

融合动态机制的改进型 Faster R-CNN 识别田间棉花顶芽

陈柯屹¹, 朱龙付¹, 宋 鹏^{1*}, 田晓敏², 黄成龙³, 聂新辉², 肖爱玲⁴, 何良荣⁵

(1. 华中农业大学植物科学技术学院, 武汉 430070; 2. 石河子大学农学院, 石河子 832003; 3. 华中农业大学工学院, 武汉 430070; 4. 塔里木大学机械电气化工程学院, 阿拉尔 843300; 5. 塔里木大学植物科学学院, 阿拉尔 843300)

摘 要: 针对田间密植环境棉花精准打顶时, 棉花顶芽因其小体积特性所带来识别困难问题, 该研究提出一种改进型快速区域卷积神经网络 (Faster Region Convolutional Neural Networks, Faster R-CNN) 目标检测算法实现大田环境棉花顶芽识别。以 Faster R-CNN 为基础框架, 使用 RegNetX-6.4GF 作为主干网络, 以提高图像特征获取性能。将特征金字塔网络 (Feature Pyramid Network, FPN) 和导向锚框定位 (Guided Anchoring, GA) 机制相融合, 实现锚框 (Anchor) 动态自适应生成。通过融合动态区域卷积神经网络 (Dynamic Region Convolutional Neural Networks, Dynamic R-CNN), 实现训练阶段检测模型自适应候选区域 (Proposal) 分布的动态变化。最后在目标候选区域 (Region of Interest, ROI) 中引入目标候选区域提取器 (Generic ROI Extractor, GROIE) 提高图像特征融合能力。采集自然环境下 7 种不同棉花总计 4 819 张图片, 建立微软常见物体图像识别库 2017 (Microsoft Common Objects in Context 2017, MS COCO 2017) 格式的棉花顶芽图片数据集进行试验。结果表明, 该研究提出方法的平均准确率均值 (Mean Average Precision, MAP) 为 98.1%, 模型的处理帧速 (Frames Per Second, FPS) 为 10.3 帧/s。其 MAP 在交并比 (Intersection Over Union, IOU) 为 0.5 时较 Faster R-CNN、RetinaNet、Cascade R-CNN 和 RepPoints 网络分别提高 7.3%、78.9%、10.1% 和 8.3%。该研究算法在田间对于棉花顶芽识别具有较高的鲁棒性和精确度, 为棉花精准打顶作业奠定基础。

关键词: 深度学习; 算法; 棉花; 机制融合; 动态适应; 顶芽识别; Faster R-CNN

doi: 10.11975/j.issn.1002-6819.2021.16.020

中图分类号: TP274; TP391.41

文献标志码: A

文章编号: 1002-6819(2021)-16-0161-08

陈柯屹, 朱龙付, 宋鹏, 等. 融合动态机制的改进型 Faster R-CNN 识别田间棉花顶芽[J]. 农业工程学报, 2021, 37(16): 161-168. doi: 10.11975/j.issn.1002-6819.2021.16.020 http://www.tcsae.org

Chen Keyi, Zhu Longfu, Song Peng, et al. Recognition of cotton terminal bud in field using improved Faster R-CNN by integrating dynamic mechanism[J]. Transactions of the Chinese Society of Agricultural Engineering (Transactions of the CSAE), 2021, 37(16): 161-168. (in Chinese with English abstract) doi: 10.11975/j.issn.1002-6819.2021.16.020 http://www.tcsae.org

0 引 言

棉花是中国重要的经济作物和战略资源, 其种植规模日益扩大^[1]。作为棉花栽培过程的重要环节, 棉花打顶作业正朝着智能化方向发展^[2]。棉花打顶即通过去除其主茎生长点来抑制顶端优势, 使其提早结铃, 提高种植效益^[3]。田间自然环境进行棉花顶芽的精准识别, 是实现棉花顶芽精准去除作业的基础。

早期研究者主要通过激光、超声等传感器检测棉花顶芽所处高度^[4], 如孙杰^[5]采用并排激光对射传感器识别棉花顶芽位置; 李霞等^[6]采用超声波测距原理测量棉花株高来控制打顶作业, 该类方法主要通过检测棉花植株高度来指导打顶, 误差较大。图像处理技术的发展为棉花顶芽精准识别提供了新思路, 刘俊奇^[7]开发了用于棉花打顶高度检测的棉花顶芽图像识别系统, 验证了采用人工神经网络进行棉

花顶芽识别的可行性; 瞿端阳^[8]提出了一种基于机器视觉技术的棉株定位识别方法, 该类方法在室内识别精度高, 但易受田间环境影响, 进而影响棉花顶芽检测的稳定性。

卷积神经网络的提出, 为田间环境目标检测检测提供了新思路^[9-11], 越来越多的学者针对不同农业应用场景研究各种改进算法, 来进一步提高检测精度和效率^[12-14]。彭明霞等^[15]在 Faster R-CNN 识别算法中融合特征金字塔网络来提高算法的图像特征提取能力, 进行复杂棉田杂草识别; 姚青等^[16]采用基于 ResNeXt101 的特征提取网络的 RetinaNet 作为水稻冠层虫害检测模型, 通过改进网络中的特征提取网络提高了识别算法的识别精度。

本研究以中国新疆田间种植待打顶时期棉花为对象, 针对自然环境复杂、棉花田间种植密度大、顶芽区域相对较小等特性, 提出一种融合动态机制的改进型 Faster R-CNN 棉花顶芽识别方法。采用 RegNetX-6.4GF 作为特征提取模块, 融合导向锚框定位 (Guided Anchoring, GA) 机制、动态区域卷积神经网络 (Dynamic Region Convolutional Neural Networks, Dynamic R-CNN) 方法及目标候选区域提取器 (Generic ROI Extractor, GROIE), 最终实现田间棉花顶芽的准确识别, 以期棉花精准打顶作业奠定基础。

收稿日期: 2021-05-31 修订日期: 2021-08-10

基金项目: 中央高校基本科研基金 (2662019QD053)

作者简介: 陈柯屹, 研究方向为农业信息感知及农业机器人技术。

Email: davidchan_n1@outlook.com

*通信作者: 宋鹏, 博士, 副教授, 硕士生导师, 研究方向为农业信息技术及装备。Email: songp@mail.hzau.edu.cn

中国农业工程学会会员: 宋鹏 (E041201229S)

1 材料与方法

1.1 棉花顶芽数据集

1.1.1 图片采集

为建立自然环境下棉花群体中较为完备的棉花顶芽数据集,本研究于 2020 年 7 月 6 日至 2020 年 7 月 17 日期间,在新疆石河子地区采集实际待打顶时期自然环境下的陆地棉 (*Gossypium hirsutum*) 群体图像。其种植模式为一膜四行,行距(28+50+28) cm+55 cm,株距

9.5 cm,行长 5 m。该群体来源广泛,其叶片性状因遗传变异有所差别^[17]。为保证本研究具有代表性,选取该群体中 7 种叶片形状有差异的植株,采用具有不同分辨率的智能手机在田间随机取样,采集时间为北京时间 10:00—18:00。分别在距离棉花冠层 30~50 cm 处(中距离)及 50~100 cm 处(远距离)两种场景下,在均匀光照、强光斜射、强光直射、光照阴影 4 种不同光照条件,采集大田中不同类别棉花图像。采集过程不做任何遮光处理,采集的原始图片如图 1 所示。

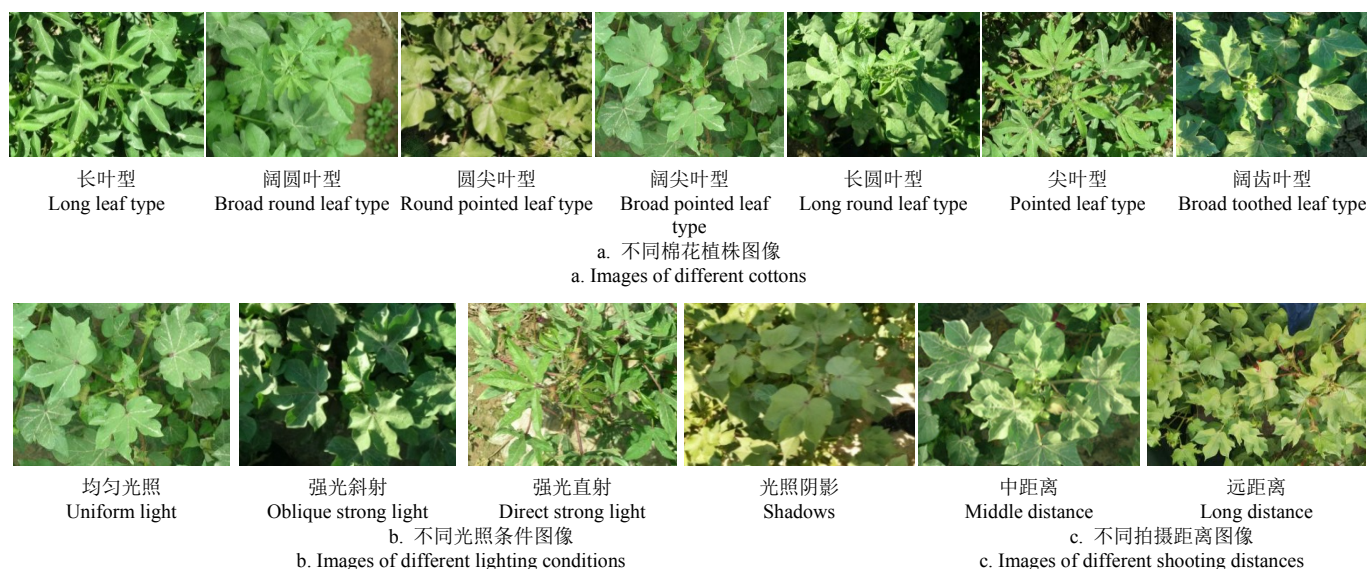


图 1 田间棉花样本图像

Fig.1 Cotton sample image in filed

1.1.2 数据集制作

本文选取包含以上各场景的棉花样本图片总计 4 819 张,其中分辨率为 2 592×1 944 (像素) 像素的图片 3 000 张,分辨率为 8 000×6 000 (像素) 的图片 1 819 张。图片集中对应图 1a 中的不同棉花的图片数量分别为 2 175、406、408、301、448、468 和 613 张。使用开源标注工具 Labelimg 对图片进行标注。选取两种不同距离拍摄场景下的 1 300 张棉花样本图片作为测试集,其中包括中距离场景图片 966 张及远距离场景图片 334 张。将余下的 3 519 张图片按照 4:1 的比例随机划分为训练集和验证集。数据集格式为 MS COCO 2017。

1.2 改进型 Faster R-CNN 田间棉花顶芽识别模型

Faster R-CNN^[18]是两阶段目标检测框架的主流框架,两个阶段可以视为两次以分类为基础的目标检测。在采用 Faster R-CNN 进行棉花顶芽识别时,检测器通过卷积神经网络对输入的棉花顶芽图片进行特征提取并生成特征图 (Feature Maps),之后特征图输入区域生成网络 (Region Proposal Network, RPN),生成多尺寸多比例的锚框 (Anchor) 网格,形成诸多候选目标区域,并对这些目标候选区域进行前景和背景的分类计算及网格边框回归计算,通过比较计算得分,生成目标候选区 (Region of Interest, ROI),最后在检测器的第二阶段,对 ROI 进行进一步分类和边框回归来预测棉花顶芽类别的置信度及其所在位置的边框。

考虑大田密植环境棉花顶芽小区域特性,针对 Faster R-CNN 算法难以有效动态适应复杂环境小目标识别问题,本文对 Faster R-CNN 算法进行优化。采用 RegNetX-6.4GF 作为特征提取模块,改善棉花图片特征提取能力。在 RPN 中使用 FPN 并融合 Guided Anchoring 机制,使 RPN 可以根据棉花顶芽特征图信息动态生成相应的 Anchor,形成以棉花顶芽图片特征为基础的候选区域 Proposals。同时在其训练阶段融合 Dynamic R-CNN 机制,使识别算法在第二阶段根据基于棉花特征信息的 Proposal 动态变化过程来改变 IOU 阈值,获得高质量的 Proposals 来提高模型的训练效果,同时引入 GROIE 机制,使其在 ROI 阶段有效提取 FPN 中的特征信息。改进型 Faster R-CNN 网络结构见图 2。

由于 Guided Anchoring 机制的引入,模型的整体 Loss 函数在 Faster R-CNN 计算训练阶段 RPN 中分类损失 L_{cls} 和回归损失 L_{reg} 的基础上,增加 Anchor 位置和形状损失函数 L_{loc} 和 L_{shape} , 如式 (1) 所示。

$$L = \lambda_1 L_{loc} + \lambda_2 L_{shape} + L_{cls} + L_{reg} \quad (1)$$

式中 λ_1 和 λ_2 分别为 L_{cls} 和 L_{reg} 的权重因子,依试验结果均设定为 1^[18]。Faster R-CNN 的整体 Loss 函数公如下:

$$L(\{p_i\}, \{t_i\}) = \frac{1}{N_{cls}} \sum_i L_{cls}(p_i, p_i^*) + \lambda \frac{1}{N_{reg}} \sum_i p_i^* L_{reg}(t_i, t_i^*) \quad (2)$$

$$L_{cls}(p_i, p_i^*) = -\log_2 [p_i^* p_i + (1 - p_i^*)(1 - p_i)] \quad (3)$$

$$L_{reg}(t, t_i^*) = R(t_i - t_i^*) \quad (4)$$

式中 p_i 为 Anchor 的预测概率, p_i^* 为样本标签, 其值 0 和 1 分别表示负样本和正样本。 $L_{cls}(p_i, p_i^*)$ 为前后背景两个类别的对数损失。Anchor 数量 N_{cls} 在训练 RPN 的阶段, 为 256, 在训练快速区域卷积神经网络 (Fast Region Convolutional Neural Networks, Fast R-CNN) 的阶段, 为 128。 λ 值设为 $1^{[18]}$, N_{reg} 为训练过程中特征图的尺寸大小。 t_i 为表示 Anchor 预测偏移量的向量, t_i^* 为维度相同的向量, 表示 Anchor 相对于真实目标框实际的偏移量, $L_{reg}(t, t_i^*)$ 为边框回归损失, 其采用 Smooth L1 Loss (式 (5)) 计算, 根据经验 σ 值设为 3。

$$\text{Smooth}_{L1}(x) = \begin{cases} 0.5x^2 \times (1/\sigma^2), & |x| < 1/\sigma^2 \\ |x| - 0.5, & \text{其他} \end{cases}, x = t_i - t_i^* \quad (5)$$

L_{loc} 的计算采用 Focal Loss 函数^[19], 如式 (6) 所示。

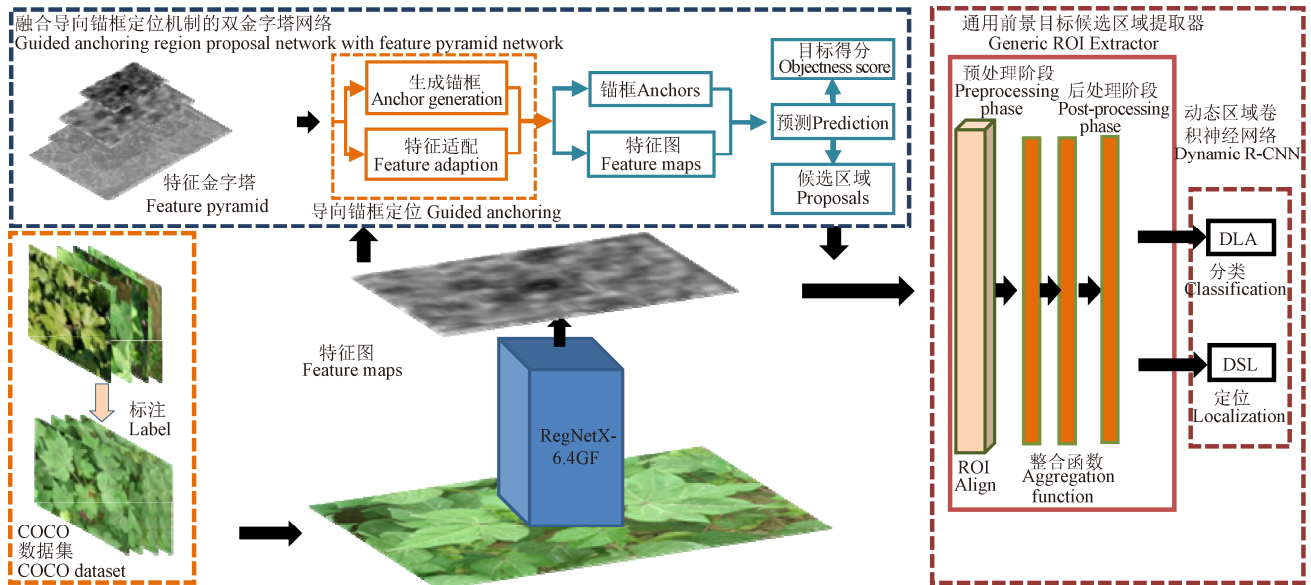
$$FL(p_i) = -\alpha_i (1 - p_i)^\gamma \log_2(p_i) \quad (6)$$

式中 p_i 表示类别概率, 其值越大样本特征越容易被学习识别。 α_i 是调节类别权重因子, 其值为第一类正样本权重, 按照试验经验设置为 0.25, γ 是调节难易样本的权重因子, 同样以试验经验设定为 $2^{[19]}$ 。Focal Loss 的使用, 使 Anchor 生成的位置与图像中实际物体位置更好地契合, 同时也更好的平衡了特征点过少带来的正负样本不平衡的问题。

L_{shape} 的计算如式 (7) 所示。

$$L_{shape} = L_1 \left(1 - \min \left(\frac{w}{w_g}, \frac{w_g}{w} \right) \right) + L_1 \left(1 - \min \left(\frac{h}{h_g}, \frac{h_g}{h} \right) \right) \quad (7)$$

式中 (w, h) 和 (w_g, h_g) 表示 Anchor 的宽长及其对应目标的真实标准宽长。 L_1 代表 Smooth L1 Loss。



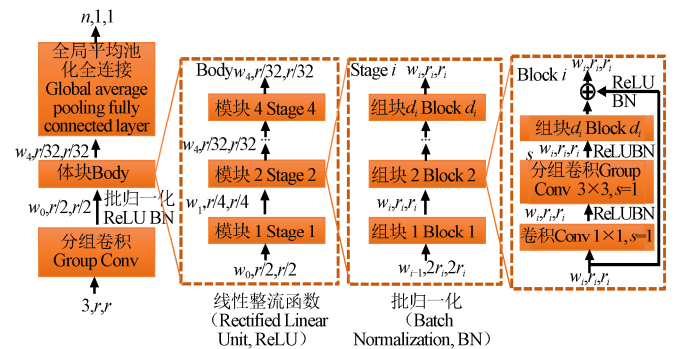
注: ROI 为目标候选区域; ROI Align 为前景目标候选区域对齐; DLA 为动态标注分配; DSL 为动态 smooth L1 损失函数。
Note: ROI is region of interest; ROI Align is region of interest align; DLA is dynamic label assignment; DSL is dynamic smooth L1 loss.

图 2 改进型 Faster R-CNN 网络结构
Fig.2 Improved Faster R-CNN structure

1.2.1 RegNetX-6.4GF 卷积网络

卷积神经网络从图片中获取不同层次特征的能力决定着模型的性能。其中, 层数深浅则影响着网络的特征提取能力^[20-21]。针对棉花顶芽在复杂背景情况下的图像特征, 本文选取 RegNetX-6.4GF^[22] 作为卷积神经网络替换 Faster R-CNN 原本采用的 VGG-16, 其结构如图 3 所示, 该网络结构中输入特征图的长度和宽度均为 r , w 为特征图的通道数, $w_i = [168, 392, 784, 1624]$ 。

RegNet 网络结构使用神经结构搜索 (Neural Architecture Search, NAS) 技术实现。给定设计空间 (Designed Search Space) 通过搜索算法去搜索出一组最佳参数组合。本文采取 RegNetX-6.4GF 网络结构作为骨干网络来提取图像特征。该网络中残差结构的引入解决了网络深度增加所带来的梯度消失的问题, 具有更强的特征提取能力,



注: w 代表特征图通道数, 其下标代表不同阶段; r 为输入特征图的长度和宽度; n 为最终输出的特征图通道数
Note: w is number of channels of characteristic matrix, and its subscripts represent different stages; r is the length and width of input feature map. n is number of channels in the final output feature map

图 3 RegNetX-6.4GF 结构
Fig.3 Structure of RegNetX-6.4GF

其在 ImageNet 数据集上的 Top-1 Error 相较于 VGG-16 的 28.07% 和 ResNeXt-50 的 21.9% 下降为 20.8%^[22]。

1.2.2 Guided Anchoring-FPN 模块

RPN 依据卷积神经网络提取到的特征图生成棉花顶芽的前景目标候选框。此阶段目标候选区域的质量高低直接影响检测器的性能，目标候选区域的质量高低表现为两个方面，即在更高的 IOU 阈值下是否有更多的目标候选区域和锚框 (Anchor) 形状与目标形状匹配。为解决上述问题，本文通过在 RPN 中将 FPN 与 Guided Anchoring 相融合，来改变原有 Anchor 生成机制。

FPN^[23]通过搭建一条自顶向下的路径将各层次特征图进行连接，实现了浅层特征和深层特征的融合。在生成的每一特征图层上，Faster R-CNN 按照特定的步长以滑框 (Sliding Window) 的方式生成的 Anchor。在棉花顶芽识别中，这种方式会带来两个问题，一是棉花顶芽所在区域过小，而生成 Anchor 则是在全图均匀分布，这样会造成多数 Anchor 分布于背景之中从而引起正负样本不均衡进而影响分类器的计算速度。二是生成的 Anchor 均是固定尺度和比例，不适用于自然状态下非固定尺度的棉花顶芽。故对于棉花顶芽的识别，需要形状位置动态可变的 Anchor 生成机制。

为解决上述问题，在 FPN 中引入 Guided Anchoring 机制^[24]。Anchor 的概率分布被分解为两个条件概率分布，即给定图像特征之后 Anchor 中心点的概率分布和给定图像特征和中心点之后的形状概率分布 (图 4)。

$$p(x, y, l, h | I) = p(x, y | I) \cdot p(l, h | x, y, I) \quad (8)$$

式中 x, y, l, h 分别表示 Anchor 的中心点坐标和宽高。

对于 Anchor 的中心点坐标的确定，采用 1×1 卷积对特征图进行处理，获得每点处是否含有物体的得分，与设定预设的阈值比较，本文将预设阈值设置为 0.01，将可能含有物体的位置作为中心点。在确定中心点坐标之后，使用 Bounded IOU Loss 对边框长宽进行回归计算，以获得最佳的长和宽。最后使用 3×3 的可变形卷积修正原始的特征图，使得特征图适应每个 Anchor 的形状。

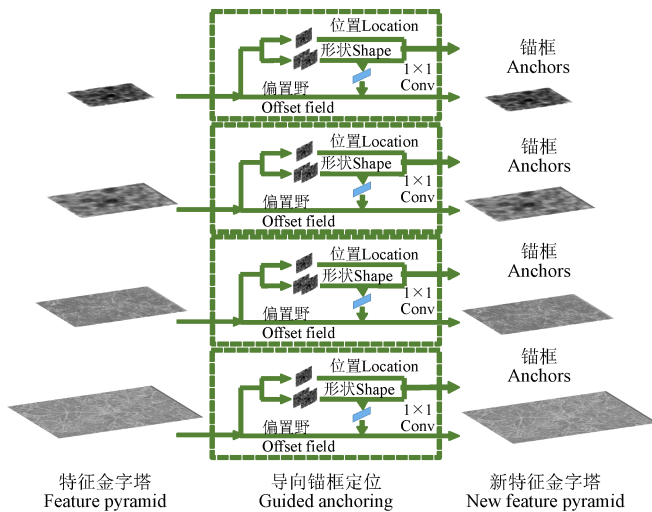


图4 导向锚框定位机制结构

Fig.4 Structure of Guided Anchoring (GA)

1.2.3 Generic ROI Extractor 模块

传统 Faster R-CNN 中通过式 (9) 从 RPN 选取特征图来提取 ROI。

$$k = k_0 + \log_2 \left(\frac{\sqrt{w_c h_c}}{224} \right) \quad (9)$$

式中 k 为所选取特征图， k_0 表示最高层次特征图，224 则是 ImageNet 中与训练卷积神经网络的图片尺寸。 w_c, h_c 代表 ROI 的宽度和高度。这种选择单一特征图的方法对于棉花顶芽识别会引起特征信息损失。为了综合多尺度特征，本文引入目标候选区域提取器 (Generic ROI Extractor)^[25] 来对改进型 Faster R-CNN 中的 RPN 信息进行综合提取。

如图 5 所示，Generic ROI Extractor 首先使用目标候选区域对齐^[26]将 ROI 映射到经 Guided Anchoring-FPN 处理的每层特征图对应位置处，之后再对每一层进行预处理，将不同尺寸的特征信息进行消融，最后使用注意力机制提取全局特征，最终经过分类回归计算得到棉花顶芽的识别结果。

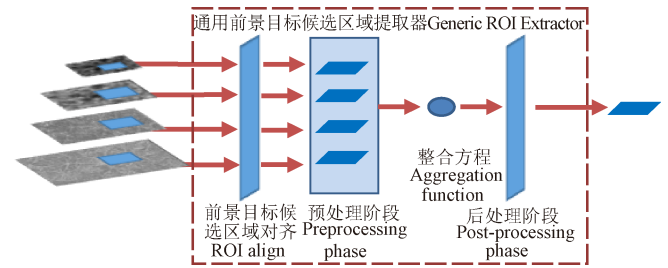


图5 Generic ROI Extractor 结构

Fig.5 Structure of Generic ROI Extractor

1.2.4 Dynamic R-CNN 机制

Faster R-CNN 中 Guided Anchoring-FPN 的引入，使得生成的 Anchor 可以根据棉花顶芽的特征信息进行动态调整，这一机制的改变使得训练过程中产生的候选区域在不同 IOU 阈值下的分布也在动态变化，而 Faster R-CNN 中只有阈值为 0.5 的 IOU。针对此问题，本文在 Faster R-CNN 中融合 Dynamic R-CNN^[27]，使得算法在第二阶段的分类器 (Classifier) 和回归计算器 (Regressor) 能动态适应 Proposal 的分布变化。在分类器中使用动态标注分类 (Dynamic Label Assignment, DLA)，见式 (10)。

$$\text{Label} = \begin{cases} 1, & \text{IOU}_{\max}(b, G) \geq T_{\text{now}} \\ 0, & \text{IOU}_{\max}(b, G) < T_{\text{now}} \end{cases} \quad (10)$$

式中 b 表示标框 (Bounding Box)， G 表示真实标注 (Ground Truths) 的集合， T_{now} 表示当前 IOU 阈值。

为保证分类器性能，参照 Dynamic R-CNN 模型已有试验效果进行相关参数选取，将初始的 IOU 阈值设定为 0.4，在训练中获取 Proposals 及其目标真实标注 (Ground Truths) 之间的 IOU 矩阵，选择第 75 个值^[27]，使用 S_l 存储该数值，每迭代 100 次使用 S_l 平均值来更新 T_{now} 。在边界框回归计算中，在传统的 Smooth L1 Loss 上增加一个参数 β_{now} ，构建动态 Smooth L1 损失函数 (Dynamic SmoothL1 Loss, DSL)，见式 (11)。

$$\text{DSL}(x, \beta_{\text{now}}) = \begin{cases} 0.5|x|^2/\beta_{\text{now}}, & |x| < \beta_{\text{now}} \\ |x| - 0.5\beta_{\text{now}}, & \text{其他} \end{cases} \quad (11)$$

式中 x 表示归类标签, β_{now} 表示当前超参数 β 值, 本文中 β_{now} 的初始设定为 1, 在训练中获得 Proposals 及其目标真实标注之间的回归误差 E , 选择 E 中第 10 个值^[27], 使用 S_E 存储该数值, 每迭代 100 次使用 S_E 中位数来更新 β_{now} 。

1.3 模型训练与评估

1.3.1 软硬件平台

本文方法所基于的软件平台: 系统软件为 Ubuntu18.04, Python3.7.9, Pytorch1.7.1 深度学习模型框架基于香港中文大学的 OpenMMLab 开源框架 MMDetection2.8。试验所用服务器硬件平台配置: 处理器 Intel Core i9-9900k, 主频为 3.6 GHz, 内存 32 G, 显卡 NVIDIA GeForce RTX 3080 (10 GB)。

1.3.2 模型训练方法

算法模型使用 MS COCO2017 格式的棉花顶芽数据集, 使用训练集 2 815 张图片对算法模型参数进行训练, 在验证集 704 张图片中比较不同算法性能, 并在测试集 1 300 张图片测试本算法实际识别效果。

在 MMDetection 框架平台中对模型进行训练, 使用随机梯度下降法训练模型, 权值衰减系数为 0.000 05, 动量参数为 0.9, 基于单 RTX 3080 显卡的 Batch 大小为 3。采用多尺度训练方案, 输入图片尺寸长边像素大小为 1 333, 短边像素尺寸每轮训练中从 [640, 672, 704, 736, 768, 800] 中随机挑选。学习率大小为 0.002 5, 设定训练 36 期 (Epochs), 学习率采用动态变化机制, 初始学习率为 0.001, 学习率在训练迭代 500 次中线性增加, 在训练迭代到 500 次时增加到 0.002 5 并保持不变, 之后学习率在第 28 和第 34 期减少为原来的 1/10, 来对模型参数进行精细调整。

1.3.3 模型评估方法

本文采用验证集来测试训练好模型的性能, 并在在测试集测试改进后的算法实际识别效果。以交并比 (Intersection Over Union, IOU) 来评价模型精度, 采用平均精度均值 (Mean Average Precision, MAP), 召回率 (Recall) 和准确率 (Precision) 进行模型评价^[13]。通常情况 MAP 数值越高越好, 本文只识别棉花顶芽, 类别数为一, 故本文采用召回率 (Recall) 和准确率 (Precision) 作为评价指标, 取值区间为 [0, 1]。

2 试验结果与分析

2.1 模型训练结果

算法模型在训练集 2 815 张图片中进行训练, 在 36 期中共迭代 35 000 次, 改进型 Faster R-CNN 由于引入了 Guided Anchoring 机制, 如式 (1) 所示, 其 Loss 函数在原有基础上增加了 Anchor 的位置损失和形状损失函数, 使得训练过程能获得更多高质量的 Anchor。图 6b, 6c 分别表示训练过程 Anchor 位置损失及形状损失曲线, 图 6d 表示训练过程正负样本的分类损失曲线, 由图可知, Anchor 位置损失、形状损失及分类损失在迭代 30 000 次

之后均收敛到稳定值。基于此机制生成 Anchor 的位置和形状能与数据集图像特征相契合, 提升了模型对棉花顶芽区域的学习能力和识别效果, 进而表现为总体损失函数值下降较快 (图 6a), 总体损失函数值在迭代 30 000 次之后达到稳定, 此时模型参数达到预期训练效果。

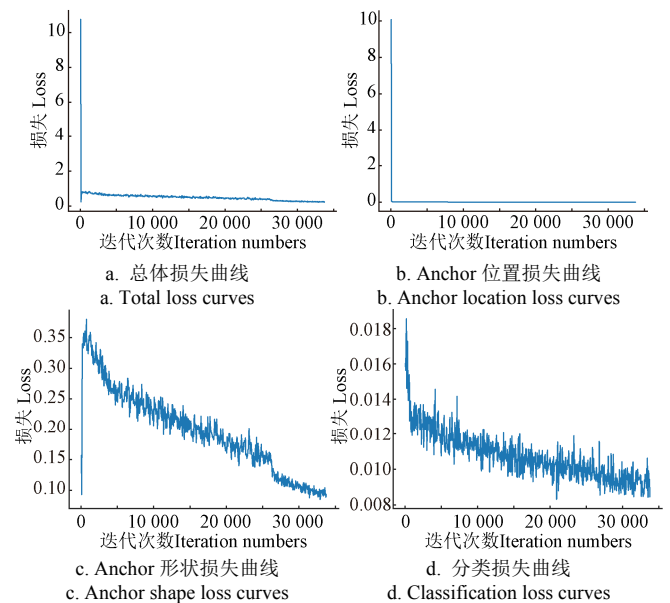


图 6 训练过程损失函数变化曲线

Fig.6 Loss curves of training

2.2 模型识别效果

为了验证模型的识别精度, 将训练好的模型应用于测试集, 对测试集中两种不同场景下的棉花图片进行检测, 场景一为中距离拍摄场景不同光照条件下的识别效果, 包含 966 张图片, 1 314 个棉花顶芽。场景二为远距离拍摄场景识别效果, 包含 334 张图片 764 个棉花顶芽。统计结果如表 1 所示。

表 1 棉花顶芽识别结果

方法 Method	场景 Scene	准确率 Precision	召回率 Recall	平均准确率均值 Mean Average Precision (MAP)
Faster R-CNN	中距离	95.3	93.6	94.9
	远距离	94.5	98.1	
改进型 Faster R-CNN Improved Faster R-CNN	中距离	98.5	96.8	98.1
	远距离	97.6	99.0	

通过对模型在测试集上的统计结果可以看出, 改进后的 Faster R-CNN 模型在大田自然环境下对棉花顶芽有较高的识别准确度。由表 1 可以看出, 本文所提方法的识别准确率在中距离和远距离场景下, 较 Faster R-CNN 分别提高了 3.2 和 3.1 百分点。针对所有样本识别的平均准确率均值可达 98.1%, 较 Faster R-CNN 高 3.2 个百分点。这是因为将 Faster R-CNN 中的卷积网络替换为 RegNetX-6.4GF, 提高了模型对目标信息的提取能力, 同时 Guided Anchoring 和 GROIE 机制的融合使得 Anchor 能够动态适应目标位置及大小变化, 增强了模型对于图像中棉花顶芽特征信息的识别能力。针对不同拍摄距离、不同光照条件的棉花顶芽识别效果如图 7 所示。



图7 不同距离不同光照条件下棉花顶芽识别效果

Fig.7 Recognition results of cotton terminal bud in different distances and light conditions

2.3 不同网络架构性能比较

目前在深度学习领域,以卷积神经网络为基础的目标识别模型主要有两大分支。一是 Anchor-based 分支,其内部又分为单阶段、双阶段及多阶段模型框架。二是 Anchor-free 分支,其主要通过确定特征点的方式来形成最终的识别框。

为验证本文算法的有效性,选取目前主流的识别算法在棉花顶芽数据集上进行训练,并在验证集上将其和本文提出的算法进行性能比较。本文选用 Anchor-based 中的基于单步多目标检测器 (Single Shot MultiBox Detector, SSD) 改进的单阶段模型 RetinaNet、双阶段模型 Faster R-CNN 和多阶段模型中采用级联检测器结构的 Cascade R-CNN^[28]以及 Anchor-free 中的 RepPoints^[29]模型框架进行对比分析。计算 IOU 值在 0.5 情况下的准确率^[13],并计算模型处理帧速 (Frames Per Second, FPS),结果如表 2 所示。

表2 不同算法在验证集上的测试结果

Table 2 Recognition results of different algorithms in the validation set

方法 Method	MAP/% (IOU=0.5)	帧速 Frames per second/(帧·s ⁻¹)
RetinaNet	45.7	21.0
Cascade R-CNN	74.3	19.0
RepPoints	75.5	21.2
Faster R-CNN	76.2	14.3
改进型 Faster R-CNN	81.8	10.3
Improved Faster R-CNN		

注: IOU 为交并比。Note: IOU is intersection over union.

通过表 2 可知,针对棉花顶芽数据集,本文提出的改进型 Faster R-CNN 在 IOU=0.5 时,其 MAP 相较于 Faster R-CNN 提高 7.3%,较 RetinaNet、Cascade R-CNN 和 RepPoints 分别提高 78.9%、10.1%和 8.3%。改进后的 Faster R-CNN 对图像信息具有更好的感知和学习能力,具有较高的精度。处理速度为 10.3 帧/s,满足棉花精准打顶实时识别速度需求。

3 结论

本文针对大田环境下棉花精准打顶作业过程的棉花

顶芽识别问题,在 Faster R-CNN 算法框架中基于棉花顶芽的小目标在复杂背景下的特点对算法进行的改进,提出一种针对棉花顶芽识别的改进型 Faster R-CNN 识别算法,主要结论如下:

1) 本文将 Faster R-CNN 中的卷积神经网络替换为 RegNetX-6.4GF,并将 Guided Anchoring 和 GROIE 机制进行融合,增强了模型对棉花顶芽的识别能力。所提方法在中距离和远距离场景下的识别准确率较 Faster R-CNN 分别提高 3.2 和 3.1 个百分点。

2) 将本文提出方法与主流识别算法 RetinaNet、Cascade R-CNN 和 RepPoints 进行比较,结果表明,本文方法检测的 MAP 在交并比为 0.5 时相比 Faster R-CNN 提高 7.3%,较 RetinaNet、Cascade R-CNN 和 RepPoints 分别提高 78.9%、10.1%和 8.3%。

3) 本文提出的改进型 Faster R-CNN 对棉花识别的平均准确率均值达 98.1%,处理帧速为 10.3 帧/s,满足棉花打顶作业对于精确度和实时性的要求,可为棉花精准打顶装备的研发奠定基础。

[参 考 文 献]

- [1] 李继福,何俊峰,陈佛文,等. 中国棉花生产格局与施肥研究现状—基于 CNKI 数据计量分析[J]. 中国棉花, 2019, 46(4): 17-24, 28.
- [2] 卢秀茹,贾肖月,牛佳慧. 中国棉花产业发展现状及展望[J]. 中国农业科学, 2018, 51(1): 26-36.
Lu Xiuru, Jia Xiaoyue, Niu Jiahui, et al. The present situation and prospects of cotton industry development in China[J]. Scientia Agricultura Sinica, 2018, 51(1): 26-36. (in Chinese with English abstract)
- [3] 刘照启,张蔚然,刘海涛,等. 棉花打顶技术应用现状与发展趋势[J]. 现代农村科技, 2020(7): 16.
- [4] 邓劲松,石媛媛,陈利苏,等. 基于近红外传感器和面向对象光谱分割技术的田间棉株识别与提取[J]. 光谱学与光谱分析, 2009, 29(7): 1754-1758.
Deng Jinsong, Shi Yuanyuan, Chen Lisu, et al. Cotton identification and extraction using near infrared sensor and object-oriented spectral segmentation technique[J]. Spectroscopy and Spectral Analysis, 2009, 29(7): 1754-1758. (in Chinese with English abstract)
- [5] 孙杰. 棉花打顶机打顶高度自动控制系统的设计与试验研究[D]. 乌鲁木齐: 新疆农业大学, 2015.
Sun Jie. Design and Experimental Research of An Automatic Height Control System for Topping of Cotton Top-cutting Machine[D]. Urumqi: Xinjiang Agricultural University, 2015. (in Chinese with English abstract)
- [6] 李霞,王维新,张若宇. 超声波测高技术棉花打顶中的应用[J]. 石河子大学学报: 自然科学版, 2009, 27(3): 358-360.
Li Xia, Wang Weixin, Zhang Ruoyu. Application of cotton altitude automation measurement based on ultrasonic technology[J]. Journal of Shihezi University: Natural Science, 2009, 27(3): 358-360. (in Chinese with English abstract)

- [7] 刘俊奇. 棉花株顶识别系统的研究[D]. 石河子: 石河子大学, 2009.
Liu Junqi. The Research of Automatic Recognition of Cotton's Top[D]. Shihezi: Shihezi University, 2009. (in Chinese with English abstract)
- [8] 瞿端阳. 基于机器视觉技术的棉株识别系统研究[D]. 石河子: 石河子大学, 2013.
Qu Duanyang. Cotton Plant Recognition System based on the Machine Vision Technology[D]. Shihezi: Shihezi University, 2013. (in Chinese with English abstract)
- [9] Alotaibi A, Mahmood A. Deep face liveness detection based on nonlinear diffusion using convolution neural network[J]. Signal, Image and Video Processing, 2017, 11(4): 713-720.
- [10] Bautista C M, Dy C A, Mañalac M I, et al. Convolutional neural network for vehicle detection in low resolution traffic videos[C]. 2016 IEEE Region 10 Symposium (TENSYP). Bali, Indonesia, IEEE, 2016: 277-281.
- [11] Tajbakhsh N, Shin J Y, Gurudu S R, et al. Convolutional neural networks for medical image analysis: Full training or fine tuning[J]. IEEE Transactions on Medical Imaging, 2016, 35(5): 1299-1312.
- [12] 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]. 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]. Transactions of the Chinese Society for Agricultural Machinery, 2019, 50(3): 144-150. (in Chinese with English abstract)
- [13] 熊俊涛, 刘振, 汤林越, 等. 自然环境下绿色柑橘视觉检测技术研究[J]. 农业机械学报, 2018, 49(4): 45-52.
Xiong Juntao, Liu Zhen, Tang Linyue, et al. Visual detection technology of green citrus under natural environment[J]. Transactions of the Chinese Society for Agricultural Machinery, 2018, 49(4): 45-52. (in Chinese with English abstract)
- [14] Nazki H, Yoon S, Fuentes A, et al. Unsupervised image translation using adversarial networks for improved plant disease recognition[J]. Computers and Electronics in Agriculture, 2020, 168: 105-117.
- [15] 彭明霞, 夏俊芳, 彭辉. 融合 FPN 的 Faster R-CNN 复杂背景下棉田杂草高效识别方法[J]. 农业工程学报, 2019, 35(20): 202-209.
Peng Mingxia, Xia Junfang, Peng Hui. Efficient recognition of cotton and weed in field based on Faster R-CNN by integrating FPN[J]. Transactions of the Chinese Society of Agricultural Engineering (Transactions of the CSAE), 2019, 35(20): 202-209. (in Chinese with English abstract)
- [16] 姚青, 谷嘉乐, 吕军, 等. 改进 RetinaNet 的水稻冠层害虫为害状自动检测模型[J]. 农业工程学报, 2020, 36(15): 182-188.
Yao Qing, Gu Jiale, Lyu Jun, et al. Automatic detection model for pest damage symptoms on rice canopy based on improved RetinaNet[J]. Transactions of the Chinese Society of Agricultural Engineering (Transactions of the CSAE), 2020, 36(15): 182-188. (in Chinese with English abstract)
- [17] 雷杰杰, 邵盘霞, 郭春平, 等. 新疆陆地棉经济性状优异等位基因位点的遗传解析[J]. 棉花学报, 2020, 32(3): 185-198.
Lei Jiejie, Shao Panxia, Guo Chunping, et al. Genetic dissection of allelic loci associated with economic traits of upland cottons in Xinjiang[J]. Cotton Science, 2020, 32(3): 185-198. (in Chinese with English abstract)
- [18] 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.
- [19] Lin T Y, Goyal P, Girshick R, et al. Focal loss for dense object detection[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2020, 42(2): 318-327.
- [20] Simonyan K, Zisserman, A. Very deep convolutional networks for large-scale image recognition[C]. International Conference on Learning Representations, San Diego, CA, USA, 2015: 1-14.
- [21] Szegedy C, Liu W, Jia Y, et al. Going deeper with convolutions[C]. Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, Boston, MA, USA, 2015: 1-9.
- [22] Radosavovic I, Kosaraju R P, Girshick R, et al. Designing network design spaces[C]. Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, Seattle, WA, USA, 2020: 10425-10433.
- [23] Lin T, Dollar P, Girshick R, et al. Feature pyramid networks for object detection[C]. IEEE Conference on Computer Vision and Pattern Recognition, Honolulu, Hawaii, USA, 2017: 936-944.
- [24] Wang J, Chen K, Yang S, et al. Region proposal by guided anchoring[C]. Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, Long Beach, CA, USA, 2019: 2960-2969.
- [25] Rossi L, Karimi A, Prati A. A novel region of interest extraction Layer for Instance Segmentation[C]. International Conference on Pattern Recognition, Milan, Italy, 2021: 2203-2209.
- [26] He K, Gkioxari G, Dollar P, et al. Mask R-CNN[C]. IEEE International Conference on Computer Vision, Venice, Italy, 2017: 2980-2988.
- [27] Zhang H, Chang H, Ma B, et al. Dynamic R-CNN: Towards high quality object detection via dynamic training[C]. European Conference on Computer Vision, Glasgow, UK, 2020: 260-275.
- [28] Cai Z, Vasconcelos N. Cascade R-CNN: Delving into high quality object detection[C]. Conference on Computer Vision and Pattern Recognition, Salt Lake City, UT, USA, 2018: 6154-6162.

- [29] Yang Z, Liu S, Hu H, et al. Reppoints: Point set representation for object detection[C]. International Conference on Computer Vision, Seoul, Korea, 2019: 9656-9665.

Recognition of cotton terminal bud in field using improved Faster R-CNN by integrating dynamic mechanism

Chen Keyi¹, Zhu Longfu¹, Song Peng^{1*}, Tian Xiaomin², Huang Chenglong³, Nie Xinhui², Xiao Ailing⁴, He Liangrong⁵

(1. College of Plant Science and Technology, Huazhong Agricultural University, Wuhan 430070, China; 2. Agricultural College of Shihezi University, Shihezi 832003, China; 3. College of Engineering, Huazhong Agricultural University, Wuhan 430070, China; 4. College of Mechanical and Electronic Engineering, Tarim University, Alaer 843300, China; 5. College of Plant Sciences, Tarim University, Alaer 843300, China)

Abstract: Accurate identification of cotton top bud is important for cotton topping operation to detect cotton terminal bud accurately in field, a recognition method using Faster R-CNN (Faster Region Convolutional Neural Networks, Faster R-CNN) by integrating dynamic mechanism was proposed to solve the recognition difficulties caused by the small size of cotton terminal bud when it is topped in densely planted fields. The RegNetX-6.4GF model was used as the backbone network to improve its image feature extractor capabilities. Due to number of proposals under a higher IOU (Intersection Over Union, IOU) as well as the matching degree between anchor and the target shape affect the performance of the detector, the method proposed in this paper changed the original anchor generation mechanism by combining FPN (Feature Pyramid Network, FPN) and Guided Anchoring in RPN (Region Proposal Network, RPN), which will cause the distribution of the proposals generated by the RPN of the algorithm under different IOUs dynamically change during the training process. To adapt the dynamic change of proposals distribution, we integrated Dynamic Region Convolutional Neural Networks (Dynamic R-CNN) in Faster R-CNN, which can dynamically adjust the value of IOU to obtain high quality proposals. And the GROIE (Generic ROI Extractor, GROIE) mechanism was inducted to extract ROI (Region of Interest, ROI) to improve the feature fusion capability. In this paper, 4 819 images of *Gossypium hirsutum* population which contain seven leaf types were taken from the top of cotton plant at distance of 30-50 cm (medium distance) and 50-100 cm (long distance) under uniform light, oblique strong light, direct strong light, and shadows. Those images were processed as MS COCO 2017 format dataset and assigned to the training set, validation set, and test set, which contained 2 815, 704, and 1 300 pictures respectively. The experimental results demonstrated that FPS (Frames Per Second, FPS) of proposed model was up to 10.3 frames/s and the Mean Average Precision (MAP) of bud identification reached to 98.1% which was 3.2 percentage points higher than original Faster R-CNN model. The validation set were used to compare performance of mainstream recognition algorithm and proposed method. Results showed that the improved Faster R-CNN's MAP was 7.3% higher than original Faster R-CNN, which was also higher than RetinaNet, Cascade R-CNN (Cascade Region Convolutional Neural Networks, Cascade R-CNN) and RepPoints by 78.9%, 10.1% and 8.3% when IOU was set to 0.5. The improved Faster R-CNN proposed in this paper meets the accuracy and real-time requirements of cotton topping operation.

Keywords: deep learning; algorithm; cotton; mechanism fusion; dynamic adaptation; terminal bud recognition; Faster R-CNN