yield estimation. Due to the density and small size of rice panicle, the size of rice panicle varies greatly at different growth stages, which brings great challenges to the effective and accurate detection of rice panicle. In order to quickly and accurately count the number of rice panicle in the video monitoring area, a rice panicle detection method based on improved Faster R – CNN was proposed. In order to deal with the problem of small target of rice panicle, dilated convolution was introduced on the basis of Inception_ResNet – v2 to optimize the solution. For the problem that rice panicle size varied greatly in different growing periods, K-means clustering aiming at the scale of label box was designed, so as to provide prior knowledge for region proposal network and improve the detection accuracy. In addition, in view of the particularity of the detection target, ROIAAlign was used instead of ROI Pooling to improve the extraction accuracy of ROI. Using the Faster R – CNN as the basic network and combining the above optimization strategy, a method was proposed for rice panicle detection based on the improved Faster R – CNN. During the experimental test, three data sets were made based on the differences in phenotypic characteristics of rice panicle at different developmental stages, and selected 10 as the best cluster number feasible in practice according to the experimental results. A large number of results showed that the rice panicle detection mAP of this algorithm reached 80.3%, which was 2.4 percentage points higher than that of the original Faster R – CNN model without improved strategy. And compared with SSD and YOLO series model, it had a greater improvement.

Key words: rice panicle detection; improved Faster R – CNN; Inception_ResNet – v2; K-means

0 引言

我国坚持水稻质量与产量并重的原则,将培育

高产优质的稻米作为水稻生产研究的主要方向^[1]。稻穗数是水稻质量与产量评估的重要指标,在水稻病害检测^[2]、营养诊断^[3]及生育期检测^[4]等方面具

收稿日期: 2020 – 09 – 02 修回日期: 2020 – 10 – 31
基金项目: 广东省重点领域研发计划项目(2019B020214002)和广东省农业厅乡村振兴专项基金项目(5600 – F19257)
作者简介: 张远琴(1992—),女,博士生,主要从事农业图像处理研究,E-mail: 969796148@qq.com
通信作者: 肖德琴(1970—),女,教授,博士生导师,主要从事物联网和农业图像处理研究,E-mail: deqinx@scau.edu.cn

有非常重要的作用。长期以来,在水稻栽培或育种研究中,稻穗数都是通过人工统计的方式来获得,该过程耗时、耗力、效率低,在大面积生产场景下,由于植株间的相似性和观察人员的主观性使稻穗计数非常困难,且容易出错。因此,无论是快速预估产量还是精准检测表参数,都迫切需要一种快速、精准的稻穗自动计数方法。

图像是对作物进行观测最直接的手段之一,基于数字图像植株穗的自动计数研究已取得一定的成果。主要分为两类:一类是基于颜色^[5-7]和纹理^[8-9]的分割方法,另一类是基于像素水平颜色特征候选区域^[10-12]和基于超像素融合生成候选区域^[13-14]的分类方法。穗分割方法处理速度较快,但需考虑光照阴影、穗粘连和叶片颜色不均等因素的影响,尤其当目标量与非目标量增加时,提取穗表型特征变得相当困难,计数精度明显下降;基于候选区域的分类方法的关键环节是候选区域的生成,候选区域的质量直接影响分类模型的训练数据质量、分类能力和穗计数精度。

近年来兴起的卷积神经网络(Convolutional neural networks, CNN)是深度学习的重要产物,应用于农业领域图像识别已取得了显著成效^[15-18]。在稻穗识别与分割方面,DUAN等^[12]从多角度分析水稻植株图像,提出一种自动确定盆栽水稻植株穗数的方法,但这种方法需要一个专门的光源和图像背景才能确保其良好的性能。XIONG等^[19]提出一种基于超像素分割和CNN的稻穗分割算法,该算法能很好地应用于不同品种及生育期的稻穗分割,但耗时较长。BAI等^[20]研究了一种水稻抽穗期自动观测方法,该方法不仅利用了稻穗的颜色和梯度特征,还利用了CNN自动学习特征,其检测精度达到75%,但受光照的影响较大,适于处理最佳自然光下拍摄的图像,而处理日常水稻图像时会耗费时间,检测精度也会下降。XU等^[21]提出一种基于深度学习和多尺度混合窗口的稳健稻穗计数算法,该算法通过增强稻穗特征来检测和计算大面积场景中的小尺度稻穗,平均计数精度达87.2%。但该算法只适于成熟期的稻穗。

稻穗自动计数的本质是稻穗的识别与定位,这与目标检测研究的目标识别与定位相一致,由此可将稻穗计数问题转化为稻穗检测问题。图像目标检测的检测方法可分为两类:一是基于区域生成的检测方法,即先由算法生成一系列候选框,再通过卷积神经网络对候选框中的目标进行分类,如Faster R-CNN^[22]等网络模型属于此类检测方法,这类方法虽不能做到实时检测,但识别率高。二是基于回归的

方法,该方法在实现目标定位的同时预测目标分类,属于一体化卷积网络的检测方法,YOLO^[23]系列、SSD^[24]等都属于此类,该类方法具有识别速度快的优势,但准确率往往低于前者。因此,目前已有算法都存在局限性。

本文针对不同生长期和高密度的小尺寸稻穗,提出一种基于改进Faster R-CNN的水稻稻穗检测方法,其目的是提高小目标稻穗检测的精度,为水稻关键发育期预测与产量预估提供研究基础,从而进一步服务于水稻育种与栽培。

1 材料与方 法

1.1 数据样本采集与预处理

本文以种植在华南农业大学资源与环境学院楼顶大棚的盆栽金农丝苗水稻为试验对象。水稻种植在直径约40 cm的花盆中,每盆分3穴栽种幼苗,幼苗间隔10~12 cm。试验场地种植的水稻按不同氮元素水平(N0、N90、N180、N270、N360)划分为5个观察区域,每个区域有9盆水稻。本试验中监测的区域为N180,属于正常范围内的氮元素水平种植的水稻,9盆水稻围成一个监测区域模拟大田环境中3行3列的区域块,在种植间隔和疏密上符合大田种植要求,以尽可能使水稻培育环境与真实大田环境相接近。本试验采用视频摄像头对盆栽水稻进行监测,视频摄像头采用海康DS-2DC4420IW-D E系列智能球机,固定在离地面2.5 m的横梁上。具体试验设备和现场图如图1所示。



图 1 拍摄现场和效果图(抽穗和扬花期)
Fig. 1 Shooting scene and effect picture
(heading and flowering period)

试验时,针对2019年的早稻(5月25日—7月10日)和晚稻(10月26日—12月6日)从抽穗期到成熟期进行定时拍摄。定时抓拍时间设置为每天08:00—18:00,每隔20 min拍摄1次,并自动保存到网络硬盘中。所拍摄图像的尺寸为1 920像素×1 080像素,图像存储格式为JPG。试验期间,共获得了2 580幅图像。其中,抽穗和扬花期图像(I类)600幅、乳熟期图像(II类)780幅、蜡熟和完熟期图像(III类)1 200幅。图2为3类不同发育期的稻穗形态。

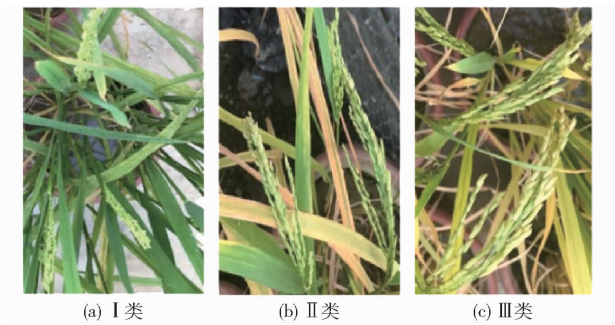


图 2 不同发育期的稻穗形态图

Fig.2 Patterns of rice panicle at different development periods

1.2 数据集构建

在真实的工程应用中,实际获得的数据往往比满足深度学习模型训练要求的数据量少得多。为了解决此问题,一般需要对训练集的样本进行数据增强来扩充数据集^[25]。

本文在模型训练前,将各类数据集分别按照 4: 1 的比例随机划分为训练集和测试集。对于训练集,采用 LabelImg 软件进行人工标注。图 3 为训练

集样本的预处理过程,通过对图像进行水平翻转、平移、旋转和裁剪来扩充数据集。表 1 为各类稻穗图像数量。

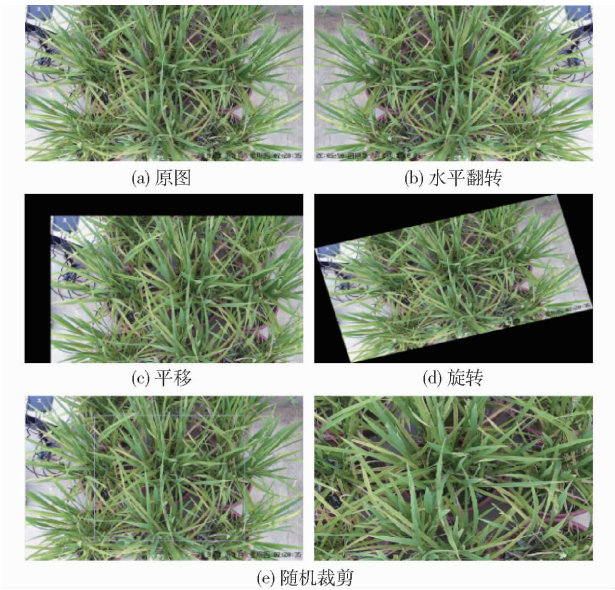


图 3 图像预处理

Fig.3 Image preprocessing

表 1 各类稻穗图像数量
Tab.1 Number of rice panicle images

数据集	I 类			II 类			III 类		
	扩充前	扩充后		扩充前	扩充后		扩充前	扩充后	
		训练集	测试集		训练集	测试集		训练集	测试集
图像数量/幅	600	2 400	120	780	3 120	156	1 200	4 800	240

1.3 水稻稻穗检测方法

Faster R - CNN 属于 Two - stage 目标检测网络的代表,其特点是识别错误率、漏识率低。与 YOLO、SSD 等 One - stage 目标检测网络不同,这类网络模型的突出特点是快速性,但是精度相对 Two - stage 目标检测网络略低。本文要求稻穗的检测精度较高,因此选取 Faster R - CNN 网络模型实现稻穗的识别与定位是合理的。

在本文的稻穗数据集中,统计标注框(刚好包含目标)的宽/高即稻穗长边范围在 26 ~ 250 像素之间,中位数在 150 像素左右,此处反映了所需识别的稻穗多为小目标,即目标尺寸未达图像尺寸的 1/10^[22]。随着水稻的生长发育,从抽穗期到成熟期,不仅各时期的稻穗尺寸会有差异,图像中稻穗的数量也越来越多。

本文考虑到高精度的 Faster R - CNN 网络在小目标物体检测上存在的问题,并结合本文稻穗检测存在目标小且密集、前后不同生长期目标大小跨度大等问题,对 Faster R - CNN 网络进行有针对性的改进。

1.3.1 优化的 Inception_ResNet - v2 特征提取网络

继 2015 年 ResNet 被提出后,2016 年由 Google 团队的 SZEGEDY 等^[26]结合 Inception 和 ResNet 的优点,提出基于 Inception v4 融合残差单元的 Inception_ResNet - v2 网络,其在 ILSVRC 图像分类基准测试中取得了当下最好的成绩。本文为解决稻穗目标较小的问题,在 Inception_ResNet - v2 的结构中引入空洞卷积来代替下采样进行优化解决。优化的 Inception_ResNet - v2 网络结构由 3 部分组成:

(1) Inception v4: Inception v4 是 Google Inception 系列网络的一种,该网络借鉴 Inception v3 的思想在网络中层(12 ~ 20 层)使用卷积核 1 × 3、3 × 1、1 × 7、7 × 1 代替卷积核 3 × 3、7 × 7 进行卷积分解。相比 Inception v3 主要不同是添加了 3 个残差单元,这样既加速了训练速度,又能够训练更深的网络,得到更丰富的空间特征信息。Inception_ResNet - v2 与 Inception v4 结构基本相同,具有相同的 Stem 模块,不同的细节在于残差单元。Stem 模块结构如图 4 所示。

(2) ResNet:单纯地增加网络深度会导致梯度消

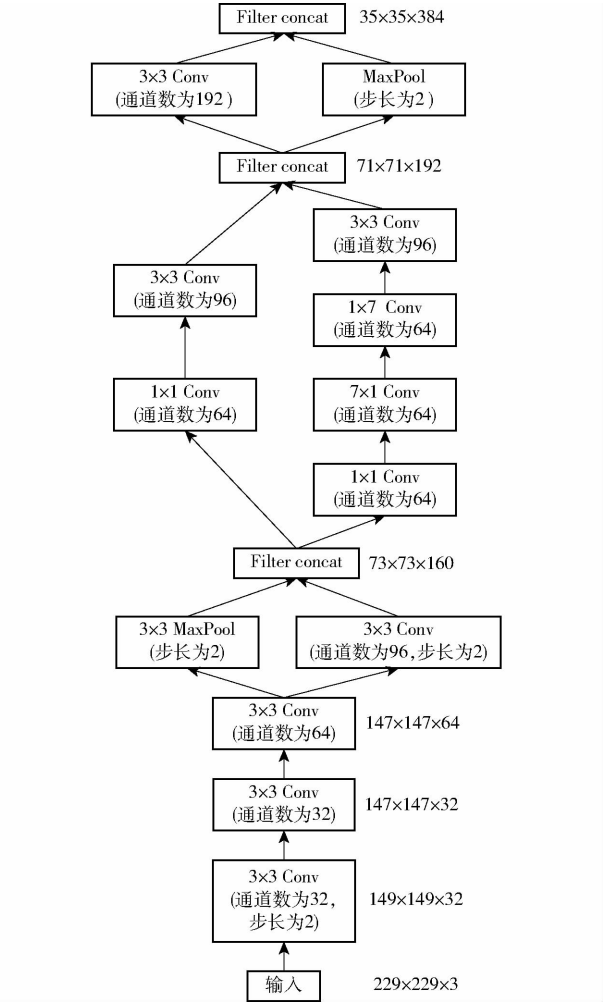


图 4 Stem 模块结构

Fig. 4 Stem module structure

失或梯度爆炸,甚至会使准确率下降。残差网络被提出,其核心是解决因增加网络深度带来的退化问题,使得通过简单地增加网络深度来提高网络性能是可行的。残差结构如图 5 所示,输入张量为 x ,学习残差函数 $F(x) = H(x) - x$,当模型精度达到饱和时,多余的网路层训练目标变为将残差结果逼近于 0,即令 $F(x) = 0$,实现恒等映射,使得随着网络加深,训练精度不会下降。

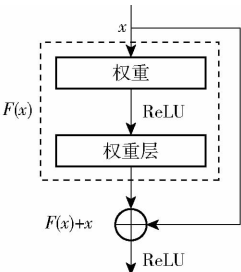


图 5 残差结构

Fig. 5 Residual structure

(3)空洞卷积:深度卷积神经网络提取图像特征往往通过下采样的方式缩小图像尺度同时扩大感受野,输出特征图具有更加丰富的高层语义信息,然

而过多的下采样操作会不可逆地丢失高层边缘信息和小目标的特征信息。空洞卷积的优点是能够在不做池化操作造成信息损失的情况下,保证特征图分辨率不改变同时扩大感受野,使卷积输出包含较丰富的特征信息。不同扩张率的 3×3 空洞卷积如图 6 所示。

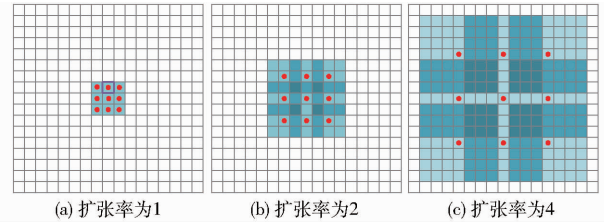


图 6 不同扩张率的 3×3 空洞卷积

Fig. 6 Dilated convolution with 3×3 kernel and different dilation rates

图 7 为引入空洞卷积优化的 Inception_ResNet-v2 特征提取网络,将其作为 Faster R-CNN 网络模型的基础网络。首先,将 Reduction-A、B 模块 (Reduction-A、B 模块见图 8) 中步长为 2 的 3×3 卷积层和最大池化层设置为步长为 1;然后,在 Inception_ResNet-B 模块引入扩张率为 2 的空洞卷积,具体的 Inception_ResNet-A、B、C 模块结构如图 9 所示。最后,将 Inception_ResNet-B 模块的输出作为 Faster R-CNN 中 RPN 网络和区域特征聚集方法 (ROIAlign) 的共享特征图,而 Inception_ResNet-B 之后的所有层作为 ROIAlign 层的后接网络。于是共享卷积层仅有 8 倍下采样,表明提取候选框的基础尺寸可以设置为 8 (原本为 16),即特征图上的每一像素点映射到原始图像上至少有 8 个像素的对应区域,事实上低于候选框下限的小目标往往会被漏检。显然,引入空洞卷积优化解决目标稻穗较小的问题是合适的。

1.3.2 基于 K-means 聚类的 RPN 网络

原始候选区域生成网络 (Region proposal network, RPN) 中的先验框的尺寸 (32, 64, 128) 和比例 (1:2, 1:1, 2:1) 都是人为设定的,在一定程度上带有主观性,不利于学习特定对象的位置和大小。如在本文的稻穗数据集中,稻穗尺寸变化范围较大,这种预先定义的边框提取策略在定位稻穗上将产生较大的偏差,甚至出现漏检的情况。而且,由于先验框的固定生成方式,RPN 网络生成的候选框重叠率过多,故训练时冗余的边框将耗费许多不必要的计算资源,增加模型的收敛时间。因此针对先验框设计问题,本文考虑采用 K-means 的聚类方法,参考稻穗标注框的尺寸,将训练集的尺寸先验性带入网络超参数中。这样,在自己的数据集上聚类计算先验

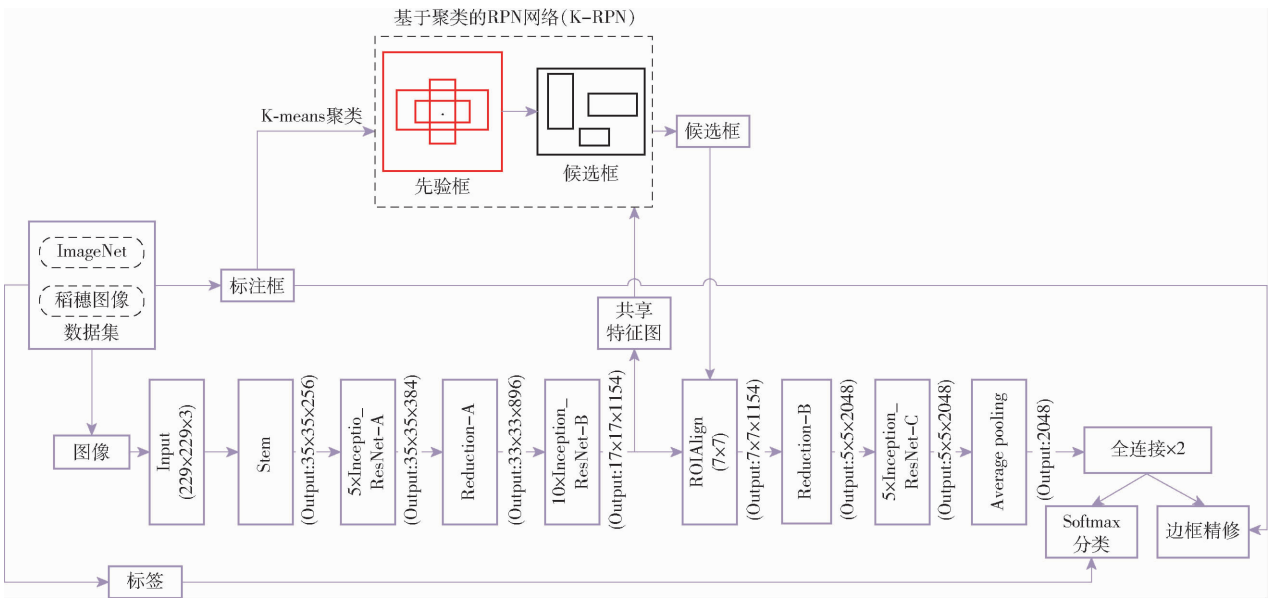


图 7 改进 Faster R - CNN 网络模型结构

Fig. 7 Improved Faster R - CNN network model structure

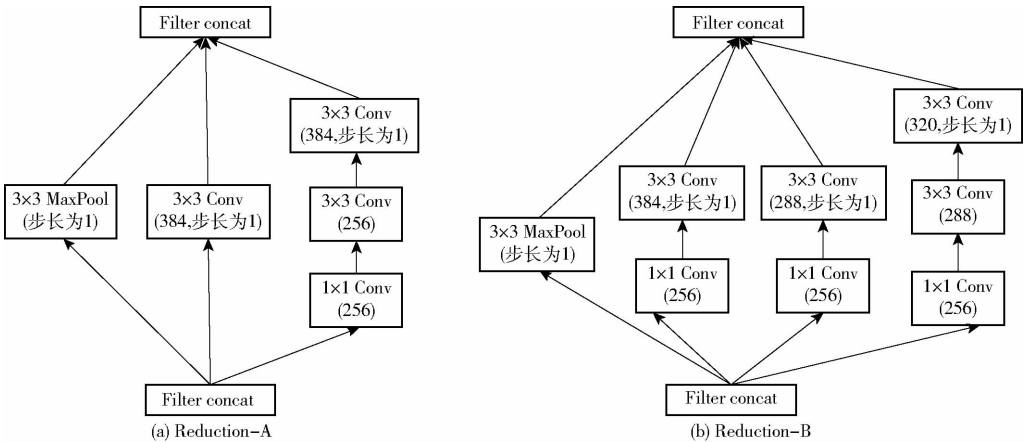


图 8 Reduction - A、Reduction - B 模块结构

Fig. 8 Reduction - A, Reduction - B module structure diagrams

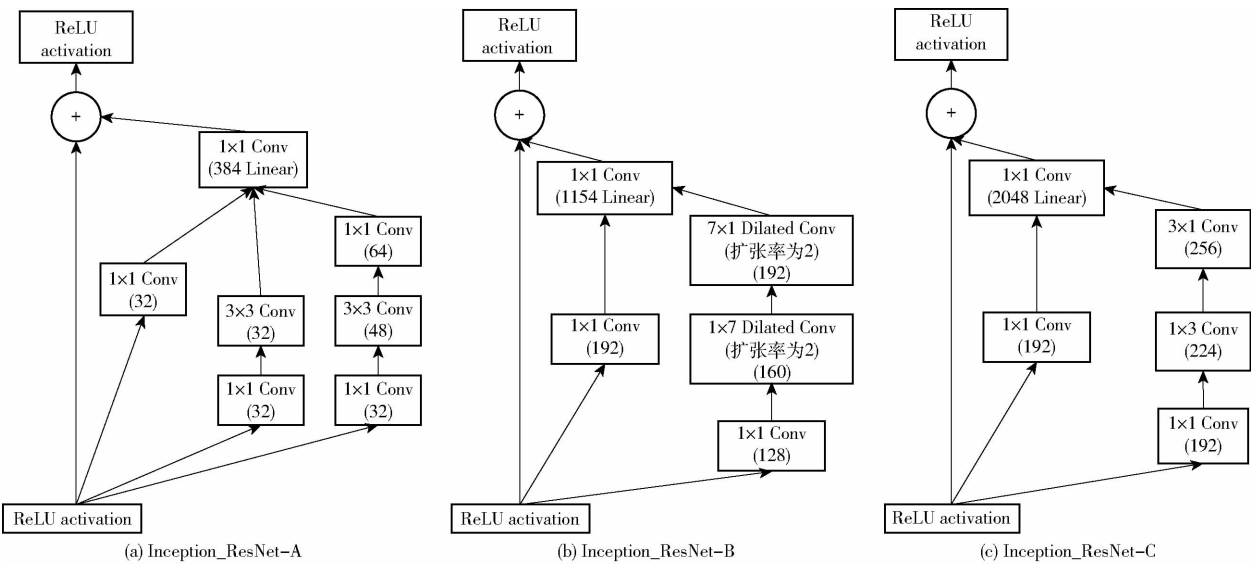


图 9 Inception_ResNet - A、B、C 模块结构

Fig. 9 Inception_ResNet - A, B and C modular structure

框将会更合适,最终能够提高预测框的检出率。

在开始训练前,对数据集样本的标注框尺寸进行 K-means 聚类。在聚类时,提取标注框的宽、高作为 K-means 算法的输入,预设参数 K 为先验框(聚类中心框)的数量。K-means 算法以相似度为依据将训练集中相似度较高的标注框归为同一个类簇,将差异性较高的标注框归为不同的类簇。最后通过计算 K 个不同类簇中所有标注框宽、高的平均值,得到先验框的合适尺寸。聚类结束后,将结果作为 RPN 网络的先验框预设规则。

在聚类的过程中,相似度的计算方法选用标注框与聚类中心框之间的交并比(IoU)作为距离指标,其距离计算公式为

$$d(\text{box}, \text{centre}) = 1 - I_{\text{ou}}(\text{box}, \text{centre})$$

式中 box——标注框

centre——聚类中心框

$d(\text{box}, \text{centre})$ ——标注框与聚类中心框之间的距离

$I_{\text{ou}}(\text{box}, \text{centre})$ ——标注框与聚类中心框的交并比

特别地,计算 IoU 时每个标注框的中心点都与聚类中心点重合。

1.3.3 选择 ROIAlign 替代感兴趣区域池化层

由于本文稻穗尺寸小,且随着生长发育图像中的稻穗也会变得密集,针对小且密的目标对象,感兴趣区域的提取精度变得尤为关键。Faster R - CNN 中的感兴趣区域池化层(ROI Pooling)可以将 RPN 生成的候选框映射到共享卷积层输出的特征图上,得到候选区域在共享特征图上的感兴趣区域,并产生固定大小的感兴趣区域,这里针对小目标稻穗尺寸的特殊性,设置 ROI Pooling 的采样参数为 7×7 。

假设原图中有一候选框尺寸为 249×249 ,映射到共享特征图上的尺寸为 26.75×26.75 ,在程序计算的时候会进行取整操作,即第 1 次量化,得到的特征图尺寸为 26×26 。根据 ROI Pooling 的采样参数为 7×7 ,即 ROI Pooling 后固定生成尺寸为 7×7 的特征图,所以,将上面在特征图上映射的 26×26 的候选区域划分成 7×7 个同等面积的小区域,每个小区域的边长为 $26/7 = 3.71$,即 3.71×3.71 ,此时,进行第 2 次量化,故小区域尺寸变成 3×3 ,那么在每个小区域里取出最大的像素作为区域代表,组成 7×7 的特征图。可以看出,经过两次量化操作,即将浮点数取整,把原本映射到共享特征图上的尺寸为 26.75×26.75 的候选框缩小成尺寸为 21×21 的候选框,这样的像素偏差定会影响到后接网络的回归定位,尤其是小目标物体。显然,这样会对小尺寸

的稻穗检测造成不利的影响。而 ROIAlign 使用双线性内插法来取代 ROI Pooling 的取整操作,保留了浮点数,于是本文选择 ROIAlign 来替代 ROI Pooling,能够获得更精准的效果。

具体地,设置 ROIAlign 的采样点为 4,即把每个 3.82×3.82 小区域平分为 4 个小单元,采用双线性插值算法计算每个小单元的中心位置像素值,得到 4 个小单元中心点的像素值。然后再取这 4 个像素值的最大值作为这个小区域的像素值,同样地, 7×7 个这样的小区域得到 7×7 个像素值,组成 7×7 的特征图。图 10 所示为 ROIAlign 采样获取小区域像素值的操作过程。

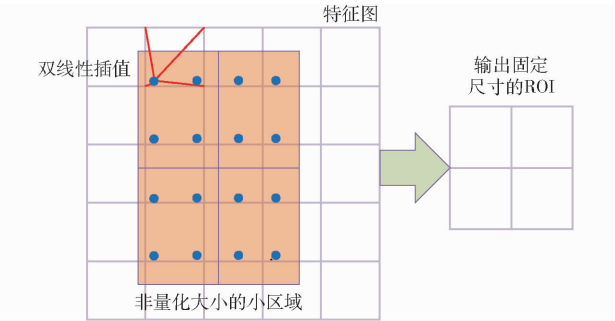


图 10 ROIAlign 采样获取小区域像素值的操作过程
Fig. 10 Operation process of ROIAlign sampling to obtain pixel value of small area

2 试验与结果分析

本文自制了 3 类数据集对提出的模型进行测试。首先,对不同聚类 K 值(选取 6 ~ 12)下优化后模型的检测精度进行试验,确定最佳聚类参数;然后分析不同优化操作下的网络检测性能以及加入聚类对优化后模型检测精度的影响,最后,为验证本文提出模型在稻穗检测上的优势,与 SSD 和 YOLO 等 One - stage 网络模型的检测精度进行对比分析。其中,目标检测精度以平均精度均值(Mean average precision, mAP)为评估标准,平均精度均值是各类别平均精度(Average precision, AP)的平均值。

2.1 试验方法与参数设置

试验运行环境为 Intel i9 9900K CPU、主频 3.40 GHz、内存 64 GB、GPU 为 NVIDIA 2080ti,操作系统为 Ubuntu 16.04 LTS, CUDA 版本为 10.0,以 TensorFlow 14.0 为深度学习框架,编程语言为 Python 3.7。

本文目标检测网络在训练前,先在 ImageNet 数据集上进行预训练,以初始化其参数,然后再在本文数据集上进行迁移学习和微调。训练阶段的网络超参数设计如表 2 所示。网络训练具体步骤为:

表 2 模型训练超参数

Tab. 2 Model training hyperparameters

参数	数值
初始化学学习率	0. 003
权重衰减速率	0. 000 5
学习率衰减速率	0. 1
学习率衰减步长	30 000
冲量	0. 9
网络中非极大值抑制(NMS)的阈值	0. 7
RPN 网络训练的正负样本阈值	0. 7
图像批量大小	256
先验框数量 K	6 ~ 12

(1) 采用 ImageNet 预训练模型进行参数初始化,独立训练 RPN 网络。

(2) 仍用 ImageNet 预训练模型进行参数初始化,但是使用步骤(1) RPN 网络产生的候选区域作为输入,训练 Fast R - CNN 网络。

(3) 使用步骤(2)的 Fast R - CNN 网络参数重新初始化 RPN 网络,但是把 RPN 和 Fast R - CNN 共享的卷积层的学习率设置为 0,意味着不更新,仅

仅更新 RPN 特有的网络层,重新训练,此时 2 个网络已经共享了所有公共的卷积层。

(4) 固定 RPN 和 Fast R - CNN 共享的网络层,加入 Fast R - CNN 特有的网络层,形成一个统一的网络,继续训练,微调 Fast R - CNN 特有的网络层。

2.2 聚类参数选取

K-means 聚类中的 K 代表了先验框划分的精细程度,理论上,其值选取得越大,对标注框聚类得到先验框的尺寸将更加具有代表性,能够更精细地表征稻穗尺寸,RPN 网络预测的候选区域就越精确,最终的目标检测精度也会越高,但是聚类过程涉及的计算资源就越多,同时模型复杂度高,网络训练所耗费的时间也不断增加。

在原始的 RPN 网络中,共享特征图上的每个点生成 9 个先验框作为初始候选区域。据此本试验的 K 以 9 为中心,取值范围为 6 ~ 12,分析不同 K 值对网络训练时间以及检测精度的影响,试验结果如图 11 所示。

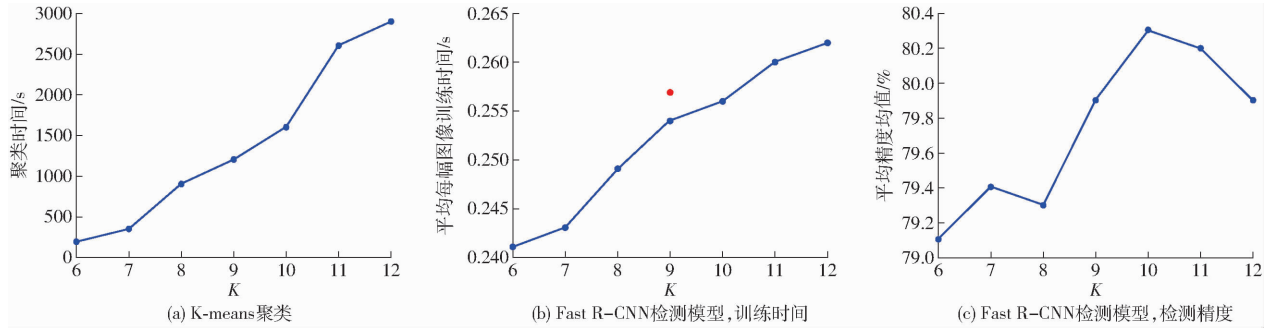


图 11 聚类检测结果
Fig. 11 Clustering detection results

由图 11a 可知,当 K 增大时,聚类时间以接近线性增长的速度越来越大。这是因为在聚类过程中,聚类数 K 表示把训练集的所有标注框分类,从 $K - 1$ 类到 K 类就需要增加计算所有标注框到第 K 个聚类中心框的距离。

由图 11b 可知,随着 K 越来越大,平均每幅图像训练时间逐步增加。原因是当先验框的数量逐渐增大时,共享特征图上每个候选框数量越来越多,候选区域产生的样本数量增大,增加了网络回归层的计算量,导致训练时间更长。图中红色点表示具有原始 RPN 的 Faster R - CNN 检测模型的训练用时,该模型较 $K = 9$ 时的模型用时更长,这说明在设定相同数量的先验框时,通过对标注框聚类统计取代人工设计先验框尺寸的方式,更适合网络自身学习,验证了本文方法的有效性。

由图 11c 可知,随着 K 越来越大,平均精度均值整体呈上升趋势;当 K 为 11 或 12 时,平均精度均值

已趋于稳定或有下降趋势。原因为:①当 K 较小时,先验框集合的数量比较少,不能很好地覆盖图像上目标可能出现的区域,造成模型对部分稻穗出现漏检的情况。②当 K 增大至 10 的过程中,尺寸更小的先验框更能够表征稻穗的尺寸,精度随之提升,其中 $K = 9$ 的检测精度比未加入聚类的模型提高了 0.5 个百分点。③随着 K 继续增长(11 或 12),本文认为在共享特征图上生成的先验框数量已趋于饱和,原因在于一方面剔除超过图像边界的先验框更多,另一方面,后续的非极大值抑制过程会筛除大量重叠率高的候选区域。因此,更多的先验框对提高检测精度并不起作用。

因聚类获取先验框的步骤能够脱离整个网络单独进行,且聚类算法运行时间远小于 Faster R - CNN 模型的训练时间,故本文权衡模型训练用时和平均精度均值来选定最优的 K 值。由图 11b、11c 可知,训练时间随 K 增大而增大,当 K 为 11 或 12 时,模型

的检测精度已有下降趋势。考虑模型训练用时和检测精度,最终选定 K 为 10。

2.3 模型检测精度对比

为了验证本文提出的优化策略能够有效地改进基于 Inception_ResNet - v2 特征提取网络的 Faster R - CNN 模型对小目标稻穗检测的不足,本文采用 3 类不同时期稻穗的数据集,通过试验分析了不同优化操作下的网络检测性能以及对加入聚类优化后的模型进行测试,结果分别如表 3、4 所示。

表 3 不同优化操作下的网络检测平均精度

Tab.3 Network detection performance under different optimization operations %

优化方式	AP			mAP
	I 类	II 类	III 类	
Inception_ResNet - v2 + ROIPooling	76.6	77.8	79.2	77.9
Inception_ResNet - v2 + 空洞卷积 + ROIPooling	78.1	78.6	80.1	78.9
Inception_ResNet - v2 + 空洞卷积 + ROIAlign	78.7	79.0	80.4	79.4

表 4 加入聚类优化后的模型检测精度对比

Tab.4 Comparison of model test results after adding clustering optimization

模型	AP/%			mAP/ %	检测速度/ (ms·f ⁻¹)
	I 类	II 类	III 类		
未加入聚类(RPN)	78.7	79.0	80.4	79.4	187
加入聚类(K - RPN,K = 9)	78.9	79.9	80.6	79.8	134
加入聚类(K - RPN,K = 10)	79.7	80.0	81.2	80.3	169

由表 3 可知,基于 Faster R - CNN 框架,仅采用 Inception_ResNet - v2 为特征提取网络的原始模型的 mAP 为 77.9%;引入空洞卷积的优化模型较原始模型的 mAP 提高了 1.0 个百分点,在此基础上,采用 ROIAlign 替代 ROIPooling 的优化操作使得模型的 mAP 提高了 0.5 个百分点;其中分别对 I、II、III 类的稻穗检测精度提高了 2.1、1.2、1.2 个百分点。表明引入空洞卷积和 ROIAlign 的优化策略能够弥补 Faster R - CNN 算法应用于本文数据集在检测上的不足,在保证了对较大目标稻穗检测精度有较好提升的情况下,在一定程度上改善了对小目标稻穗的识别精度。

由表 4 可知,加入聚类后的模型($K = 10$)其稻穗平均精度均值为 80.3%,较未加入聚类模型提升 0.9 个百分点。试验结果表明 RPN 网络引入 K-means 聚类的先验框生成方式能够更好地识别稻穗,验证了本文提出的加入聚类优化模型实现稻穗计数问题的可行性。

如图 12 所示,为了突出模型改进前后对稻穗图像检测的效果,选取关键部分细节图进行展示。图 12a 中从左到右依次为原始模型分别对 I、II、III 类稻穗图像检测结果细节图。可以看出,原始模型的检测结果存在检测框不能完全覆盖稻穗表面、稻穗漏检、叶片误识别的情况。而改进后的模型具有更优的检测效果。

为进一步验证改进模型的有效性,本文基于相同的数据集,训练了 SSD 和 YOLO 系列检测模型,

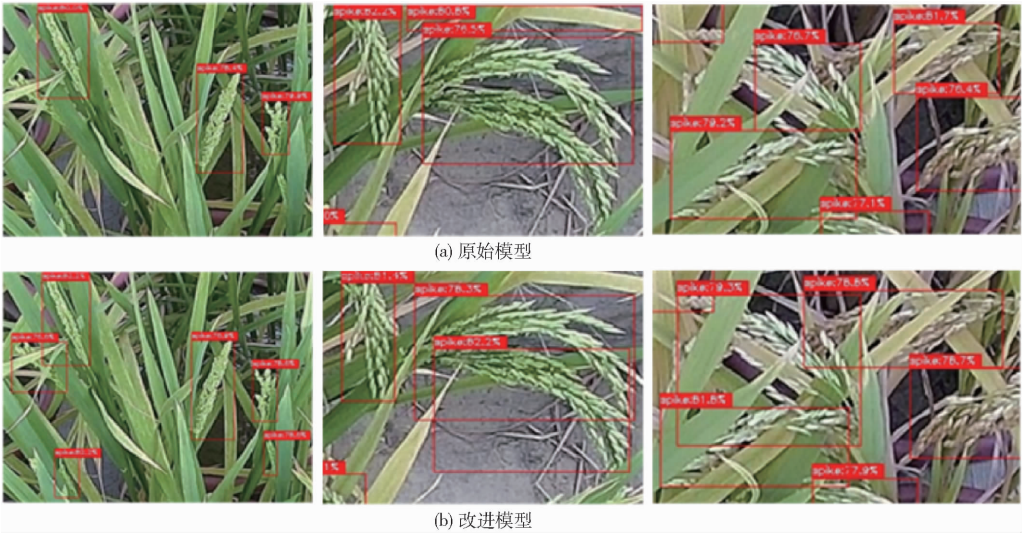


图 12 模型改进前后水稻穗图像关键部分检测效果对比

Fig.12 Comparison of key parts of rice panicle images detection effect before and after model improvement

并对其检测精度进行对比试验,试验结果如图 13 所示。由图 13 可知,One - stage 检测网络对于以稻穗为对象的数量多且尺度较小的目标识别性能较低,其中 SSD500 与 Faster R - CNN 的检测精度较接近。

试验结果验证了本文方法对稻穗检测具有优越性。

3 结束语

提出了一种基于改进 Faster R - CNN 的水稻稻

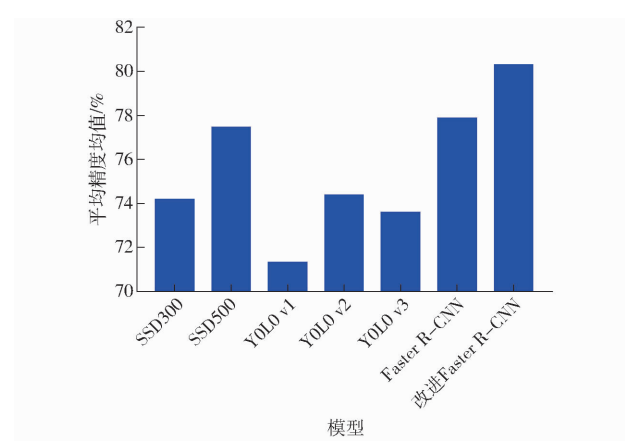


图 13 模型对比结果

Fig. 13 Model comparison results

穗检测方法。首先,引入空洞卷积来优化 Inception_ ResNet - v2 特征提取网络,在此基础上采用 ROIAlign 替代 ROI Pooling 得到优化模型;然后,对标注框进行聚类分析,选取合适的先验框尺寸及数量,以优化 RPN 网络中先验框生成方式;最终获得基于改进 Faster R - CNN 稻穗检测模型。将稻穗数据集分为 I、II、III 类,试验表明,本文方法对 I、II、III 类稻穗的检测精度分别为 79.7%、80.0%、81.2%,平均精度均值为 80.3%,相较于原模型(仅用 Inception_ResNet - v2 作为特征提取网络)的平均精度均值提升了 2.4 个百分点。与 SSD 和 YOLO 系列检测模型相比,改进 Faster R - CNN 检测模型能够更有效地识别小目标稻穗,提高了检测精度。

参 考 文 献

[1] 张明,刘道军,杨再强. 优质高产水稻栽培技术研究[J]. 农家科技旬刊,2014(3):109.
ZHANG Ming, LIU Daojun, YANG Zaiqiang. Research on high yield and high quality rice cultivation technology[J]. Agricultural Science and Technology,2014(3):109. (in Chinese)

[2] ZHOU X. First report of bacterial panicle blight of rice caused by burkholderia glumae in South Africa[J]. Plant Disease,2014, 98(4):566 - 566.

[3] ZHANG Y, LIU M, DANNENMANN M, et al. Benefit of using biodegradable film on rice grain yield and N use efficiency in ground cover rice production system[J]. Field Crops Research,2017,201:52 - 59.

[4] BONIL K, KANG S K, SANG W G, et al. Variation of panicle differentiation stage by leaf growth according to rice cultivars and transplanting time[J]. Korean Journal of Crop Science,2013,58(4):353 - 361.

[5] 陈含,吕行军,田凤珍,等. 基于 Sobel 算子边缘检测的麦穗图像分割[J]. 农机化研究,2013,35(3):33 - 36.
CHEN Han, LÜ Xingjun, TIAN Fengzhen, et al. Wheat panicle image segmentation based on Sobel operator edge detection [J]. Journal of Agricultural Mechanization Research,2013,35(3):33 - 36. (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. Wheat panicle counting in field 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?file_no=20140247&flag=1. DOI:10.6041/j.issn.1000-1298.2014.02.047. (in Chinese)

[7] TANG W, ZHANG Y, ZHANG D, et al. Corn tassel detection based on image processing[C]//International Workshop on Image Processing and Optical Engineering. International Society for Optics and Photonics,2012.

[8] COINTAULT F, GUERIN D, GUILLEMIN J P, et al. In-field triticum aestivum panicle counting using colour-texture image analysis[J]. New Zealand Journal of Crop & Horticultural Science,2008,36(2):117 - 130.

[9] FRESERIC C, FRESERIC J, GILLES R, et al. Texture, color and frequential proxy-detection image processing for crop characterization in a context of precision agriculture[M]. Godwin Aflakpui. Agricultural Science, Alexandria LA:InTech,2012: 49 - 70.

[10] 赵锋,王克俭,苑迎春. 基于颜色特征和 AdaBoost 算法的麦穗识别的研究[J]. 作物杂志,2014(1):141 - 144.
ZHAO Feng, WANG Kejian, YUAN Yingchun. Research on wheat panicle recognition based on color characteristics and AdaBoost algorithm[J]. Crops,2014(1):141 - 144. (in Chinese)

[11] 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.

[12] DUAN L, HUANG C, CHEN G, et al. Determination of rice panicle numbers during heading by multi-angle imaging[J]. The Crop Journal,2015,3(3):211 - 219.

[13] LU H, CAO Z, XIAO Y, et al. Fine-grained maize tassel trait characterization with multi-view representations[J]. Computers & Electronics in Agriculture,2015,118:143 - 158.

[14] 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. Statistical method of panicle of wheat in the field based on superpixel segmentation[J]. Scientia Agricultura Sinica, 2019, 52(1):21 - 33. (in Chinese)

[15] 许景辉,邵明烨,王一琛,等. 基于迁移学习的卷积神经网络玉米病害图像识别[J/OL]. 农业机械学报,2020,51(2):230 - 236,253.
XU Jinghui, SHAO Mingye, WANG Yichen, et al. Recognition of corn leaf spot and rust based on transfer learning with

- convolutional neural network[J/OL]. Transactions of the Chinese Society for Agricultural Machinery, 2020, 51(2):230 – 236,253. http://www.j-csam.org/jcsam/ch/reader/view_abstract.aspx?file_no=2002025&flag=1. DOI:10.6041/j.issn.1000-1298.2020.02.025. (in Chinese)
- [16] 袁培森,申成吉,徐焕良. 基于迁移学习和双线性 CNN 的细粒度菌菇表型识别[J/OL]. 农业机械学报,2021,52(7):151 – 158. YUAN Peisen, SHEN Chengji, XU Huanliang. Fine-grained mushroom phenotype recognition based on transfer learning and bilinear CNN[J/OL]. Transactions of the Chinese Society for Agricultural Machinery, 2021, 52(7):151 – 158. http://www.j-csam.org/jcsam/ch/reader/view_abstract.aspx?file_no=20210715&flag=1. DOI:10.6041/j.issn.1000-1298.2021.07.015. (in Chinese)
- [17] 肖德琴,黄一桂,张远琴,等. 基于改进 Faster R – CNN 的田间黄板害虫检测算法[J/OL]. 农业机械学报,2021,52(6):242 – 251. XIAO Deqin, HUANG Yigui, ZHANG Yuanqin, et al. Pest detection algorithm of yellow plate in field based on improved Faster R – CNN[J/OL]. Transactions of the Chinese Society for Agricultural Machinery, 2021, 52(6):242 – 251. http://www.j-csam.org/jcsam/ch/reader/view_abstract.aspx?file_no=20210625&flag=1. DOI:10.6041/j.issn.1000-1298.2021.06.025. (in Chinese)
- [18] 陈娟,陈良勇,王生生,等. 基于改进残差网络的园林害虫图像识别[J/OL]. 农业机械学报,2019,50(5):187 – 195. CHEN Juan, CHEN Liangyong, WANG Shengsheng, et al. Pest image recognition of garden based on improved residual network [J/OL]. Transactions of the Chinese Society for Agricultural Machinery, 2019, 50(5):187 – 195. http://www.j-csam.org/jcsam/ch/reader/view_abstract.aspx?file_no=20190522&flag=1. DOI:10.6041/j.issn.1000-1298.2019.05.022. (in Chinese)
- [19] XIONG X, DUAN L, LIU L, 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(1):104.
- [20] BAI X, CAO Z, ZHAO L, et al. Rice heading stage automatic observation by multi-classifier cascade based rice panicle detection method[J]. Agricultural & Forest Meteorology, 2018, 259:260 – 270.
- [21] XU C, JIANG H, YUEN P, et al. MHW – PD: a robust rice panicles counting algorithm based on deep learning and multiscale hybrid window[J]. Computers and Electronics in Agriculture, 2020, 173:105375.
- [22] 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, 2017, 39(6):1137 – 1149.
- [23] REDMON J, FARHADI A. YOLOv3: an incremental improvement[J]. arXiv: 1804.02767.
- [24] LIU W, ANGUELOV D, ERHAN D, et al. SSD: single shot multibox detector[C]// European Conference on Computer Vision. Springer International Publishing, 2016, 9905:21 – 37.
- [25] LOSHIN D. The practitioners guide to data quality improvement[M]. Burlington: Morgan Kaufmann, 2011.
- [26] SZEGEDY C, IOFFE S, VANHOUCKE V, et al. Inception – v4, Inception – ResNet and the impact of residual connections on learning[C]// AAAI Conference on Artificial Intelligence, 2016.