

doi:10.6041/j.issn.1000-1298.2023.03.026

基于无人机多光谱影像和关键点检测的雪茄烟株数提取

饶雄飞¹ 周龙宇² 杨春雷¹ 廖世鹏^{2,3} 李小坤^{2,3} 刘诗诗²

(1. 湖北省烟草科学研究院, 武汉 430030; 2. 华中农业大学资源与环境学院, 武汉 430070;
3. 农业农村部长江中下游耕地保育重点实验室, 武汉 430070)

摘要: 为从无人机遥感影像中准确识别烟草,实现植株定位与计数,以雪茄烟草植株为研究对象,提出一种新的深度学习模型。区别于传统的利用检测框识别目标,本文模型利用少量的关键点学习烟草中心形态学特征,并采用轻量级的编、解码器从无人机遥感影像快速识别烟草并定位计数。首先,提出的模型针对烟草植物形态学特点,通过中心关键点标注的方法,使用高斯函数生成概率密度图,引入更多监督信息。其次,对比不同主干网络在模型中的效果,ResNet18 作为主干网络时平均精度大于 99.5%,精度和置信度都高于测试的其他主干网络。而 MobileNetV2 在 CPU 环境下达到运行效率最优,但平均置信度相对较低。使用损失函数 Focal Loss 与 MSE Loss 结合的 Union Loss 时,平均精度大于 99.5%。最后,利用不同波段组合作为训练数据,对比结果发现使用红边波段更有助于模型快速收敛且能够很好地区分烟草和杂草。由于红边波段与植株冠层结构相关,使用红边、红、绿波段时平均精度达到 99.6%。本文提出的深度学习模型能够准确地检测无人机遥感影像中的烟草,可为烟草的农情监测提供数据支持。

关键词: 雪茄烟草; 无人机多光谱遥感; 关键点检测; 深度学习

中图分类号: TP391.4 文献标识码: A 文章编号: 1000-1298(2023)03-0266-08 OSID: 

Counting Cigar Tobacco Plants from UAV Multispectral Images via Key Points Detection Approach

RAO Xiongfei¹ ZHOU Longyu² YANG Chunlei¹ LIAO Shipeng^{2,3} LI Xiaokun^{2,3} LIU Shishi²

(1. Tobacco Research Institute of Hubei Province, Wuhan 430030, China
2. School of Resources and Environment, Huazhong Agricultural University, Wuhan 430070, China
3. Key Laboratory of Arable Land Conservation in Middle and Lower Reaches of Yangtze River, Ministry of Agriculture and Rural Affairs, Wuhan 430070, China)

Abstract: Tobacco is an important industrial crop in China. The survival rate and growing status of tobacco plants after being transplanted to the field are essential for the field management and yield predictions. However, counting the number of live plants is traditionally conducted by labors, which is time consuming and expensive. Unmanned aerial vehicle is a cost-effective option for monitoring croplands and plantations. However, visual inspection for such images can be a challenging and biased task, specifically for locating and detecting plants. As tobacco plant has a characteristic center-oriented feature, a novel deep-learning algorithm was developed to locate and count tobacco plants via key points detection method, instead of using a common bounding-box object-detection approach. The proposed deep learning algorithm was tested on the cigar plants. In the algorithm, the center of each plant was firstly annotated with a point, and a Gaussian probability density was derived to provide useful information of morphological features. Secondly, different backbones and loss functions in the proposed algorithm were evaluated. Using ResNet18 as a backbone provided the most accurate prediction of the plant number (average precision higher than 99.5%). MobileNetV2 was the most efficient backbone, but the uncertainty of predictions was higher than that of ResNet18. The combination of Focal Loss function and MSE Loss function (Union loss) reached the highest accuracy (average precision higher than 99.5%)

收稿日期: 2022-05-22 修回日期: 2022-07-07
基金项目: 国家自然科学基金面上项目(42171350)和湖北省烟草公司科技项目(027Y2021-009)
作者简介: 饶雄飞(1981—),男,高级农艺师,主要从事烟草栽培研究,E-mail: raoxiongfei@163.com
通信作者: 刘诗诗(1983—),女,副教授,主要从事农业遥感研究,E-mail: sslu@mail.hzau.edu.cn

while reduced the uncertainty. Finally, the evaluation of different combinations of multispectral bands showed that the combination of red-edge, red, and green bands had a better performance than using red, green, and blue bands in differentiating tobacco plants and weeds, resulting in less uncertainty in the tobacco plant detection. The proposed algorithm can accurately locate and count tobacco plant in the UAV images, providing an effective tool and a valuable data support for planting high-quality tobaccos.

Key words: cigar tobacco; UAV multispectral remote sensing; key point detection; deep learning

0 引言

烟草是我国重要经济作物,在我国国民经济中占有重要的地位^[1]。烟叶生产周期长、种植劳动强度大、技术要求高。烟草的产量与烟草移栽后烟苗成活率紧密相关,而移栽后的长势影响烟叶的追肥和管理。目前烟草株数统计主要依靠人工,工作量大、耗时耗力^[2]。随着无人机在轻简化、稳定性方面的快速发展,无人机遥感技术已经在农作物植保、施肥、监测中广泛应用^[3-5]。利用无人机遥感影像自动识别作物株数,监测出苗情况不仅可以节约人力物力成本,还可以为大范围的出苗和长势监测提供准确信息,特别适用于附加值高的经济作物种植管理。

目前,利用无人机遥感自动识别作物植株的研究大致可以分为基于传统数字图像处理的方法和基于机器学习方法两类。数字图像分割方法(如大津法)能够有效区分作物植株和背景,但会有植株粘连和杂草影响的问题^[6]。除了数字图像的分割方法,角点检测算法(如 Moravec 算子)也可以有效识别作物植株^[7]。基于机器学习的方法提取作物株数则是提取无人机影像上作物的形态学特征和光谱特征,利用机器学习算法建立特征和株数间的关系,达到估测株数的目的^[8-10]。

近年来,随着深度学习在计算机视觉领域的不断进步,用于作物计数的方法也在不断更新。目前,物体计数的方法大体可分为:基于目标检测的方法、基于图像分割的方法和密度图计数的方法^[11]。基于目标检测的计数方法,如 YOLOv3^[12]、Faster R-CNN^[13]等,通过对物体生成外接矩形框,进而得到物体位置并统计得到数量。基于图像分割的方法,如 FCN^[14]、U-Net^[15]等,通过分割出前景与背景区域,利用连通域来得到物体边界进而统计数量。基于密度图计数的方法,如 MCNN^[16]、TasselNet^[11],通过对卷积网络得到的密度图求积分,从而得到图像上物体的数量。在上述方法中,基于目标检测与图像分割的方法会受到植株粘连与严重遮挡的影响^[11],基于密度图求积分的方法主要针对过于复杂场景中的计数问题,如麦穗计数、玉米抽穗计数等。

然而对于烟草定位与计数问题,由于烟草对象

的复杂特性,如植株非刚体、边界遮掩现象明显等,无需清楚地划分烟草对象的边界信息,而更应该关注其空间位置信息。由于烟草是茄科烟草属植物,具有强大的圆柱形直立主茎,在无人机影像上表现为由中心向外扩展的特点,通过关键点检测来识别烟草中心区域,从而识别烟草并统计烟草株数。目前,关键点检测主要方法可归为两类,即利用 CNN 对关键点坐标进行回归^[17]或在标注点坐标位置通过二维高斯函数生成概率密度图^[18]进而定位,前者训练与推理速度快而泛化能力不足,且目标对象所在区域难以被某一个像素位置所定义,而后者模型更加复杂但更能充分利用空间信息,更加符合卷积神经网络对空间捕捉的特点,使得网络更容易收敛。

本文针对雪茄烟草识别的关键期采用概率密度图的思想,提出一种新的轻量级高效识别烟草的模型来解决烟草检测问题。该模型结构在使用不同主干分类网络作为编码器的基础上,将轻量级特征融合模块作为解码器部分,使得网络模型能有效减少学习参数的数量。利用烟草植物形态学特点使用点级别注释,使数据标注更加便捷,同时引入多光谱信息、随机遮掩等数据增强方法,使得模型不仅能准确有效识别烟草并且更具有鲁棒性。

1 数据获取与方法

1.1 研究区概况

研究区位于湖北省十堰市丹江口市习家店镇(32.778 427 8°N, 111.183 133 0°E)的烟草种植示范区(图 1, I 区为训练集, II 区与 III 区为验证集)。习家店镇属于亚热带季风气候,年平均气温 21.6℃,年平均降水量 750 ~ 900 mm。习家店镇烟草种植示范区总面积约为 64 hm²,种植的烟草品种为茄衣-26 和茄芯-14。烟苗均由习家店镇五龙池村育苗基地提供。茄衣是雪茄烟的最外层,也是雪茄烟叶中最昂贵的部分。茄芯是雪茄最中心的部分,是每支雪茄的核心。本研究在烟苗移栽成活后的伸根期开展。伸根期烟草生长迅速,需要加强田间管理,为下一阶段的旺盛生长打好基础。在此期间识别烟草不仅能够计算成活率,还可以为精准的田间管理和长势监测提供数据支持。

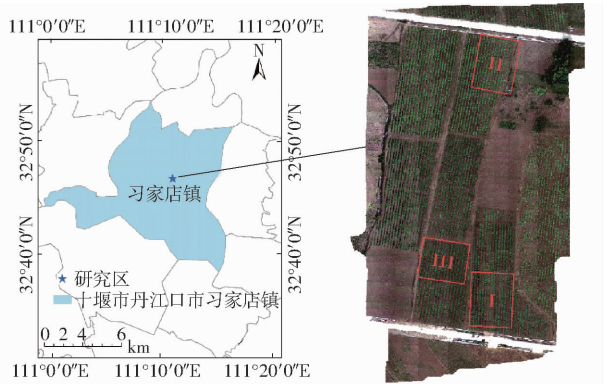


图 1 研究区域

Fig. 1 Map of study area

1.2 数据获取及标注

1.2.1 无人机多光谱影像采集与预处理

本研究于 2020 年 5 月 29 日在烟草伸根期利用大疆精灵 4 多光谱版无人机拍摄烟草种植区(部分区域约 0.53 hm²)。精灵 4 多光谱版是一款配备一体式多光谱成像系统的航测无人机,集成 1 个可见光镜头及 5 个多光谱镜头。5 个多光谱波段分别是蓝波段(中心波长 450 nm)、绿波段(中心波长 560 nm)、红波段(中心波长 650 nm)、红边波段(中心波长 730 nm)、近红外波段(中心波长 840 nm)。无人机飞行高度为 20 m,地面空间分辨率约为 1.1 cm,影像之间有 80% 航向重叠和 70% 旁向重叠。选择晴朗无云天气来拍摄无人机影像,拍摄时间为当天 12:00—14:00,在拍摄期间使用自动曝光。拍摄研究区域大概需要 30 min,假定光照在 30 min 内的变化忽略不计。拍摄区域内放置 3 个定标毯(Terracam Inc., Chatsworth, 美国),反射率分别为 0.06、0.36、0.56,用于影像的辐射定标。辐射定标是通过建立定标毯反射率和无人机影像各个波段像元亮度值(Digital number, DN)的线性关系,将各个波段的 DN 转换反射率的过程。辐射定标后,将拍摄的所有影像进行拼接。辐射定标和影像拼接均利用 DJI Terra 版本 3.1.4 完成。

1.2.2 数据标注

由于烟草是具有中心特征的植物,本文采用与人体姿态估计^[17]的点标注相同的方法,对烟草中心位置进行标注。点标注的方法利用烟草中心信息,所以对图像的品质没有较高要求。因为点标注只需要对其中心进行标注,故标注速度快,节省人力成本。本研究对拼接后的无人机影像选择 3 个子区域进行标注,如图 1 所示,Ⅰ区用作训练集,Ⅱ区与Ⅲ区用作验证集。Ⅰ区和Ⅱ区种植的是茄衣-2,Ⅲ区种植的是茄芯-14。Ⅱ区选在烟草种植区与农户闲田种植区的交界处,用于检验所提出的方法是否能

将烟草与杂草进行区分。通过人工标注,训练集区域内共 752 株烟草,验证集区域内共 1 127 株烟草。在伸根期,茄衣和茄芯的冠层特征和色泽都极为相似,因此本研究对茄衣和茄芯不做区分,而是统一进行识别。

1.2.3 数据增强

深度学习模型训练需要足够的数据进行支撑,但实际情况中获取数据存在一定难度。在烟草定位与计数中,无人机影像会受到光照、拍摄角度、影像拼接等因素的影响,因此需要对原始影像进行数据增强。采用的数据增强方法如图 2 所示,主要包括:随机裁剪、随机遮掩、随机改变亮度、随机改变对比度、随机改变饱和度等^[19]。

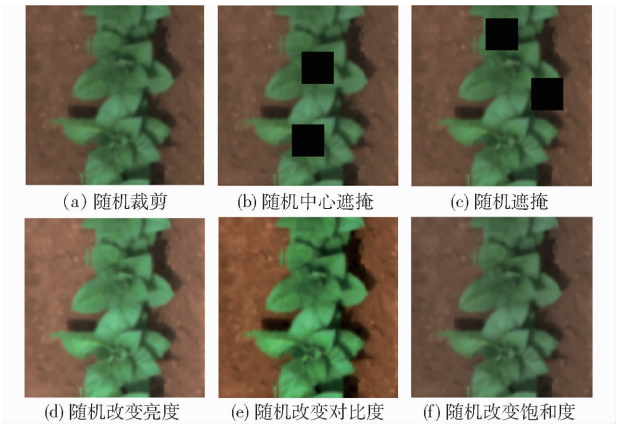


图 2 数据增强示例

Fig. 2 Examples of image augmentation

1.3 烟草计数与定位方法

1.3.1 概率密度图

采用与 CornerNet^[20]相似的方式,将烟草点标注的数据用于生成概率密度图,与之不同的是预测概率密度图结果保留与原图尺寸相同。首先,输入的图像 $I \in \mathbf{R}^{W \times H \times 3}$,点标注生成概率密度图为 $Y \in [0,1]^{W \times H}$,其中 W 和 H 分别为图像的宽与高。对于每幅图像标签点 $p(x_p, y_p)$,通过高斯核函数在其附近生成一个范围在 0~1 间的二维正态分布。高斯核函数为

$$Y_{xy} = \exp \left(-\frac{(x - x_p)^2 + (y - y_p)^2}{2\sigma_p^2} \right) \tag{1}$$

式中 (x, y) ——像素坐标
 Y_{xy} ——高斯核函数计算值
 σ_p ——控制高斯核尺寸参数

根据烟草平均尺寸,设置 σ_p 为 10。高斯概率密度图表示为越靠近标签中心点区域,其预测为中心关键点的概率越大,越远离中心点的区域其预测为中心关键点的概率越小。

1.3.2 烟草识别模型

随着深度学习中卷积神经网络的发展,许多神

神经网络模型在图像识别领域表现优越。为了准确定位并识别烟草的中心,本文基于人体关键点检测^[18]的策略,构建出多尺度特征融合且性能与精度兼顾的模型,如图3所示。首先,将影像随机裁剪后得到的尺寸为112像素×112像素的图像两倍上采样至尺寸224像素×224像素,从而增强烟草中心特征信息。在编码器部分,本研究使用3种不同的分类网络模型作为主干网络,即VGG^[21]、ResNet^[22]、MobileNetV2^[23],其中MobileNetV2是一种轻量级分类网络。在解码器部分,为快速且高效识别烟草,本研究提出更加轻量的特征融合模块。以ResNet18作为主干网络为例,如图3所示,“Conv”表示卷积

层,“F”表示特征图像,“BN”为批归一化层,“ReLU”为激活函数,“CBR”是由卷积核、批归一化层和激活函数组成的模块,将网络主干部分得到的多尺度特征统一上采样至尺寸56像素×56像素进行叠加,使得模型能感受不同尺度特征,并通过1×1卷积核进行特征降维,将不同尺度的特征进行融合。再经过2倍上采样至尺寸112像素×112像素的特征图保持输出图像尺寸与原图像一致。在网络预测末端部分,经过3×3卷积与1×1卷积核进行降维得到特征图,最后通过Sigmoid激活函数将预测概率限制在[0,1]范围内,从而得到概率密度图。

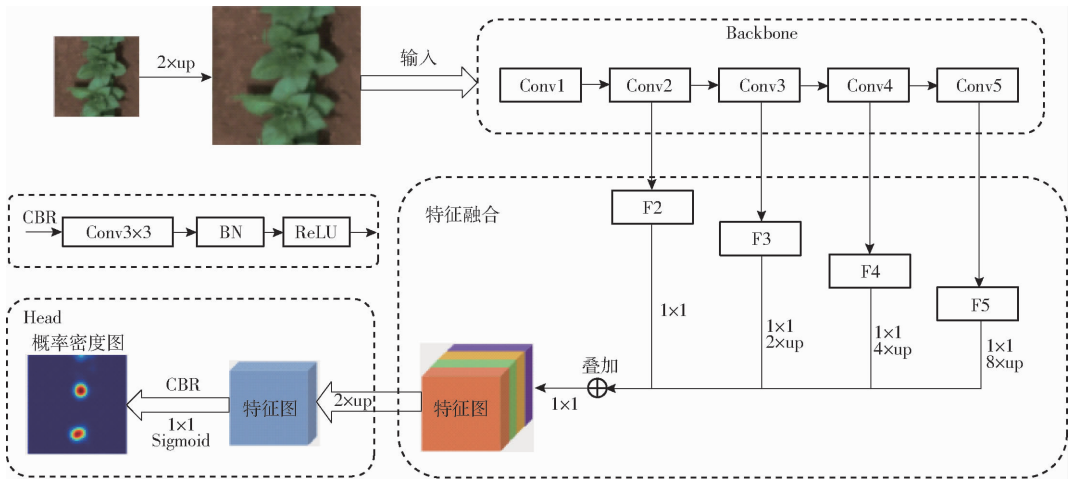


图3 烟草识别模型网络结构图

Fig.3 Architecture of deep learning model for detecting tobacco plants

由于无人机影像分辨率高,在本研究中按像元尺寸为80的步长依次裁剪影像,裁剪后的影像作为模型输入进行预测,同时使用滑动窗口预测策略消除拼接带来的裂隙问题^[15]。将模型预测得到的概率密度图拼接完毕后,通过高斯滤波对预测概率密度图进行平滑处理。以0.1为阈值,对低于阈值的概率进行过滤,经过局部峰值检测得到无人机影像上烟草中心关键点。统计关键点数量,得到烟草株数。

1.3.3 损失函数

采用常用的两种损失函数,即Focal Loss^[24]与MSE Loss,并在两者基础上提出用于烟草中心检测的联合损失函数Union Loss。Focal Loss为动态缩放的交叉熵损失函数,用于解决样本不平衡问题。本文设置概率 $Y > 0.6$ 部分为正样本区,其余部分为负样本区。同时给予正样本区间有更高的权重,损失函数具体形式参照文献[24],其中权重系数 α 设置为0.8。MSE Loss是用于回归计算的损失函数,是用于回归每个像元对应的概率,其具体形式参照文献[24]。

在MSE Loss作为损失函数训练模型中,模型在低概率与高概率区域出现过拟合,不利于对目标物体进行定位与识别。因此本研究提出使用Union Loss,其结合了Focal Loss与MSE Loss的部分计算公式,在 $Y \geq 0.9$ 高概率区域内即接近烟草中心的区域使用Focal Loss,提高模型预测值,而在 $Y < 0.2$ 低概率区域即远离烟草中心的区域使用Focal Loss,来降低模型预测值。同时在 $Y \in [0.2, 0.9]$ 区间内即过渡区域中使用MSE Loss来进行回归,能有效减缓MSE Loss作为损失函数过拟合情况。其具体形式为

$$L_{UL} = \frac{1}{N_{wh}} \sum_{i=1}^{N_{wh}} f(\hat{Y}, Y) \tag{2}$$

其中

$$f(\hat{Y}, Y) = \begin{cases} -\alpha_1 \ln(1 - \hat{Y}) & (Y < \delta_1) \\ \alpha_2 (\hat{Y} - Y)^2 & (\delta_1 \leq Y < \delta_2) \\ -\alpha_3 \ln \hat{Y} & (Y \geq \delta_2) \end{cases} \tag{3}$$

$$\alpha_1 + \alpha_2 + \alpha_3 = 1 \tag{4}$$

式中 δ_1, δ_2 ——区间阈值,分别取0.2、0.9
 $\alpha_1, \alpha_2, \alpha_3$ ——比例因子,分别取0.5、0.3、0.2
 $\hat{Y}$ ——模型预测到的概率

L_{UL} ——Union Loss 损失函数
 N_{wh} ——图像中像元数量

1.3.4 评价指标

为了评估烟草识别模型的性能,本文根据验证集数据计算了平均精度 (Average precision, AP)、精确率 (Precision)、召回率 (Recall)、平均置信值以及运行效率 (Frames per second, FPS)。平均精度是所有类别的精度与召回率曲线下的面积,在本研究中,目标对象仅为烟草,故无需求取平均精度均值,综合评价指标 F1 值为用于权衡精确率与召回率的评估量。在本研究中设置预测关键点与标注中心点距离在 10 个像元内为阳性样本,否则为错误样本。精确率为预测的阳性样本与实际阳性样本的比例,召回率为观察到的阳性样本与正确预测为阳性样本的比例。

置信度得分是用来描述和确认当前检测目标为某一个标记类别的接近程度,在本研究中指模型检测烟草为中心点的置信程度,在数值上指模型对预测结果的把握程度。平均置信度得分用来表示所有关键点置信度得分的均值,其具体形式为

$$F_c = \frac{1}{N_t} \sum_{i=1}^{N_t} p_i \tag{5}$$

式中 N_t ——影像中所有烟草数量
 p_i ——模型预测关键点置信度得分

为了验证模型的运行效率,在本研究中在 CPU 与 GPU 环境下分别求取模型的运行效率。运行效率 f_{FPS} 为 1 s 内检测图像的数量,具体形式为

$$f_{FPS} = \frac{1}{T(x)} \tag{6}$$

式中 T ——预测单幅图像所需的时间

1.3.5 模型训练环境与方法

本文实验所使用网络均在同一环境下训练,模型使用 PaddlePaddle 2.2.2 搭建网络,开发语言为 Python 3.7.4, GPU 型号为 Tesla V100, CPU 为 2 Cores,内存为 16 GB。在模型训练过程中,使用随机变量对网络模型进行初始化,模型主干部分未使用迁移学习,并采用 Momentum 梯度下降算法,动量因子为 0.9,初始学习率为 0.02,使用 Cosine Annealing 学习率衰减策略,权重衰减设为 0.000 1。为了防止模型过度拟合烟草特征,模型训练迭代次数总计 10 个迭代周期 (Epoch),批量大小 (Batch size) 设置为 32。

2 结果与分析

2.1 对照实验

2.1.1 不同主干网络精度对比

为验证不同网络的精度与运行效率,选择 ResNet18、ResNet34、VGG16 与 MobileNetV2 作为主干网络进行对比,并分别在 GPU 与 CPU 环境下求取运行效率。训练数据为训练集 I 区 RGB 波段同一随机裁剪得到的结果,使用学习方法与学习率以及后处理参数完全相同。模型检验基于验证集 II 与 III 区域中对概率密度图拼接后的统计结果。不同主干网络的烟草识别结果如表 1 所示。

表 1 不同主干网络评价指标对比
Tab.1 Performance of architectures with different backbones

主干网络	验证集	平均精度/%	精确率/%	召回率/%	F1 值/%	F_c	$f_{FPS}/(f \cdot s^{-1})$	
							GPU	CPU
ResNet18	II	99.9	90.9	100	95.2	0.867	99	9
	III	99.6	94.0	99.7	96.8	0.921	93	9
ResNet34	II	99.5	87.7	99.8	93.4	0.855	68	6
	III	99.8	93.7	99.8	96.7	0.940	62	5
VGG16	II	98.7	87.5	99.4	93.1	0.803	143	2
	III	99.5	92.3	99.8	95.9	0.930	137	2
MobileNetV2	II	99.9	92.1	100	95.9	0.845	64	12
	III	99.4	93.4	99.8	96.5	0.861	62	12

由表 1 可知,以 ResNet18 作为主干网络在精度与运行效率上表现相对最优,而以 MobileNetV2 作为主干网络在模型参数量更具有优势,便于在 CPU 环境下进行部署。对比 ResNet18 与 ResNet34 作为主干网络,随着主干网络特征层增多,烟草的检测精确率与召回率并没有显著提升,而运行效率在 CPU 与 GPU 上明显降低。由于 VGG16 无残差快捷连接模块,在 GPU 环境下表现为运行效率最优,但平均

置信度得分相对不足。MobileNetV2 作为轻量级主干网络,通过使用深度可分离卷积减少运算量与参数量,在 CPU 环境下达到运行效率最优,同时参数量最少,但平均置信度相对较低。

2.1.2 不同损失函数对比

为了验证不同损失函数对模型精度的影响,在实验中以同一 GPU 环境下使用 ResNet18 作为主干网络进行训练。训练数据为训练集 I 区 RGB 波段

影像随机裁剪得到的结果,使用学习率以及后处理参数完全相同。利用验证集评价不同损失函数的烟草识别效果,结果表明本研究提出的 Union Loss 在验证集上的精度、召回率、置信度得分均为最高(表 2)。使用 Focal Loss 用于解决样本不平衡问题较好,但在训练过程中出现未收敛的情况,精度与召回率较低。使用 MSE Loss 在验证集上精度、召回率相对较高,但置信度得分相对较低。结合 Focal Loss 与 MSE Loss 两者的优势,在本研究中精度、召回率、置信度得分均达到最高。

表 2 不同损失函数对比

Tab.2 Performance of architectures with different loss functions						
损失函数	验证集	平均 精度/%	精确 率/%	召回 率/%	F1 值/ %	F_c
Focal Loss	Ⅱ	12.9	20.5	69.5	31.7	0.552
	Ⅲ	18.8	26.0	70.3	37.9	0.521
MSE Loss	Ⅱ	99.5	85.7	99.6	92.1	0.785
	Ⅲ	99.9	91.7	100	95.7	0.846
Union Loss	Ⅱ	99.9	91.2	100	95.4	0.892
	Ⅲ	99.4	94.5	99.5	96.9	0.933

2.1.3 不同波段训练对比

为了验证影像不同波段组合对模型精度的影响,以 GPU 环境下使用 ResNet18 作为主干网络进行训练,损失函数为 Union Loss。训练数据为训练集 I 区不同波段组合的随机裁剪得到的结果,使用学习率以及后处理参数完全相同,利用验证数据集对比不同波段组合对于烟草识别的影响。结果显示红边、红、绿波段组合在验证集上精度相对较高(表 3)。在验证集 Ⅱ 中,使用红、绿、蓝波段进行组合,模型将周围杂草等也识别为烟草,故精确率相对

表 3 不同波段组合对比

Tab.3 Performance of ResNet18 models with different combinations of bands						
波段组合	验证集	平均 精度/%	精确 率/%	召回 率/%	F1 值/ %	F_c
红、绿、蓝	Ⅱ	99.5	92.9	99.6	96.2	0.860
	Ⅲ	99.8	95.4	99.8	97.6	0.884
红边、红、绿	Ⅱ	99.6	93.8	99.6	99.6	0.880
	Ⅲ	99.5	94.3	99.7	96.9	0.892
红边、绿、蓝	Ⅱ	98.7	95.3	99.1	97.1	0.758
	Ⅲ	99.3	93.1	99.5	96.2	0.839
红外、红、绿	Ⅱ	99.5	90.8	99.6	95.0	0.850
	Ⅲ	99.5	93.7	99.7	96.6	0.905
红外、红边、绿	Ⅱ	98.5	83.3	99.1	90.5	0.782
	Ⅲ	99.6	91.7	99.7	95.5	0.891
红外、红边、红	Ⅱ	98.4	77.8	98.9	87.1	0.720
	Ⅲ	99.7	91.7	99.8	95.6	0.903

较低。使用红边、红、绿与红边、绿、蓝波段进行组合,在验证集 Ⅱ 与 Ⅲ 上均表现较高精度,但红边、绿、蓝波段组合置信度得分较低,说明该波段组合稳定性不足。在红外波段与其它波段的组合中,当使用红外与红边波段组合时,由于 2 个波段具有相似性,存在信息冗余现象,使模型在相同训练次数下无法迅速捕捉特征,故在验证集 Ⅱ 上呈现相对较低精度,而使用红外、红、绿组合时,波段信息具有独立性,精度更高,说明单独使用红边波段或红外波段引入组合能提高模型训练效率与精度。

通过对比不同波段下典型地物反射光谱变化,发现在波长为 730 nm 时,烟草出现与其它地物相比更高的反射率,有助于进行地物类型区分(图 4),这主要是因为红边波段对植物冠层结构和绿度敏感^[25]。而在波长为 840 nm 时,则表现为烟草与杂草相近的反射率,可区分性较低。在表 3 中,使用红边、红、绿波段与红外、红、绿波段组合时,在验证集 Ⅱ 与 Ⅲ 中,使用红边波段精确率为 94.1%,相较于使用红外波段精确率提高 1.8 个百分点。说明引入红边波段作为训练信息更有助于提高烟草识别精度与泛化性。综上所述,使用红边、红、绿波段组合识别烟草相对于红、绿、蓝波段组合效果更好。

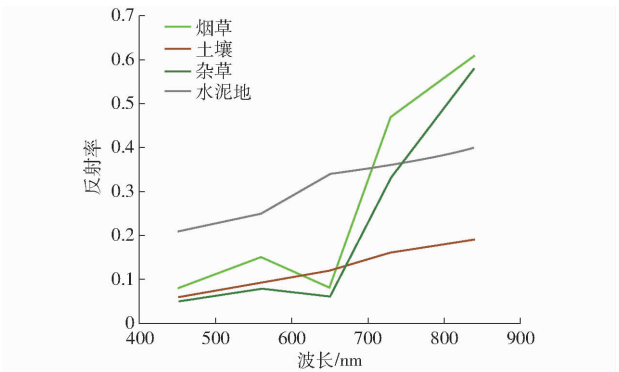


图 4 研究区典型地物的光谱反射率变化曲线
Fig.4 Typical figure reflection spectral variations at different wavelength bands

2.2 烟草识别结果可视化

图 5 展示了不同策略下模型预测概率密度图可视化结果。在不同主干网络可视化对比中,使用 ResNet18 相比其它主干网络具有更高的峰值。在不同损失函数进行对比中,使用 Union Loss 进行烟草中心检测相比其它损失函数更为集中,且峰值更突出。在不同波段组合对比中,使用红边、红、绿波段组合预测烟草中心更为精确,且峰值更高,在后处理局部峰值检测中预测结果更准确。综上所述,使用 ResNet18 作为骨干网络、Union Loss 作为损失函数与红边、红、绿波段组合作为组合策略效果达到最

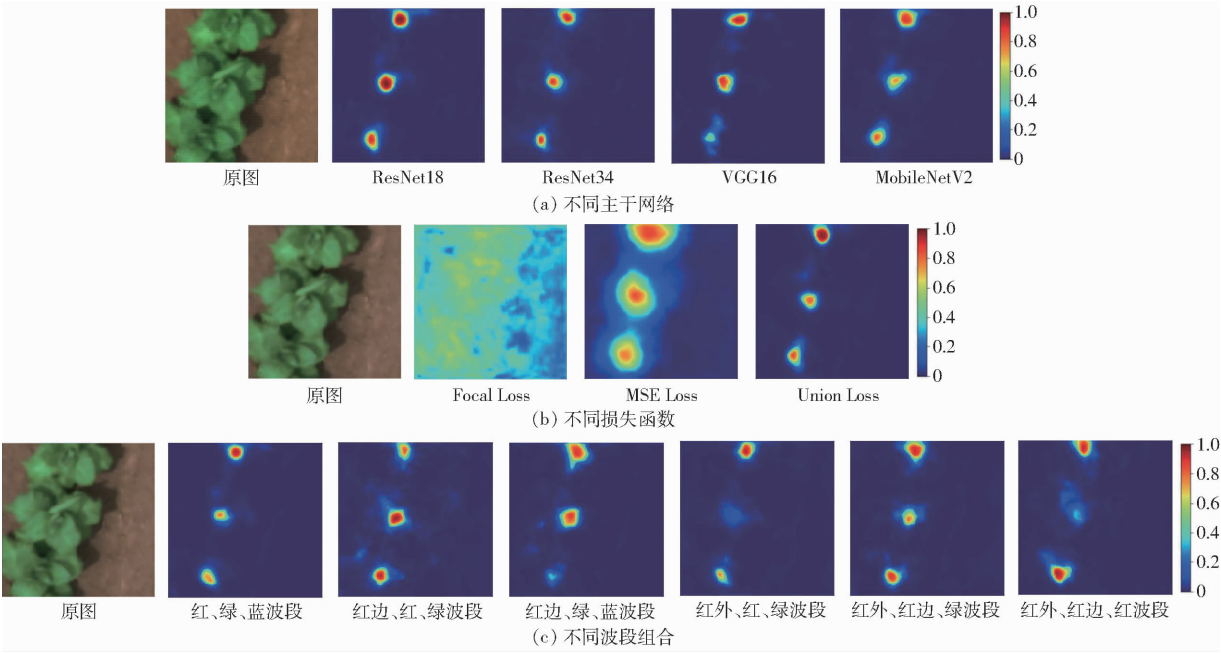


图 5 不同策略预测概率密度图比较

Fig. 5 Comparisons of probability density generated by different models, loss functions, and combinations of bands

佳。如图 6 所示,使用红边、红、绿波段组合模型对整个研究区烟草区域进行预测,通过计数统计得到烟草数量为 13 361 株。

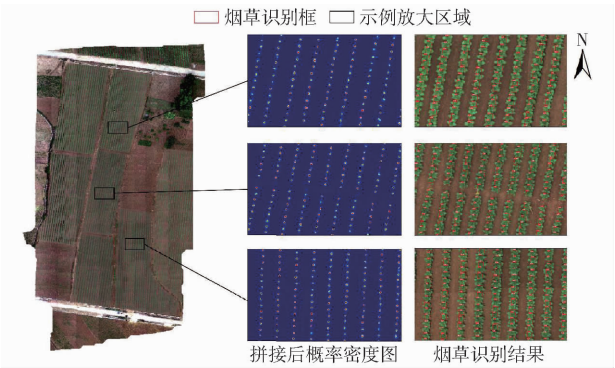


图 6 烟草识别结果

Fig. 6 Results of tobacco recognition

3 结束语

为快速识别烟草并统计株数,利用无人机多光

谱影像,根据烟草形态学特点,提出基于深度学习的轻量级快速识别烟草的模型。与常用的目标检测矩形框标注的方法不同,本研究利用点标注能够减少更多的人力成本,并采用少量数据标注,在少量训练次数的情况下,通过使用不同主干网络对比与轻量化特征融合模块,在训练集与验证集上保持高精度并在 GPU 与 CPU 上都具有较高推理速度。通过对不同主干网络进行对比发现,使用 ResNet18 作为主干网络能在 GPU 环境下取得更高精度(AP 为 99.6% ~ 99.9%),使用 MobileNetV2 作为主干网络在 CPU 上具有更高的运行效率,达 12 f/s。使用本研究提出的 Union Loss 作为损失函数能有效提高识别烟草中心的准确性,并通过红边、红、绿波段组合使用能提高模型泛化性与精确性。本文提出的烟草中心特征识别模型能够直接从无人机多光谱影像中读取研究区的烟草株数,为烟草农情监测提供了有效的技术支持。

参 考 文 献

[1] 童治军,方敦煌,陈学军,等. 6 个烟草重要产量相关性状的遗传分析[J]. 中国烟草学报, 2020, 26(5):72-81. TONG Zhijun, FANG Dunhuang, CHEN Xuejun, et al. Genetic analysis of six important yield-related traits in tobacco (*Nicotiana tabacum* L.)[J]. Chinese Tobacco Journal, 2020, 26(5): 72-81. (in Chinese)

[2] 薛宇飞,张军,张萍,等. 基于 Sentinel-2 遥感影像的烟草种植信息精准提取[J]. 中国烟草科学, 2022, 43(1): 96-106. XUE Yufei, ZHANG Jun, ZHANG Ping, et al. Object-oriented accurate extraction of tobacco information based on Sentinel-2 remote sensing Images[J]. Chinese Tobacco Science, 2022, 43(1): 96-106. (in Chinese)

[3] 杨普,赵远洋,李一鸣,等. 基于多源信息融合的农业空地一体化研究综述[J]. 农业机械学报, 2021, 52(增刊): 185-196. YANG Pu, ZHAO Yuanyang, LI Yiming, et al. Review of research on integration of agricultural air-ground integration based on multi-source information fusion[J]. Transactions of the Chinese Society for Agricultural Machinery, 2021, 52(Supp.): 185-196. (in Chinese)

[4] 董超,赵庚星,宿宝巍,等. 基于无人机多光谱影像的冬小麦返青期变量施氮决策模型研究[J]. 光谱学与光谱分析, 2023, 23(1): 1-8.

- 2019,39(11):3599–3605.
- DONG Chao, ZHAO Gengxing, SU Baowei, et al. Decision model of variable nitrogen fertilizer in winter wheat returning green stage based on UAV multi-spectral images[J]. Spectroscopy and Spectral Analysis, 2019, 39(11):3599–3605. (in Chinese)
- [5] 冯权泷,牛博文,朱德海,等.土地利用/覆被深度学习遥感分类研究综述[J].农业机械学报,2022,53(3):1–17.
- FENG Quanlong, NIU Bowen, ZHU Dehai, et al. Review for deep learning in land use and land cover remote sensing classification[J]. Transactions of the Chinese Society for Agricultural Machinery, 2022, 53(3):1–17. (in Chinese)
- [6] 赵晓伟,黄杨,汪永强,等.基于无人机多光谱数据的玉米苗株估算[J].自然资源遥感,2022,34(1):106–114.
- ZHAO Xiaowei, HUANG Yang, WANG Yongqiang, et al. Estimation of maize seedling number based on UAV multispectral data[J]. Remote Sensing for Natural Resources, 2022, 34(1):106–114. (in Chinese)
- [7] 刘帅兵,杨贵军,周成全,等.基于无人机遥感影像的玉米苗期株数信息提取[J].农业工程学报,2018,34(22):69–77.
- LIU Shuaibing, YANG Guijun, ZHOU Chengquan, et al. Extraction of maize seedling number information based on UAV imagery[J]. Transactions of the CSAE, 2018, 34(22):69–77. (in Chinese)
- [8] JIN Xiuliang, LIU Shouyang, BARET Frédéric, et al. Estimates of plant density of wheat crops at emergence from very low altitude UAV imagery[J]. Remote Sensing of Environment, 2017, 198:105–114.
- [9] 郑晓岚,张显峰,程俊毅,等.利用无人机多光谱影像数据构建棉苗株数估算模型[J].中国图象图形学报,2020,25(3):520–534.
- ZHENG Xiaolan, ZHANG Xianfeng, CHENG Junyi, et al. Using the multispectral image data acquired by unmanned aerial vehicle to build an estimation model of the number of seedling stage cotton plants[J]. Journal of Image and Graphics, 2020, 25(3):520–534. (in Chinese)
- [10] 戴建国,薛金利,赵庆展,等.利用无人机可见光遥感影像提取棉花苗情信息[J].农业工程学报,2020,36(4):63–71.
- DAI Jianguo, XUE Jinli, ZHAO Qingzhan, et al. Extraction of cotton seedling growth information using UAV visible light remote sensing images[J]. Transactions of the CSAE, 2020, 36(4):63–71. (in Chinese)
- [11] LU H, CAO Z, XIAO Y, et al. TasselNet: counting maize tassels in the wild via local counts regression network[J]. Plant Methods, 2017, 13(1):1–17.
- [12] REDMON J, FARHADI A. YOLOv3: an incremental improvement[EB/OL]. (2018–04–08). <https://arxiv.org/pdf/1804.02767.pdf>.
- [13] 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.
- [14] LONG J, SHELHAMER E, DARRELL T. Fully convolutional networks for semantic segmentation[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2015, 39(4):640–651.
- [15] RONNEBERGER O, FISCHER P, BROX T. U-Net: convolutional networks for biomedical image segmentation[C]//Medical Image Computing and Computer-Assisted Intervention-MICCAI 2015: 18th International Conference, Munich, Germany, 2015:234–241.
- [16] ZHANG Y, ZHOU D, CHEN S, et al. Single-image crowd counting via multi-column convolutional neural network[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2016:589–597.
- [17] TOSHEY A, SZEGEDY C. DeepPose: human pose estimation via deep neural networks[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Columbus, OH, USA. 2014:24–27.
- [18] TOMpson J, JAIN A, LECUN Y, et al. Joint training of a convolutional network and a graphical model for human pose estimation[J]. Advances in Neural Information Processing Systems, 2014,27:1799–1807.
- [19] CUBUK E D, ZOPH B, MANE D, et al. Autoaugment: learning augmentation strategies from data[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2019:113–123.
- [20] LAW H, DENG J. CornerNet: detecting objects as paired keypoints[J]. International Journal of Computer Vision, 2020, 128(3):642–656.
- [21] SIMONYAN K, ZISSERMAN A. Very deep convolutional networks for large-scale image recognition[J]. arXiv preprint arXiv:1409.1556, 2014.
- [22] HE K, ZHANG X, REN S, et al. Deep residual learning for image recognition[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2016:770–778.
- [23] SANDLER M, HOWARD A, ZHU M, et al. Mobilenetv2: inverted residuals and linear bottlenecks[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2018:4510–4520.
- [24] LIN T Y, GOYAL P, GIRSHICK R, et al. Focal loss for dense object detection[C]//Proceedings of the IEEE International Conference on Computer Vision, 2017:2980–2988.
- [25] MONNO Y, TERANAKA H, YOSHIKAZI K, et al. Single-sensor RGB–NIR imaging: high-quality system design and prototype implementation[J]. IEEE Sensors Journal, 2018, 19(2):497–507.