

地力夏提·依马木, 周建平, 许燕, 等. 基于 Logistic 算法与遥感影像的棉花虫害监测研究 [J]. 华南农业大学学报, 2022, 43(2): 87-95.
DILIXIATI Yimamu, ZHOU Jianping, XU Yan, et al. Cotton pest monitoring based on Logistic algorithm and remote sensing image[J]. Journal of South China Agricultural University, 2022, 43(2): 87-95.

基于 Logistic 算法与遥感影像的棉花虫害监测研究

地力夏提·依马木¹, 周建平^{1,2}, 许燕¹, 樊湘鹏¹, 亚里坤·沙吾提³

(1 新疆大学 机械工程学院, 新疆 乌鲁木齐 830047; 2 新疆大学 新疆维吾尔自治区农牧机器人及智能装备工程研究中心, 新疆 乌鲁木齐 830047; 3 新疆维吾尔自治区农牧业机械化技术推广总站, 新疆 乌鲁木齐 830054)

摘要:【目的】借助多光谱遥感影像和 Logistic 算法, 实现对棉田虫害的田间监测。【方法】以虫害棉花区域为研究对象, 利用无人机获取棉田多光谱遥感影像, 并对影像进行预处理; 结合受虫害棉花光谱特征, 利用虫害敏感波段反射率与植被指数构建 Logistic 回归模型, 开展棉花虫害识别监测研究。【结果】由土壤调节植被指数 (Soil adjusted vegetation index, SAVI) 模型和归一化植被指数 (Normalized vegetation index, NDVI) 模型构建的棉蚜虫、棉红蜘蛛、棉铃虫识别模型为最优模型, 其训练样本准确率达到 93.7%, 测试样本准确率达到 90.5%, 召回率为 96.6%, $F1$ 值为 93.5%, 对棉蚜虫、棉红蜘蛛和棉铃虫的识别模型决定系数分别为 0.942、0.851 和 0.663。【结论】该模型可满足棉田中棉蚜虫、棉红蜘蛛和棉铃虫 3 种虫害的发生区域识别, 且可基本满足棉田精准植保作业相关要求。

关键词: 无人机; 光谱特征; 遥感影像; 植被模型; Logistic 回归模型算法; 虫害监测

中图分类号: S251; TP751.1

文献标志码: A

文章编号: 1001-411X(2022)02-0087-09

Cotton pest monitoring based on Logistic algorithm and remote sensing image

DILIXIATI Yimamu¹, ZHOU Jianping^{1,2}, XU Yan¹, FAN Xiangpeng¹, YALIKUN Shawuti³

(1 College of Mechanical Engineering, Xinjiang University, Urumqi 830047, China; 2 Agriculture and Animal Husbandry Robot and Intelligent Equipment Engineering Technology Center of Xinjiang Uygur Autonomous Region, Xinjiang University, Urumqi 830047, China; 3 Xinjiang Uygur Autonomous Region Agricultural and Animal Husbandry Mechanization Technology Promotion Terminal, Urumqi 830054, China)

Abstract: 【Objective】The purpose of this article is to monitor cotton pest in field based on Logistic algorithms and multi-spectral remote sensing images. 【Method】The cotton areas with insect pests were selected as the research object. The multi-spectral remote sensing images of cotton field were acquired by UAV, and then pre-processed. Based on the spectral characteristics of cotton pests, the Logistic regression model was constructed by the reflectivity of pest-sensitive band and vegetation index to identify and monitor cotton pests. 【Result】The cotton aphid, cotton red spider mite, and cotton bollworm identification models constructed by

收稿日期: 2021-06-03 网络首发时间: 2022-01-06 17:45:00

网络首发地址: <https://kns.cnki.net/kcms/detail/44.1110.S.20220106.1232.004.html>

作者简介: 地力夏提·依马木, 硕士, 主要从事水肥一体化数字滴灌系统研究, E-mail: 965001772@qq.com; 通信作者: 周建平, 教授, 博士, 主要从事水肥一体化数字滴灌系统研究, E-mail: linkzhou@163.com

基金项目: 国家自然科学基金 (51765063); 国家级大学生创新创业训练计划项目 (201810755079S); 新疆维吾尔自治区研究生科研创新项目 (XJ2019G033)

the soil adjusted vegetation index (SAVI) model and the normalized vegetation index (NDVI) model were the optimal models, and their accuracy for training sample and test sample reached 93.7% and 90.5% respectively the recall rate and *F1* value were 96.6% and 93.5% respectively and the determination coefficients of recognition models for three types of pests were 0.942, 0.851 and 0.663 respectively. 【Conclusion】 This model can identify the occurrence area of cotton aphid, cotton red spider mite and cotton bollworm, which can basically meet the requirements of precision plant protection operation in cotton field.

Key words: UAV; Spectral characteristic; Remote sensing image; Vegetation model; Logistic regression model algorithm; Pest monitoring

近年来,提升大田作物品质与保障农田作物产量是农业机械化领域的热点研究问题,给农业生产造成损失的主要原因是虫害的发生与蔓延。据现有统计信息,2020 年中国棉花总产量为 591 万吨,因病虫害损失 15%~20%^[1]。作为中国最大的优质棉生产区域,新疆棉花播种面积占全国的 76% 左右,总产占全国棉花产能的 84.9%^[2-3]。棉花高产量与优产率得益于新疆独特的地域优势与气候环境优势^[4-5]。棉蚜虫、棉红蜘蛛、棉铃虫是棉花的主要害虫,对棉花生产造成的危害较大。精准、快速地对棉花虫害发生情况进行监测,及时了解棉花虫害动态变化,对于棉花大面积稳健生长具有重要意义^[6-7]。传统的虫害识别监测、定位方法需要耗费大量的人力、物力,仅适用于小片区棉田,难以满足新疆棉田规模化植保防治的要求^[8]。赵亮等^[9]基于便携式光谱仪的光谱信息分析螨害严重度与棉叶螨叶片光谱反射率的相关性,构建了基于敏感波段的虫害诊断模型,从而实现了棉花螨害的识别提取。Prabhakar 等^[10]利用叶片波谱分析和受害叶片症状解读方法,构建叶蝉敏感波段,并利用比值指数组合构建叶蝉指数 (Leafhopper index, LHI) 有效识别叶蝉发生区域;竞霞等^[11]利用卫星遥感数据提取归一化差值植被指数 (Normalized difference vegetation index, NDVI)、重归一化差值植被指数 (Renormalized difference vegetation index, RDVI)、增强植被指数 (Enhanced vegetation index, EVI)、差值植被指数 (Difference vegetation index, DVI)、全球环境监测植被指数 (Global environment monitoring index, GEMI) 和修改型土壤植被指数 (Modified soil-adjusted vegetation index, MSAVI) 共 6 种影响因子,使用偏最小二乘法 (Partial least squares, PLS) 和变量投影重要性 (Variable importance projection, VIP) 准则构建棉花黄萎症影响度估测模型,有效识别了病害发生区

域并分析了其精度;胡根生等^[12]利用植被中对蚜虫影响较大的植被指数生成 8 种特征因子,基于环保卫星远程遥感影像资料,提取相应光谱特征信息,利用最小二乘孪生支持向量机建立小麦蚜虫识别模型,且模型精度达到 86.4%,较好地实现了对小麦蚜虫的监测。Stella 等^[13]基于农作物受病虫害侵害的时间构建作物病虫害的 Logistic 模型,以总农作物种群为常数,以病虫害媒介种群为变量,以基本病虫繁殖数为平衡点分析病虫害发生情况,并在棉花、小麦、玉米等农作物上进行模型验证;冯炼等^[14]在冬小麦病虫害面积监测研究中选择二值 Logistic 回归法,分别建立植被指数与病虫害发生区域的 Logistic 模型,从模型识别结果来看,提取的冬小麦受病虫害胁迫面积与实测值结果相符,能满足农作物病虫害发生区域遥感监测要求,该模型在棉田病虫害识别方面也得到了较好的精度。学者们在预测虫害发生趋势研究方面利用 Logistic 模型确定性和随机性、虫害敏感波段、植被指数影响因子等方法,对比基于无人机遥感获取的高光谱、多光谱传感器影像数据,识别提取作物虫害发生区域并预测虫害发生趋势。结合无人机光谱遥感精准、高时效、高分辨率的优势,可以实现虫害区域预测研究^[9, 11, 15-17]。而 Logistic 模型可同时引用敏感波段、植被指数等数据^[14, 18]构建作物虫害识别模型,且从精度数据来看,对作物虫害识别的准确率均可满足虫害发生区域监测要求,故本研究采用 Logistic 模型实现棉花虫害监测与分类。现阶段无人机遥感影像数据在识别监测棉田虫害研究领域仅停留在对单一虫害的识别监测,鲜见对多种虫害的识别监测,且鲜见对影响遥感影像精度的因素及其解决方案的系统性研究。借助遥感影像处理软件分析影响遥感影像精度的因素并对所获取棉田遥感影像进行几何校正、辐射校正,提升影像精度,其次利用 Logistic 模

型、自相关性时间序列检测的互相关函数 (Cross correlation function, CCF) 等算法, 实现棉田多种虫害快速精准识别监测, 以期为棉花稳产高质提供技术服务, 为虫害识别提供数据支撑。

1 研究平台与方法

1.1 研究区概况

研究区位于新疆维吾尔自治区巴音郭楞蒙古自治州尉犁县境内, 地处北纬 $41^{\circ}29'27''\sim 41^{\circ}29'58''$, 东经 $86^{\circ}15'30''\sim 86^{\circ}16'3''$, 占地 127.3 hm^2 。属暖温带大陆性荒漠气候, 冷热差异悬殊, 全年平均气温 $10.1\text{ }^{\circ}\text{C}$, 年平均降水量 43 mm , 年平均蒸发量 2700 mm 。本次研究选择 ‘新陆中 68 号’ 棉花作为研究对象, 生育期 $134\sim 143\text{ d}$, 研究区棉花采取 2 膜 12 行种植模式, 行距配置为 $(10+66+10+66+10)\text{ cm}$, 种植行间距为宽行 66 cm , 窄行 10 cm , 株距 10 cm 。种植区域采用滴灌模式, 其他管理措施按照当地大田棉花高产栽培模式进行。

1.2 试验设计

棉田遥感影像数据获取试验于 2020 年 7 月 15 日早间进行, 天气晴, 气温 $22\text{ }^{\circ}\text{C}$, 风向西北, 风速 3.1 m/s , 适宜无人机航拍作业。无人机遥感平台选用 DJI 大疆经纬 Matrice 600 多旋翼无人机, 搭载 MicaSense RedEdge-M 5 通道光谱相机构成多旋翼无人机遥感系统。RedEdge-M 5 通道光谱相机分辨率为 $1\ 280\times 960$, 光谱波段包括蓝光 475 nm 、绿光 560 nm 、红光 668 nm 、近红外 840 nm 和红光 717 nm , 其对应的带宽分别为 20 、 20 、 10 、 40 和 10 nm 。

数据采集时, 根据预定航线采取垂直拍摄模式, 覆盖整片研究区地块, 飞行高度 50 m , 影像航向重叠率 80% , 旁向重叠率 70% , 满足影像处理需求。从研究区棉田划分 4.67 hm^2 棉田作为验证区, 取 20 个采样节点, 每个采样节点为 $3\text{ m}\times 3\text{ m}$ 空间范围, 采样点地理位置信息采取空地设备结合方式, 地面选取轻测 Lite RTK 设备获取地理位置信息, 该设备满足 5 Hz 刷新率, RTK 精度达到 $\pm(10+1\times 10^{-6}D)\text{ mm}$, 其中 D 为被测点间距; 选用本研究执飞无人机 GPS 模块获取空间地理位置信息。获取的地理位置数据经由几何校正后, 生成试验区正射影像图, 从而保证采样点定位精度。选取此空间范围内所有棉花植株作为采样对象, 采样对象包含棉花植株叶片、花蕾、花铃所发生的虫害, 采样节点位置示意图如图 1 所示。采样过程样叶分成左基部、右基部两端, 每端各测 2 次, 每次测 2 条光谱, 生成 4 组光谱波长数据, 取其平均值作为样叶光谱值, 棉

花叶片采样位置如图 2 所示。在 20 个采样节点中选取 300 株棉花验证模型准确度, 从中选取总样本数 $2/3$ 的棉花植株 (200 株) 用于模型训练, 其中, 100 个样本选取受棉蚜虫、棉红蜘蛛、棉铃虫虫害的棉花植株, 100 个样本为健康棉花植株; 剩余 $1/3$ 的棉花植株样本数据用来测试精准度, 其中, 50 个样本为受棉蚜虫、棉红蜘蛛、棉铃虫虫害的棉花植株, 50 个样本为健康棉花植株。本研究中健康棉叶与受虫害棉叶区分标准为: 受虫害棉叶指叶片受虫害侵蚀区域面积大于单个叶片面积的 5% , 健康棉叶指叶片受虫害侵蚀区域面积少于单个叶片面积的 5% 。

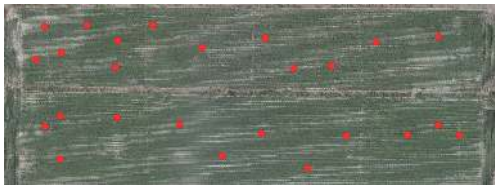


图 1 采样节点位置示意图
Fig. 1 Schematic diagram of sampling node location

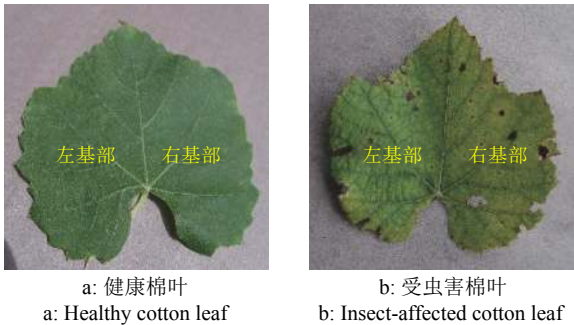


图 2 棉花叶片采样位置示意图
Fig. 2 Schematic diagram of sampling position of cotton leaves

1.3 研究方法

针对现阶段无人机遥感在规模化农田虫害识别研究中存在的遥感影像质量严重依赖天气、图片质量影响虫害识别精度及当前研究停留于单一虫害识别等问题, 本研究利用几何校正、辐射校正方法, 提升图像地理位置精度与像素覆盖密度, 有效提升遥感影像质量; 利用逻辑回归 (Logistic regression, LR) 方法识别虫害发生区域, 对虫害识别精度及虫害发生区域特征进行分析。

LR 又称为广义线性回归, 属于判别分析, 用于处理因变量为分类变量的数据。在实际应用过程中响应变量为二值变量, 为解决二分类问题, 延伸出二值逻辑回归法^[19]。在解决二分类问题过程中, 因变量为二分类变量, 利用 0 与 1 代替表达不同状态。本研究中棉蚜虫、棉红蜘蛛、棉铃虫 3 种虫害发

生与否为非线性变化求解问题, 因此, 以上 3 种虫害监测问题属于非线性二分类问题, 基于此建立 Logistic 回归判定模型:

$$P = \frac{\exp(\beta_0 + \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_p x_p)}{1 + \exp(\beta_0 + \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_p x_p)}, \quad (1)$$

式中, β_0 为常数项或截距, $\beta_1, \beta_2, \cdots, \beta_p$ 为模型回归系数, $x_1, x_2, \cdots, x_p$ 代表 p 个自变量, P 表示在 p 个自变量作用下的概率。也可表示为在影响 Y 取值范围的 p 个自变量条件作用下棉蚜虫、棉红蜘蛛、棉铃虫发生的概率为 $P=P(Y=1|x_1, x_2, \cdots, x_p)$ 。

对式 (1) 进行 Logistic 变换, 其线性表达式如下:

$$Y = \text{Logit}(P) = \ln\left(\frac{P}{1+P}\right) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_p x_p. \quad (2)$$

从式 (2) 可以看出, 回归系数 β 在控制自变量条件下反映变量影响力, 响应变量 Y 求解二分类问题时分别取值 0 和 1, 表示发生与不发生, P 为发生概率, 以 $P=0.5$ 作为界限, $P \geq 0.5$ 表示逻辑为真, 即棉花植株发生虫害, 当 $P < 0.5$ 表示棉花植株健康 (未发生虫害)。

2 结果与分析

2.1 遥感影像预处理

影像拼接及几何校正过程使用瑞士 Pix4D Mapper 软件进行, 可实现空三加密, 结合控制点信息与所采集图像地理位置信息数据进行几何校正, 生成研究区正射影像^[20-22], 几何校正过程利用归一化函数提高配准点精度, 从而有效提高图像精度。

经几何校正后, 有效提升图像最优模型精度, 模型精度达到 (0.059 ± 0.012) m。所生成正射影像如图 3 所示, 几何校正后绝对位置和方向不确定性参数如表 1 所示。

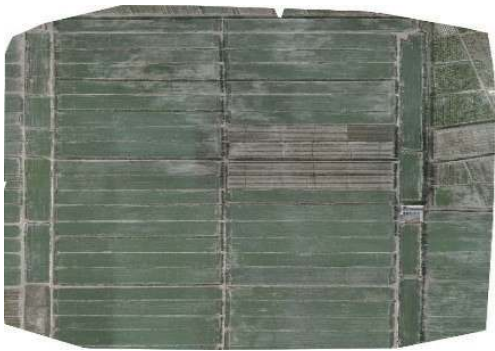


图 3 研究区正射影像图
Fig. 3 Orthophoto map of the study area

表 1 绝对位置和方向不确定性参数
Table 1 Absolute geographic location and directional uncertainty parameters

指标 Item	X/m	Y/m	Z/m	偏转角/(°) Deflection angle	平移角/(°) Translation angle	扭转角/(°) Twist angle
均值 Average value	0.134	0.133	0.319	0.079	0.068	0.015
离散值(σ) Discrete value	0.022	0.026	0.067	0.009	0.014	0.002

由于受光照强度及其他因素影响, 遥感影像光谱响应误差表现为像元值像素覆盖密度数值波动, 造成图像解析不准确, 无法精准有效反映光谱反射率。为消除数值波动造成的像素密度误差, 本研究利用 Back propagation(BP) 神经网络模型对棉田遥感影像数据进行监督分类。辐射校正过程使用 ENVI 软件 Band Math 工具定义辐射定标公式:

$$L_{\lambda} = \left(\frac{LMAX_{\lambda} - LMIN_{\lambda}}{QCALMAX} \right) QCAL + LMIN_{\lambda}, \quad (3)$$

式中, L_{λ} 为单元像元 (Digital number, DN) 内像素覆盖密度, $LMAX_{\lambda}$ 为单元像元最大值光谱反射率, $LMIN_{\lambda}$ 为单元像元最小值光谱反射率, $QCAL$ 为单元像元值, $QCALMAX$ 是最大量化定标像素值。红光波段、绿光波段、蓝光波段、近红外波段辐射校正前单位像元内覆盖密度分析结果如图 4a 所示。

对所获取无人机遥感棉田影像提取的红光波段 (Red)、绿光波段 (Green)、蓝光波段 (Blue)、近红外波段 (Near-infrared, Nir) 组合作为输入层构建 Red-BP、Green-BP、Blue-BP、Nir-BP 多层感知器模型优化波段反射率, 光谱波段反射率作为输出层。以辐射校正前影像像素反射率数值作为预测值, 该预测值数据由 Micro-Hyperspec 小型光谱检测仪实测叶片波长数据计算得出, 辐射校正后 BP 模型反射率数值作为实测值, 选用 Sigmoid 函数作为激活函数, 实数值自变量区间为 $[0, 1]$, 采用梯度下降法训练数据集。经辐射校正后红光波段、绿光波段、蓝光波段、近红外波段单位像元内覆盖密度结果如图 4b 所示。由辐射校正模型预测值-实测值散点图 (图 5) 可知, 红光波段、蓝光波段、近红外波段和绿光波段基于 BP 法的 Red-BP、Blue-BP、Nir-

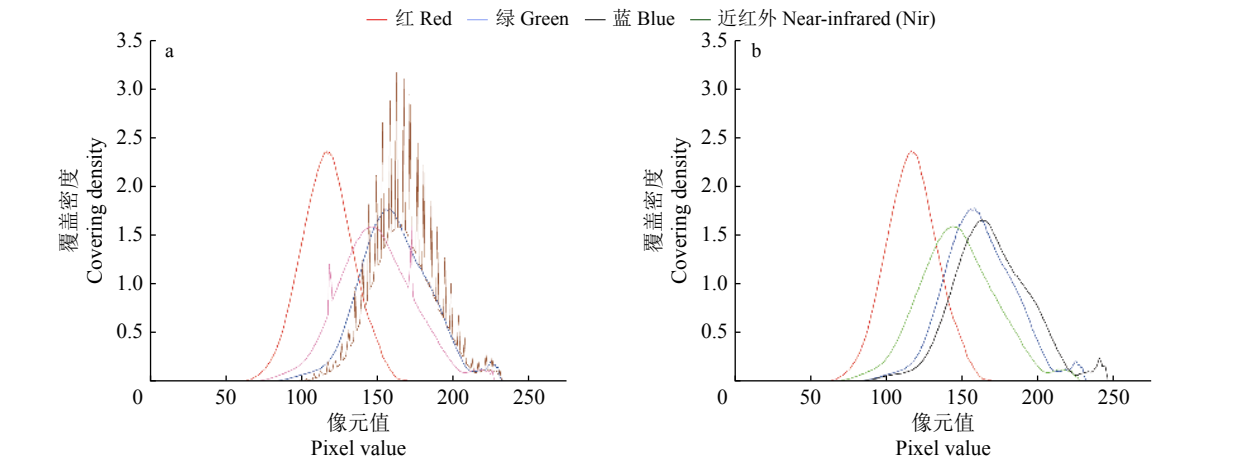


图 4 辐射校正前 (a)、后 (b) 单位像元内像素覆盖密度

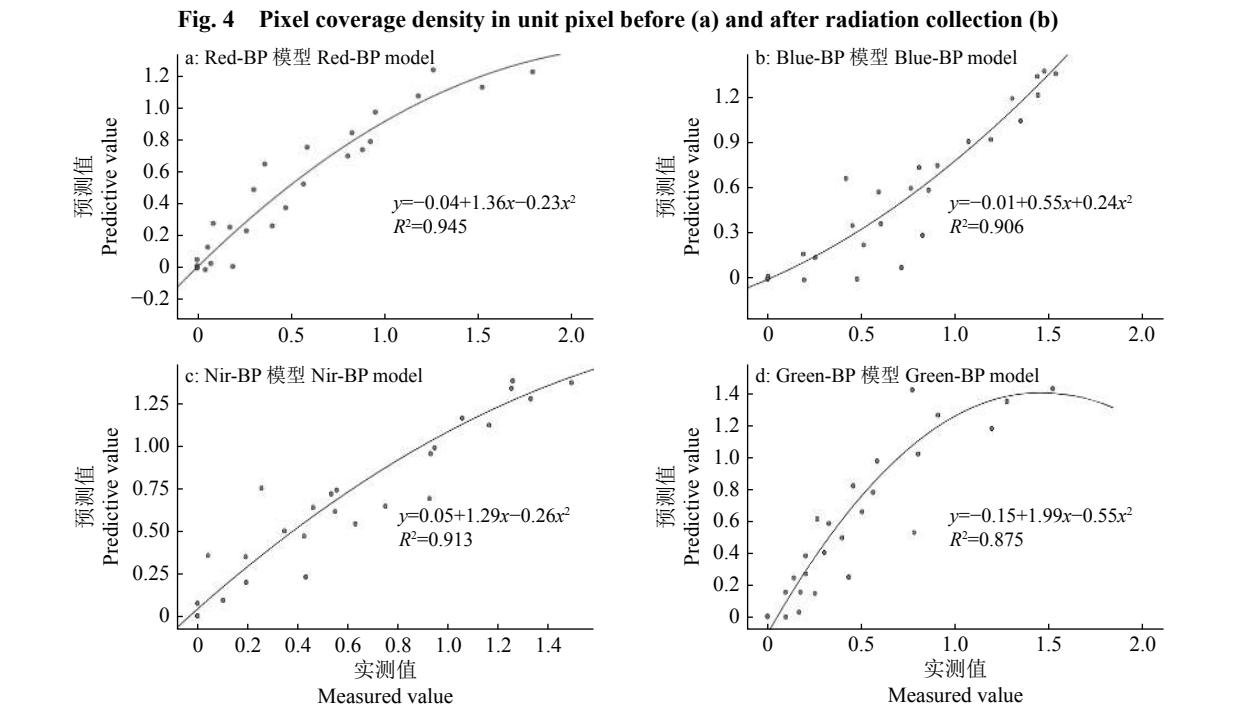


图 5 BP 神经网络模型对辐射校正模型光谱反射率估测结果

Fig. 5 Estimation results of spectral reflectance of image pixel after radiation correction by BP neural network model

BP 和 Green-BP 模型光谱反射率预测值与实测值分布较为均匀, 其预测值与实测值有较好的拟合效果, 且决定系数 (R^2) 分别为 0.945、0.906、0.913 和 0.875, 模型精度较高, 满足图像解析要求。

根据图 4 提取辐射校正前、后单元 DN 内像素覆盖密度数据, 再针对像素覆盖密度数据绘制 ROC 曲线 (图 6), 对比数据预测价值比较如表 2 所示。

辐射校正后单元 DN 内红光波段、蓝光波段、近红外波段和绿光波段像素覆盖密度的 ROC 曲线下面积 (The area under the ROC curve, AUC) 分别为 0.875、0.750、0.688 和 0.875; 在最佳点位置, 像素覆盖密度灵敏度分别为 88.3%、79.3%、80.9% 和 87.1%。从灵敏度、ROC 曲线下面积及校正情况

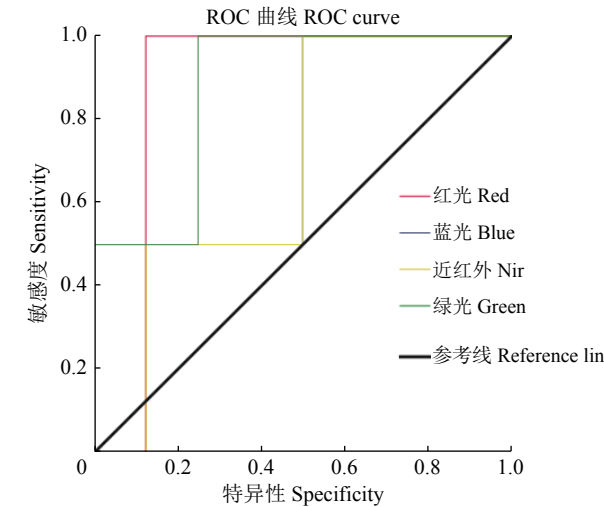


图 6 4 种波段像素覆盖密度

Fig. 6 ROC curves of pixel coverage densities of four bands

表 2 4 种波段像素覆盖密度预测价值比较

Table 2 Comparison of the prediction values for the pixel coverage densities of four bands

辐射校正波段 Radiation correction band	最佳截点 Best cut point	举例数 No. of cases	校正情况例数 No.of correction situation		灵敏度/% Sensitivity	ROC曲线下面积 The area under the ROC curve
			好 Good	差 Bad		
红光 Red	<0.1	25	22	3	88.3	0.875
	≥0.1	225	199	26		
蓝光 Blue	0	125	100	25	79.3	0.750
	≥0	125	100	25		
近红外 Nir	<0.5	13	11	2	80.9	0.688
	≥0.5	237	192	45		
绿光 Green	<0.5	63	55	8	87.1	0.875
	≥0.5	187	163	24		

来看, 辐射校正可有效降低光谱响应误差, 增强影像像素密度。

2.2 指数模型构建

为更好地反映出大田棉花种植区域、棉花长势情况及棉叶光谱反射率基本情况, 为 Logistic 模型提供数据, 本研究构建 NDVI、EVI、SAVI 模型。NDVI 是体现农作物植被覆盖度的参数, 是描述植被生长状态的一个重要指标^[23-25], 对于本研究识别棉花虫害区域有着重要的意义。为增强植被反射率信息, EVI 在归一化指数基础上加入蓝光波段数据, 矫正土壤参数和气溶胶散射数据的影响, EVI 常用于植被覆盖密度较大区域^[24, 26]。SAVI 是基于 NDVI 和大量遥感数据观测修正后提出用以减小土壤背景影响的指数模型。利用 ENVI 软件 Spectral Math 工具构建指数模型, 包括 NDVI 模型、EVI 模型和 SAVI 模型, 3 种植被指数模型公式如下:

$$NDVI = \frac{Nir - Red}{Nir + Red}, \tag{4}$$

$$EVI = \frac{2.5(Nir - Red)}{Nir + 6Red - 7.8Blue + 1}, \tag{5}$$

$$SAVI = \frac{(Nir - Red)(1 + L)}{Nir + Red + L}, \tag{6}$$

式中, L 是随着植被覆盖度变化的参数, 取值阈值为 0~1, 当植被密度很高时为 0, 反之为 1。很明显, 如果 $L=0$, $SAVI=NDVI$ 。对于棉花植株, L 取 0.5 时 SAVI 消除土壤背景影响的效果较好^[27]。

NDVI 模型可直观地表述棉花与土壤分布情况, EVI 模型可直观地表述棉花生长趋势情况,

SAVI 模型可直观地表述土壤分布及棉花生长趋势; 3 种模型均可精准提取棉花长势信息及分布面积, 在棉田虫害模型构建过程中提供棉花分布情况数据及土壤边界。

2.3 棉花虫害敏感波段提取

使用 Micro-Hyperspec 小型光谱检测仪对采样节点受虫害区域及健康棉花区域进行光谱值测定, 利用 Excel 软件对样叶光谱波长数据进行统计分析, 利用反射率公式生成虫害反射率及健康棉花叶片波长数据, 为 Logistic 模型提供敏感波段数据公式如下:

$$L = \frac{(LMAX - LMIN)(QCAL - QCALMIN)}{QCALMAX - QCALMIN} + LMIN, \tag{7}$$

式中, LMAX、LMIN 分别为像元辐射亮度最大值和最小值; QCAL 为某一像元 DN 值; QCALMAX、QCALMIN 分别为像元所取最大值和最小值。

使用 Origin 软件构建棉田虫害光谱敏感波段模型及健康棉叶光谱波长信息模型, 由图 7 可以看出, 棉蚜虫的严重程度与 713~831 nm 波段的棉花蚜虫叶片光谱反射率显著相关; 棉红蜘蛛的严重程度与 605~724 nm 波段的棉花红蜘蛛叶片光谱反射率显著相关; 棉铃虫的严重程度与 709~794 nm 波段的棉花铃虫叶片光谱反射率显著相关, 此范围内遥感因子可以作为辨认棉花棉蚜虫、棉红蜘蛛、棉铃虫 3 种虫害的敏感波段。

2.4 Logistic 算法虫害识别研究

2.4.1 基于 Logistic 算法虫害识别模型 模型构建前, 为使构建模型特征相同, 对特征变量进行归一化分析。基于此, 利用 SPSS 软件分析 300 株棉花数据, 将本研究中应用的 3 种植被指数作为回归模

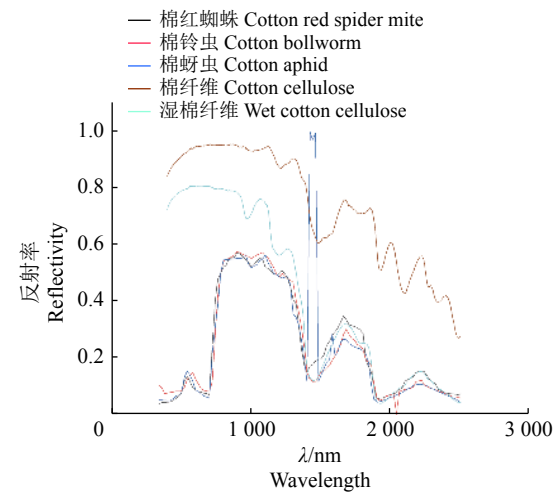


图 7 健康棉叶及棉花虫害光谱波长反射率曲线

Fig. 7 Spectral wavelength reflectance curves of healthy cotton leaves and cotton pests

型输入变量因子, 构建虫害识别 Logistic 回归模型, 同时计算均方根误差 (Root mean square error, RMSE)。当引入 1 个输入变量因子时, 回归模型为 $y=10.25+9.4SAVI$, RMSE 为 0.193; 输入 2 个变量因子时, 回归模型为 $y=-13.16+155.1SAVI-24.6NDVI$, RMSE 为 0.187; 输入 3 个变量因子时, 回归模型为 $y=-9.85+138.38SAVI-75.7NDVI-44.8EVI$, RMSE 为 0.376。NDVI、EVI、SAVI 与棉花受害情

况的相关系数分别为-0.914, -0.813 和 0.937, 分别达到 0.01、0.05 和 0.01 水平。综上, 构建虫害识别模型选取 2 个变量因子的识别效果最佳, 由 SAVI 模型和 NDVI 模型建立的 Logistic 回归模型是识别棉蚜虫、棉红蜘蛛、棉铃虫的最优模型。

利用 ENVI 软件基于规则面向对象工作流程工具, 波谱类型选取 SAVI 模型、NDVI 模型与棉蚜虫、棉红蜘蛛、棉铃虫 3 种虫害光谱反射率曲线, 基于 Logistic 回归判定的虫害识别模型, 生成虫害识别区域如图 8 所示, 由图 8 可明显看出受虫害影响区域。

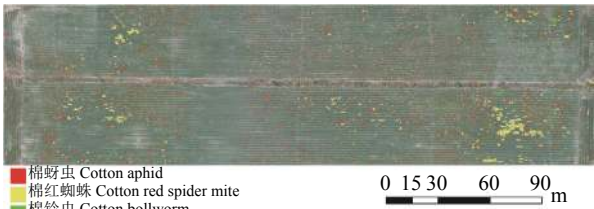


图 8 3 种害虫为害区域识别结果

Fig. 8 Identification results of three pests infested arer

2.4.2 模型分类结果 基于 RMSE 选择具有代表性的 Logistic 回归虫害识别模型进行评价比较, 具体分类结果分为选择 1、2、3 个输入变量因子, 模型分类比较结果如表 3 所示。从比较结果中可以看出, 输入变量为 2 的虫害识别模型识别效果最

表 3 模型分类比较结果¹⁾

Table 3 Comparison results of model classification

模型 Model	实际样本 Actual sample	训练样本 Training sample						测试样本 Test sample						召回率/% Recall rate	F1
		健康 Health	B ₁	B ₂	B ₃	总和 Sum	准确率/% Accuracy	健康 Health	B ₁	B ₂	B ₃	总和 Sum	准确率/% Accuracy		
RMSE=0.193 n=1	健康Health	50	3	1	1	55	91.4	28	1	1	2	32	85.8	93.9	89.8
	B ₁	3	46	2	1	52		2	24	1	1	28			
	B ₂	2	0	46	0	48		1	1	18	1	21			
	B ₃	3	2	0	40	45		2	0	1	16	19			
	总和Sum	58	51	49	42	200		33	26	21	20	100			
RMSE=0.187 n=2	健康Health	50	1	1	1	53	93.7	25	1	0	0	26	90.5	96.6	93.5
	B ₁	3	53	1	0	57		2	24	1	1	28			
	B ₂	3	0	45	0	48		2	1	29	1	33			
	B ₃	1	0	1	40	42		0	1	0	12	13			
	总和Sum	57	54	48	41	200		29	27	30	14	100			
RMSE=0.376 n=3	健康Health	67	6	5	2	80	83.1	22	3	3	2	30	75.1	90.4	82.8
	B ₁	4	48	5	3	60		2	16	1	2	21			
	B ₂	3	1	25	1	30		2	2	20	2	26			
	B ₃	2	0	2	26	30		3	1	2	17	23			
	总和Sum	76	55	37	32	200		29	22	26	23	100			

1) B₁、B₂和B₃ 分别表示棉蚜虫、棉红蜘蛛和棉铃虫
1) B₁, B₂ and B₃ indicate cotton aphid, cotton red spider mite and cotton bollworm respectively

佳, 而输入变量为 3 的虫害识别模型在准确率、召回率及 F1 值上均低于其他 2 个虫害识别模型。结合模型构建过程, 由 SAVI 模型和 NDVI 模型建立的棉蚜虫、棉红蜘蛛、棉铃虫识别模型为最优模型, 其训练样本准确率和测试样本准确率分别达到 93.7% 和 90.5%, 召回率、F1 值分别为 96.6% 和 93.5%。分别对 3 种虫害发生区域像元数据参数估算,

分析基于 Logistic 模型识别的棉田虫害区域预测值-实测值数据, 验证 Logistic 模型在棉田虫害区域监测的精准度。3 种虫害预测值-实测值相关性分析如图 9 所示。由图 9 可知, Logistic 模型在棉田虫害区域监测的精准度较高, 对棉蚜虫、棉红蜘蛛、棉铃虫发生区域识别模型的决定系数分别为 0.942、0.851 和 0.663。

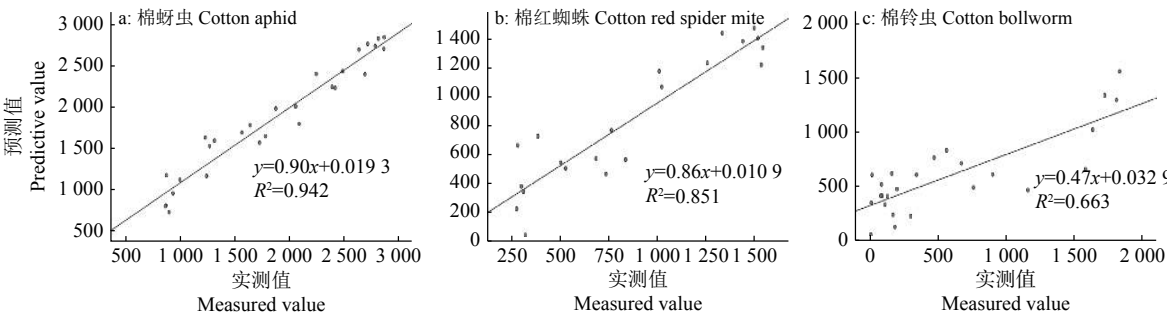


图 9 3 种虫害发生区域识别精准度

Fig. 9 Accuracy of the occurrence area of three types of pests

2.4.3 虫害发生区域特征分析 对棉田受棉蚜虫、棉红蜘蛛、棉铃虫灾害区域进行自相关性时间序列检测 (CCF), 分析随时间变化的棉田受虫害面积的相关性, 结果见图 10。由图 10 可知, 棉蚜虫、棉红蜘蛛与棉铃虫的发生与蔓延均具有相似的条件。随时间推移, 棉蚜虫与棉红蜘蛛的发生蔓延在 1 周内

相关系数达到 0.863, 随后逐渐降低, 棉蚜虫与棉铃虫的发生蔓延相关系数从开始的 0.394 逐渐降低, 而棉红蜘蛛与棉铃虫的发生蔓延相关系数在 2 周内保持为 0.35~0.40; 从 CCF 相关性系数变化范围与对比数据来看, 3 种虫害发生初期其发生条件受彼此影响。

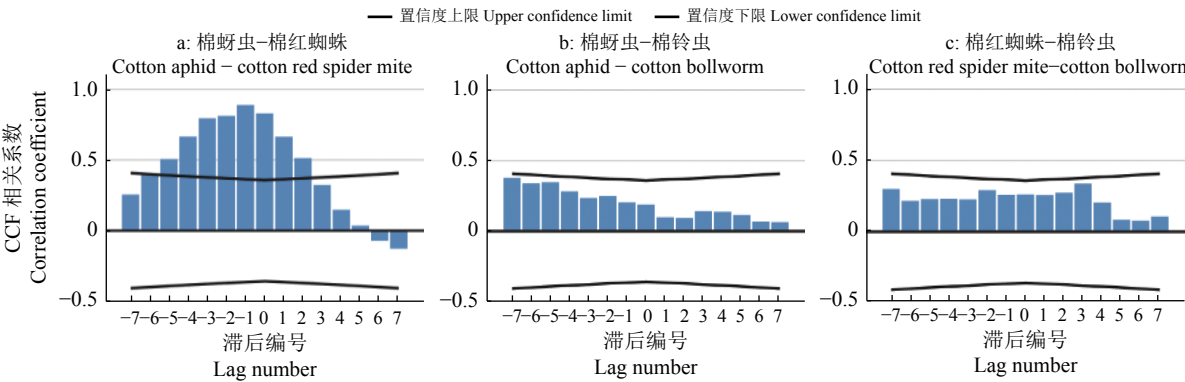


图 10 3 种棉田虫害受灾面积相关性分析

Fig. 10 Correlation a Analysis of the areas affected by three types of pests in cotton field

本研究主要开展棉田虫害识别监测与分类, 对于虫害识别, 健康棉叶与受虫害棉叶光谱反射率有明显偏差, 受虫害侵蚀越严重, 其反射率偏差越明显。虫害发生早期症状不明显, 其反射率数值偏差较低, 但在开展虫害识别监测与分类研究时, 经量化细分, 反射率数值偏差仍能较好反映, 因此对样叶的采样标准会对虫害识别模型结果造成影响, 但

仅表现于虫害发生早期, 不会造成显著影响。

3 结论

利用 Logistic 回归判定方法提取棉蚜虫、棉红蜘蛛、棉铃虫发生区域, 由 SAVI 模型和 NDVI 模型建立的棉蚜虫、棉红蜘蛛、棉铃虫识别模型为最优模型, 其训练样本准确率和测试样本准确率分别达

到 93.7% 和 90.5%, 召回率、F1 值分别为 96.6% 和 93.5%。本研究在识别 3 种虫害发生区域过程中使用了 Logistic 模型, 并利用自相关性时间序列检测法对虫害发生区域进行动态识别监测, 从试验数据来看, Logistic 模型在棉田虫害区域监测的精准度达到较高水平, 该方法可应用于多种虫害的监测识别, 并可基本满足棉田精准植保作业相关要求。借助识别区域 Shapefile 矢量文件, 可为精准植保作业提供经纬度等数据, 便于实现智慧农田管理的相关作业。

本研究仍存在一定的局限性, 遥感影像数据优化过程中使用方法仍存在改进空间, 此外对于会造成虫害发生与蔓延的人为土地管理干预、虫害天敌、极端气候气象条件等影响因子未做进一步研究, 后续开展研究时, 应多方面考虑虫害发生与蔓延的相关因子, 以提高虫害动态监测模型精度。

参考文献:

[1] 乔艳艳, 吴洁, 操宇林, 等. 不同药剂处理对棉种出苗及苗期病虫害发生的影响[J]. 安徽农业科学, 2020, 48(4): 125-127.

[2] 新疆农业科学院. 新疆棉花产量全国占比 84.9% 比重创新高[J]. 新疆农机化, 2020(1): 6.

[3] 杜江涛. 棉花灌溉决策指标研究[J]. 农业科学, 2021, 11(4): 6.

[4] 杨忠娜, 唐继军, 喻晓玲. 新疆棉花产业对国民经济的影响及对策研究[J]. 农业现代化研究, 2013, 34(3): 298-302.

[5] 宋桂红. 新疆奎屯垦区机采棉品种引用、推广现状及对棉花品质的影响[J]. 中国种业, 2016(8): 44-45.

[6] 王磊, 周建平, 许燕, 等. 农用无人机的应用现状与展望[J]. 农药, 2019, 58(9): 6-11.

[7] 崔美娜, 戴建国, 王守会, 等. 基于无人机多光谱影像的棉叶螨识别方法[J]. 新疆农业科学, 2018(8): 1457-1466.

[8] 兰玉彬, 邓小玲, 曾国亮. 无人机农业遥感在农作物病虫害诊断应用研究进展[J]. 智慧农业, 2019, 1(2): 1-19.

[9] 赵亮, 陈兵, 肖春华, 等. 棉花棉叶螨叶片遥感监测技术研究[J]. 甘肃农业大学学报, 2015(5): 94-99.

[10] PRABHAKAR M, PRASAD Y G, THIRUPATHI M, et al. Use of ground based hyperspectral remote sensing for detection of stress in cotton caused by leafhopper (Hemiptera: Cicadellidae)[J]. Computers & Electronics in Agriculture, 2011, 79(2): 189-198.

[11] 竞霞, 黄文江, 琚存勇, 等. 基于 PLS 算法的棉花黄萎病高空间分辨率遥感监测[J]. 农业工程学报, 2010(8): 229-235.

[12] 胡根生, 吴问天, 罗菊花, 等. 结合 HJ 卫星影像和最小二乘孪生支持向量机的小麦蚜虫遥感监测[J]. 浙江大学学报(农业与生命科学版), 2017(2): 211-219.

[13] STELLA I R, GHOSH M. Modeling and analysis of plant disease with delay and logistic growth of insect vector[J]. Communications in Mathematical Biology and Neuroscience, 2018: 2018.

[14] 冯炼, 吴玮, 陈晓玲, 等. 基于 HJ 卫星 CCD 数据的冬小麦病虫害面积监测[J]. 农业工程学报, 2010, 26(7): 213-219.

[15] 马勇, 徐鑫, 徐文君, 等. Logistic 方程模拟森林病虫害发生[J]. 辽宁林业科技, 2001(3): 17-18.

[16] MARCUS R. Deterministic and stochastic logistic models for describing increase of plant diseases[J]. 1991, 10(2): 155-159.

[17] YANG C, EVERITT J H, FERNANDEZ C J. Comparison of airborne multispectral and hyperspectral imagery for mapping cotton root rot[J]. Biosystems Engineering, 2010, 107(2): 131-139.

[18] 张连翔, 惠兴学, 赵树清. 植物病害 I-S 关系: Logistic 衍生模型的研究[J]. 沈阳农业大学学报, 2001(4): 270-273.

[19] 谭宏卫, 曾捷. Logistic 回归模型的影响分析[J]. 数理统计与管理, 2013(3): 476-485.

[20] LI Y X, SHI H M. Research on data processing of low altitude photogrammetry based on Pix4D[J]. International Journal of Intelligent Information and Management Science, 2020, 9(1): 195-197.

[21] NISHAR A, RICHARDS S, BREEN D, et al. Thermal infrared imaging of geothermal environments and by an unmanned aerial vehicle (UAV): A case study of the Wairakei – Tauhara geothermal field, Taupo, New Zealand[J]. Renewable Energy, 2016, 86: 1256-1264. doi: 10.1016/j.renene.2015.09.042.

[22] 刘坦. Inpho、PhotoScan 及 Pix4D 无人机正射影像处理软件对比[J]. 海峡科技与产业, 2017(11): 82-84.

[23] 刘世生, 王仁宗, 李叶民, 等. 一种基于卫星影像数据的新疆棉花区域识别方法及系统: CN111345214A[P]. 2020 - 06 - 30 [2020 - 03 - 04] .

[24] 杨杰, 张莹莹, 王建雄, 等. 利用 NDVI 与 EVI 再合成的植被指数算法[J]. 遥感信息, 2020, 35(5): 131-137.

[25] 邓江, 谷海斌, 王泽, 等. 基于无人机遥感的棉花主要生育时期地上生物量估算及验证[J]. 干旱地区农业研究, 2019, 37(5): 55-61.

[26] 唐亮, 何明珠, 许华, 等. 基于无人机低空遥感的荒漠植被覆盖度与归一化植被指数验证及其对水热梯度的响应[J]. 应用生态学报, 2020(1): 35-44.

[27] 贾洁琼, 刘万青, 孟庆岩, 等. 基于 GF-1 WFV 影像和机器学习算法的玉米叶面积指数估算[J]. 中国图象图形学报, 2018, 23(5): 107-117.