

吕伟,宋轩,杨欢. 基于深度学习和多源遥感数据的玉米种植面积提取[J]. 江苏农业科学,2023,51(23):171-178.
doi:10.15889/j.issn.1002-1302.2023.23.026

基于深度学习和多源遥感数据的玉米种植面积提取

吕伟¹, 宋轩¹, 杨欢²

(1. 郑州大学网络空间安全学院,河南郑州 450001; 2. 郑州大学水利与土木工程学院,河南郑州 450001)

摘要:玉米作为我国主要粮食作物之一,及时准确监测其种植范围及面积对农业产能评估、保障粮食安全具有重要意义。以华北平原典型农业区——原阳县为例,基于欧空局 Sentinel-1 SAR 和 Sentinel-2 MSI 遥感影像数据,在谷歌地球引擎云平台的支持下通过提取雷达后向散射系数时序曲线以及归一化植被指数(NDVI)时序曲线,搭建卷积神经网络(CNN)模型,并将时序数据输入模型得到典型地物分类结果,提取了研究区玉米种植区域,利用野外调查数据进行精度验证,并与随机森林分类对种植区的提取结果进行对比。结果表明,基于光学和 SAR 融合遥感影像数据的识别效果最佳,总体精度达到 93.33%, κ 系数为 0.911;与随机森林分类法相比,卷积神经网络分类的总体精度更高,分类效果更好。因此,采用卷积神经网络以及多源遥感数据的融合能够实现玉米种植面积的准确监测。

关键词:深度学习;多源遥感;融合数据;卷积神经网络;种植识别;时序数据;玉米

中图分类号:TP79;S127 **文献标志码:**A **文章编号:**1002-1302(2023)23-0171-07

玉米作为我国主要粮食作物之一,及时准确监测玉米的种植面积对掌握玉米产量、保障粮食安全以及经济和环境的可持续发展有至关重要的意义^[1]。传统玉米种植面积监测主要通过问卷调查和实地访谈统计信息并逐层上报,需要大量的人力物力^[2],且耗时较长^[3]。遥感分类法是基于 1 幅或多幅玉米种植影像进行面积提取的方法^[4],卫星遥感技术以其覆盖面大、重访周期短、多时空分辨率等优势,在植被面积监测方面取得了巨大成就^[5]。遥感分类法主要包括 2 种,一种选取作物光谱特征差异明显的 1 期影像来实现分类,另一种是基于多期遥感影像,提取作物生育期内光谱变化的时序特征进行面积提取。后者精度更高,是目前作物精细分类的主要方式^[6],但是需要人工预处理并提取相关特征^[7],数据重建工作复杂,不利于快速与自动化分类作物。近年来,深度学习作为机器学习和数据挖掘领域的突破性技术,广泛应用于图像分类、目标识别等领域。深度学习在特征表示方面具有特有的灵活性,基于原始输入数据和输入标签,使

用卷积运算自动提取特征,实现模型端到端训练,不需要专家先验知识即可实现高精度、自动化分类^[8]。Kussul 等基于多源多时序遥感数据对比分析后认为,卷积神经网络(CNNs)要优于随机森林方法,各种作物的目标分类精度均在 85% 以上^[9]。Garnot 等使用 CNN、RNN、R-CNN 对时序哨兵二号数据进行分类,发现所有深度学习模型均优于随机森林模型,突出了深度学习模型在农业地块分类方面的潜力^[10]。目前国内外作物面积提取主要采用多光谱遥感影像,但是夏季作物生长关键期云雨天气频繁^[11],难以获得有效的多时相遥感数据^[12]。相比之下,合成孔径雷达能穿透云、雾,不受天气影响,但雷达影像中存在很大的散斑噪声^[13]和几何形变现象(如透视收缩、叠掩和阴影等)^[14],会影响地物识别的精度。如何将多光谱和雷达图像进行融合,从而提高作物种植面积提取精度显得尤为重要。因此,本研究以华北平原典型玉米种植区的识别为例,基于 Sentinel-1 SAR 和 Sentinel-2 MSI 融合的时间序列遥感数据,采用卷积神经网络算法(CNN),评估 CNN 在时间序列分类任务中的适用性,探索光学和雷达遥感融合数据在作物种植类型识别中的优势。

1 研究区域和数据处理

1.1 研究区域概况

研究区选择河南省新乡市原阳县,该县地处豫

收稿日期:2023-02-24

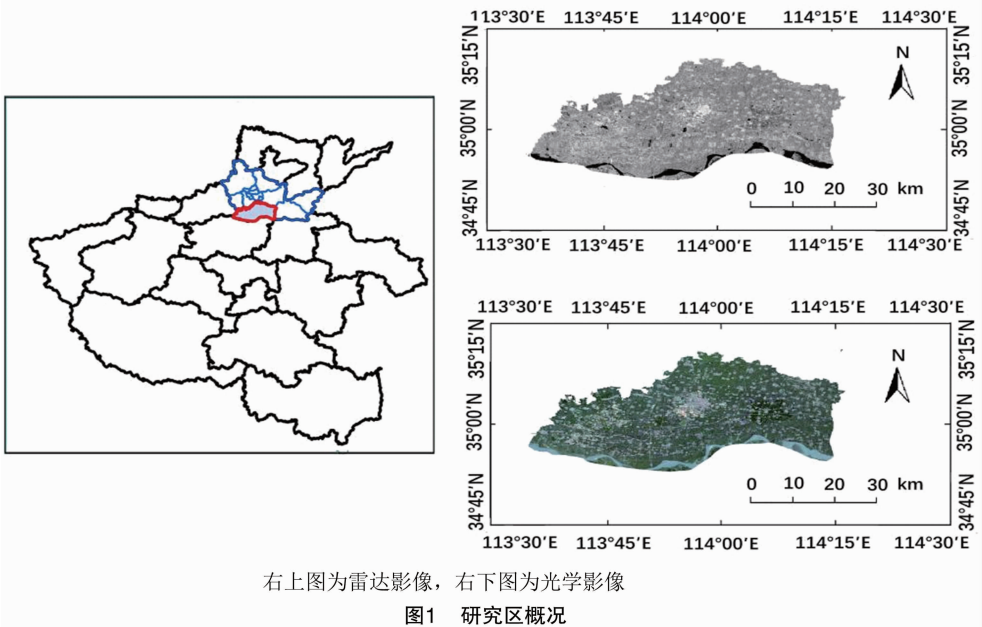
基金项目:国家重点研发计划(编号:2017YFD0800605-5)。

作者简介:吕伟(1998—),女,山东临沂人,硕士,主要从事人工智能与视觉计算研究。E-mail:terminus_lv@163.com。

通信作者:宋轩,博士,教授,博士生导师,主要从事开源地理信息系统研究及应用、生态环境遥感、资源与环境系统计算机模拟与分析研究。E-mail:songxuan@zzu.edu.cn。

北平原,南临黄河,位于 34.55°~35.11°N、113.36°~114.15°E,东接封丘县,西邻武陟县、获嘉县,背靠新乡县、延津县,南与郑州市隔河相望,总面积 1 022 km²。地貌属冲积平原,地势西南高、东北低。温带大陆性季风气候,年均气温约为 14.3℃,平均降水量为 556 mm,全年无霜期 227 d。

原阳县农业发达,是华北平原主要农作物种植区之一,以 1 年 2 熟的作物轮作模式为主,玉米为夏季主要作物之一,一般于 6 月中上旬播种,8 月进入生长旺盛期,10 月中下旬收获,生育期为每年的 6—10 月(图 1)。



右上图为雷达影像,右下图为光学影像
图1 研究区概况

1.2 数据来源及预处理

1.2.1 遥感影像数据 选取哨兵一号雷达影像(Sentinel-1 SAR)和哨兵二号多光谱影像(Sentinel-2 MSI),数据访问、处理及模型构建通过谷歌地球引擎(Google Earth Engine,GEE)云平台进行。哨兵系列影像具有重访周期短、空间分辨率高、影像覆盖范围大、数据免费等优点,是现今较有优势的一种对地观测数据。

1.2.1.1 Sentinel-1 SAR GRD 数据 该数据通过 GEE 云平台获取,采用干涉测量宽幅模式(interferometric wide swath,IW)下的 VH(垂直发射水平接收)极化方式和 VV(垂直发射垂直接收)极化方式,利用 Sentinel-1 Toolbox(S1TBX)工具进行预处理,包括热噪声去除、辐射定标、地形校正等,空间分辨率为 10 m,重访周期为 12 d。

1.2.1.2 Sentinel-2 MSI 数据 Sentinel-2 MSI 采用 Level-1C 产品,共 13 个波段,其中分辨率为 10 m 的波段有 4 个,分辨率为 20 m 的有 6 个,分辨率为 60 m 的有 3 个,光谱覆盖范围从可见光到近红外、短波红外波段^[15],时间分辨率 5 d,其自带质量评估(QA)波段识别云像元。本研究通过 GEE 云平台筛选出作物生长期云量少于 25% 的 8 幅影像。

1.2.2 野外数据 为准确获取样本信息,2020 年 7 月利用 GPS 采集作物分布信息,同时利用 Sentinel-1/2 数据选取其他地物典型像元作为样本,并将所选样本导入 Google Earth,调用同年份高分辨率影像,检查样本正确性。最终获得的样本信息见表 1。以 6:2:2 的比例从中选择 720 个样本进行训练,240 个样本用于验证分类精度,240 个样本用于测试模型准确性。

表 1 样本数据信息

| 地物类型 | 样本类型 | 样本数量(个) | 训练样本数量(个) | 验证样本数量(个) | 测试样本数量(个) | 分类总计(个) |
|-------|------|---------|-----------|-----------|-----------|---------|
| 玉米 | 玉米 | 292 | 172 | 60 | 60 | 292 |
| 非玉米植被 | 水稻 | 284 | 164 | 60 | 60 | 284 |
| 非植被 | 建筑 | 309 | 189 | 60 | 60 | 624 |
| | 水体 | 315 | 195 | 60 | 60 | 60 |

2 研究方法和技术路线

2.1 技术路线

本研究技术路线见图 2。第一,数据集的构建,利用 GEE 云计算平台对 Sentinel-1/2 数据进行预处理,并根据样本坐标点信息提取各种地物类型的不同遥感数据作物生长期时序曲线。第二,对各类时序数据进行组合,在 GEE 云平台上基于不同组合数据集构建单变量或多变量卷积神经网络,对数据进行作物分类,对比试验采用随机森林方法。第三,分析测试数据的分类结果,得到作物分类识别的最优数据组合。

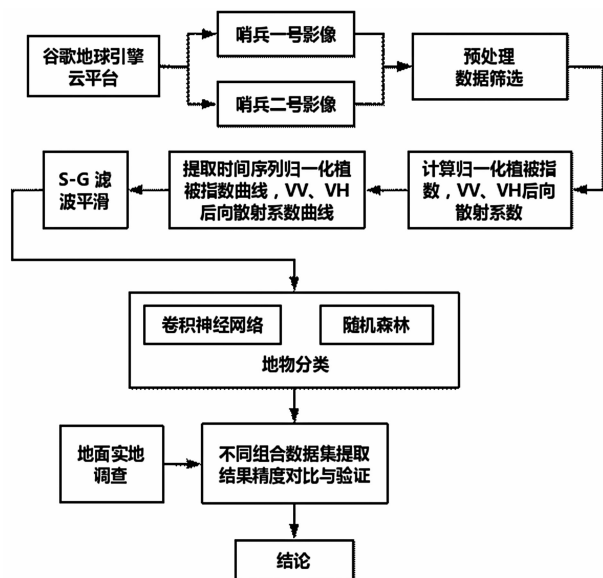


图2 研究方法流程

2.2 时间序列 NDVI/后向散射曲线提取

2.2.1 GEE 提取时序数据 归一化植被指数(normalized difference vegetation index, NDVI)能较好地反映植被绿度变化,是植被生长状态及植被覆盖度的最佳指示因子^[16],其结果在 $[-1, 1]$ 之间。计算公式为

$$NDVI = (\rho_{NIR} - \rho_R) / (\rho_{NIR} + \rho_R) \quad (1)$$

式中:NDVI 表示归一化植被指数; ρ_{NIR} 表示近红外波段反射率; ρ_R 表示红外波段反射率。

将样本坐标点处理成 shp 文件,上传至 GEE 云平台显示标注地图,提取每一地物样本数据集(以实地考察样点为主)相应的时序数据。并通过光学影像质量标识剔除因云雨天气导致质量不高的数据^[17],最终得到相应的 NDVI、VV 后向散射系数、VH 后向散射系数时序数据。

对不同时序数据进行组合,共产生 5 组数据集:

NDVI 数据、交叉极化数据(VH 数据)、单极化数据(VV 数据)、融合 2 类极化数据(VH + VV 数据)、融合 2 类影像提取的 NDVI 和极化数据得到的融合数据(NDVI + VV + VH)。

2.2.2 Savitzky - Golay 滤波平滑 由于云、气溶胶、太阳高度角等因素,原始遥感数据集存在很多噪声,在此情况下提取的时序曲线,波动性较大,不具有代表性。因此,需要对时间序列进行去噪处理。Savitzky - Golay(S - G)滤波被广泛应用于数据流平滑去噪,可以在消除噪声的同时确保信号的形状、宽度不变,是遥感植被指数时间序列的主要滤波方法。它是一种时域低通滤波法,通过局部多项式回归模型平滑时序数据,是移动窗口的加权平均算法,通过将给定高阶多项式的最小二乘法在滑动窗口内拟合得到加权系数。基本思想是:基于多项式在滤波窗口内利用最小二乘法对数据进行拟合^[18]。

$$Y'_j = \frac{\sum_{i=-n}^n C_i Y_{j+1}}{N} \quad (2)$$

式中: Y 表示原始数值; Y' 表示拟合值; C_i 表示第 i 个点的权重; $N = 2n + 1$ 表示滤波窗口的大小。

进行数据清洗后,将每类地物所有样本的时序数据在同一时间上分别求平均,通过 S - G 滤波器对平均时序数据进行滤波处理,得到较平滑的时序曲线^[19]。为得到较好的平滑结果,经多次试验后,滤波核左右各选取 5 点,设平滑多项式次数为 2,最终的标准时序曲线分别见图 3、图 4、图 5。

2.3 卷积神经网络

卷积神经网络(CNN)是深度学习方法中最成功的网络架构之一,是建立在传统人工神经网络上的一种深度学习算法,也是第 1 个成功训练多层网络的算法^[20]。CNN 具有局部连接、权值共享和池化层降采样的特点,与传统神经网络相比,可以减少参数量、降低模型复杂度,并赋予模型对平移和形变的容忍性^[21]。卷积层通过计算卷积核与覆盖区域信号值的点击来确定神经元的输出。

CNN 一般由用于分层提取特征的卷积滤波层和用于计算输出值的全连接层 2 个部分组成。网络层数的选择、卷积核的数量和大小、激活函数等,在 CNN 结构设计中尤为重要。经多次调参尝试,本研究选择 4 层 Conv1D 来提取特征值,每 2 层 Conv1D 后添加 1 层 MaxPooling1D 保留主要特征,减少计算量。每层卷积层使用修正线性单元 ReLU 函数作为激活函数来提高神经网络对模型的表达能力。池化

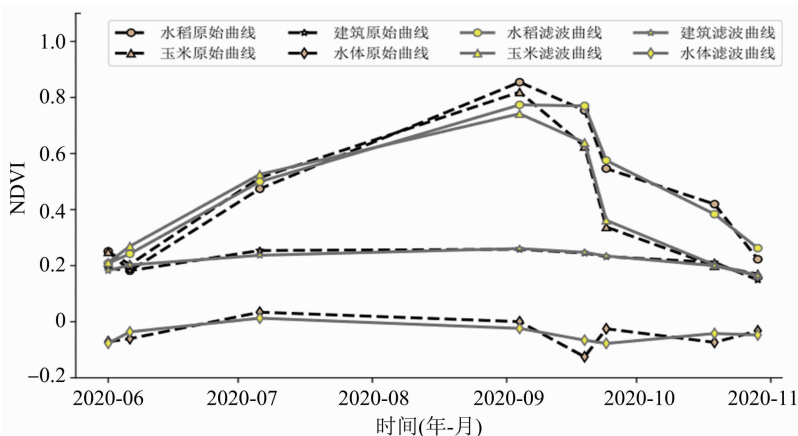


图3 4种典型地物的NDVI时序曲线

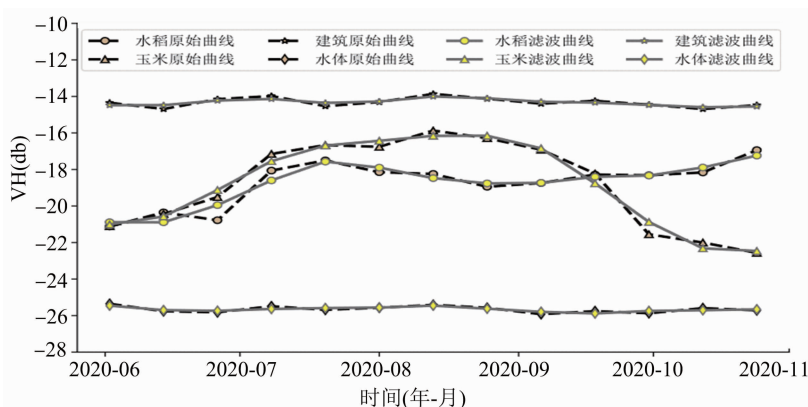


图4 4种典型地物的VH后向散射系数时序曲线

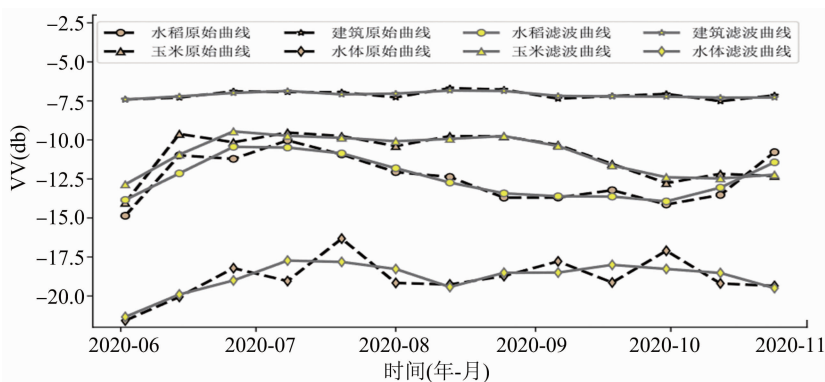


图5 4种典型地物的VV后向散射系数时序曲线

层被固定为最大池化层,窗口大小为2。后面接1个全连接层、1个Dropout层、1个Dense层进行Softmax分类。模型训练选择Adam优化器,损失函数采用交叉熵损失函数(CE_{loss}),公式如下^[22]。

$$CE_{loss} = -\frac{1}{N} \sum_{n=1}^N \sum_{c=1}^C I_c^{(n)} \cdot [\ln P_c^{(n)}] \quad (3)$$

式中: N 表示每批训练样本的数量; C 表示所有类别的集合; $I_c^{(n)}$ 表示当前批次第 n 个样本独热标签类; $P_c^{(n)}$ 表示模型预测样本 n 为 c 类的概率, $\ln P_c^{(n)}$ 用来表示惩罚模型对错误分类的样本的预测。

Dropout是一种正则化技术,通过随机失活一些神经元,防止数据集较小时容易造成过拟合^[23-24],以提高神经网络的性能,设置为0.5。全连接层经过Dropout层后,输入Softmax层,最终输出4种类别的概率结果。经过滤波处理的5组不同特征组合的时序数据集均输入该模型,模型架构见图6。

2.4 精度评价指标

构建混淆矩阵时选取用户精度(user accuracy, UA)、生产者精度(producer accuracy, PA)、总体精度(overall accuracy, OA)、 κ 系数、 F_1 分数(F_1 -score,

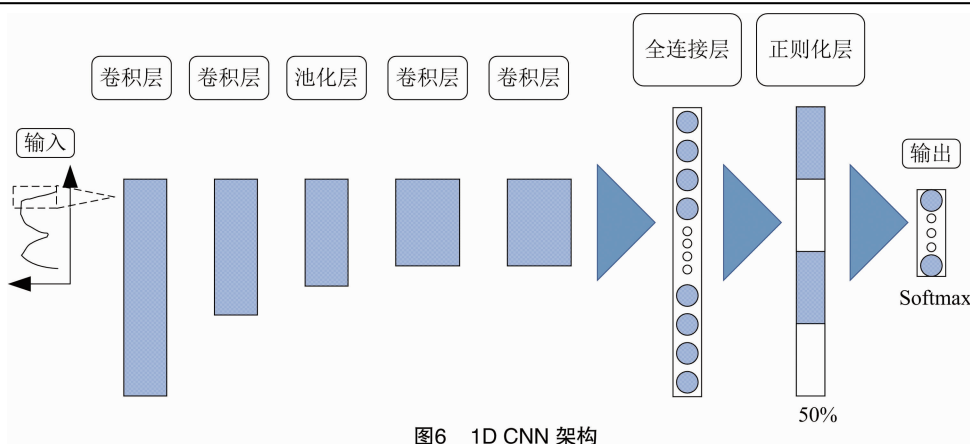


图6 1D CNN 架构

别称平衡 F 分数)5 个评价指标。其中,OA 用于描述验证样本与分类结果一致的概率;PA 表示地面的某类别被正确分类的概率;UA 表示正确分类某类别的概率,主要用来评价分类结果的可信度; F_1 分数指标为综合考虑 PA 和 UA 的调和值^[24]; κ 系数用于描述混淆矩阵的一致性,表征分类结果的可信度,当 $\kappa > 80\%$ 时,表示分类精度最高;当 $40\% < \kappa < 80\%$ 时,表示分类精度中等;当 $\kappa < 40\%$ 时,表示分类精度最差^[25]。

3 结果与分析

整体而言,采用 1D - CNN 分类方法的各种组合数据集均获得较好的分类效果,总体分类精度均高于 85%, κ 系数高于 80%,所有试验结果中,NDVI 数据与 SAR 极化数据融合分类效果最好。分类结果分别见表 2、表 3。随机森林分类效果较差,其总体分类精度及 κ 系数结果见表 3。

表 2 研究区内典型地物分类生产者精度、用户精度、 F_1 分数统计

| 地物 | 类别 | VV 数据 | VH 数据 | NDVI 数据 | (VV + VH) 数据 | (NDVI + VV + VH) 数据 |
|----|-------------|--------|--------|---------|--------------|---------------------|
| 水稻 | 生产者精度(%) | 76.67 | 80.00 | 80.00 | 85.00 | 81.67 |
| | 用户精度(%) | 83.64 | 82.76 | 92.31 | 83.61 | 92.45 |
| | F_1 分数(%) | 80.00 | 81.36 | 85.71 | 84.30 | 86.76 |
| 玉米 | 生产者精度(%) | 83.33 | 83.33 | 91.67 | 85.00 | 91.67 |
| | 用户精度(%) | 70.42 | 89.29 | 83.33 | 86.44 | 85.94 |
| | F_1 分数(%) | 76.34 | 86.21 | 87.30 | 85.71 | 88.71 |
| 建筑 | 生产者精度(%) | 86.67 | 93.33 | 98.33 | 98.33 | 100.00 |
| | 用户精度(%) | 94.55 | 86.15 | 96.72 | 95.16 | 95.24 |
| | F_1 分数(%) | 90.43 | 89.60 | 97.52 | 96.72 | 97.56 |
| 水体 | 生产者精度(%) | 98.33 | 100.00 | 98.33 | 96.61 | 100.00 |
| | 用户精度(%) | 100.00 | 98.36 | 96.72 | 100.00 | 100.00 |
| | F_1 分数(%) | 99.16 | 99.17 | 97.52 | 98.28 | 100.00 |

3.1 基于时序 NDVI 数据的典型地物分类

由表 3 可知,基于时序 NDVI 数据的典型地物分类总精度为 92.08%, κ 系数为 0.894,高于 SAR 极化数据分类精度。从单个类别的分类效果看,采用 NDVI 数据分类在水稻、玉米、建筑等 3 类地物的提取中有较好的分类效果,仅低于 NDVI 数据和极化数据融合的分类效果,其中玉米提取的 F_1 分数为 87.3%,但在水体提取上效果最差。

3.2 基于时序 SAR 后向散射数据的典型地物分类

首先比较 2 种极化方式分类结果,发现 VV 极

化数据分类效果最差,基于时序 VV 极化数据的识别总精度为 86.25%, κ 系数 0.817;VH 极化数据分类效果较好,优于 VV 极化效果,识别总精度为 89.17%, κ 系数 0.856;与单一 VH 极化数据相比,VH + VV 的组合数据分类效果更好,总精度为 91.21%, κ 系数为 0.883,相比于单独利用 VH 或 VV 极化数据,VH + VV 的组合分类总精度分别提高 4.96、2.04 百分点, κ 系数分别提高 0.066、0.027。从单个类别的分类效果来看,后向散射系数对水体的分类效果较好,因为水体与其他地物后向

表 3 研究区内卷积神经网络及随机森林分类结果统计

| 数据类型 | 分类方法 | 总精度 (%) | κ 系数 |
|---------------------|--------|---------|-------------|
| VV 数据 | 随机森林 | 69.08 | 0.585 |
| | 卷积神经网络 | 86.25 | 0.817 |
| VH 数据 | 随机森林 | 70.51 | 0.604 |
| | 卷积神经网络 | 89.17 | 0.856 |
| NDVI 数据 | 随机森林 | 74.24 | 0.656 |
| | 卷积神经网络 | 92.08 | 0.894 |
| (VV + VH) 数据 | 随机森林 | 76.27 | 0.683 |
| | 卷积神经网络 | 91.21 | 0.883 |
| (NDVI + VV + VH) 数据 | 随机森林 | 89.88 | 0.865 |
| | 卷积神经网络 | 93.33 | 0.911 |

散射系数差异明显,单独使用 VV 极化数据和单独使用 VH 极化数据,对水体分类的 F_1 分数均能达到 99% 以上,但是对玉米这类旱地作物分类效果较差,尤其是 VV 极化数据,对玉米分类的 F_1 分数仅能达到 76.34%,但 VH 极化数据对玉米分类的 F_1 分数能达到 86.21%,说明在提取玉米面积时,雷达图像具有可分性。

3.3 基于融合时序 NDVI 和 SAR 后向散射数据的典型地物分类

由表 3 可知,融合数据分类总精度为 93.33%, κ 系数为 0.911。相比 NDVI 数据总精度提高 1.25 百分点, κ 系数提高约 0.017;相对于单一极化数据,融合数据总精度分别提高 7.08、4.86 百分点, κ 系数分别提高 0.094、0.055;相对于 VV + VH 组合数据,融合数据总精度提高了 2.12 百分点, κ 系数提高 0.028。融合数据的玉米提取的 F_1 分数为 88.71%,较其他数据方案均有明显提升。综上,融合 2 类遥感影像提取出的数据用于作物分类效果更佳,对玉米提取精度的提高有明显效果。

3.4 CNN 在玉米面积提取中的表现

整体而言,利用 CNN 对 5 类组合数据进行地物分类,均得到较好的分类效果,总体分类精度均高于 85%,高于随机森林分类方法, κ 系数均高于 80%,说明使用卷积神经网络能够有效地应用于典型地物识别。各数据组合方案中,除单一 VV 极化数据外,玉米提取 F_1 分数均在 85% 以上,说明卷积神经网络能有效区分玉米,能达到玉米识别精度要求。不同数据集整体分类结果较接近,但 NDVI 数据分类结果有一定的错分现象,含有后向散射系数数据的分类结果椒盐现象较明显,这是因为雷达影

像存在斑点噪声,单个像元受噪声影响较大,影响了时序曲线的构建。结合 NDVI 时序数据以及后向散射系数数据的分类精度更高,椒盐现象及错分现象都得到明显改善,减少了将建筑错分为玉米的现象,有利于玉米的遥感提取与制图。

3.5 玉米面积提取

综上,本研究采用卷积神经网络和多源遥感时序数据,获取原阳县的玉米种植面积及分布(图 7),可见玉米为原阳县主要夏季农作物,主要分布于除太平镇和葛埠口乡以外的大部分地区。《2021 年新乡统计年鉴》记录原阳县 2020 年的玉米种植面积为 477.2 km²,分类结果中玉米种植面积为 446.1 km²,提取玉米面积与统计数据的一致性为 93.5%。

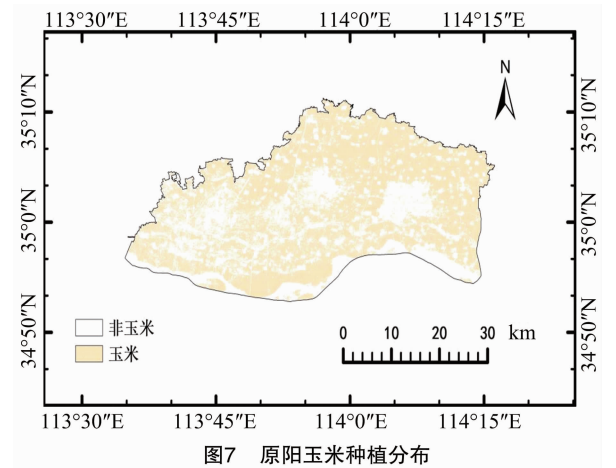


图7 原阳玉米种植分布

4 结论与讨论

4.1 结论

本研究基于 Sentinel-1/2 数据,采用 5 种组合数据方案,结合 GEE 云平台以及 1D CNN 分类器,以华北平原典型农作物种植区为例,通过多光谱和雷达影像的数据融合,基于深度学习对玉米种植面积进行识别与提取,主要结论如下:(1)不同数据组合方式对分类识别结果存在差异,融合数据效果最好,总精度为 93.33%, κ 系数为 0.911。SAR 极化数据识别总精度均在 85% 以上,VV + VH 的组合识别总精度最高,为 91.21%, κ 系数为 0.883;NDVI 数据分类精度高于 SAR 极化数据,总精度为 92.08%, κ 系数为 0.894。(2)针对玉米提取,除 VV 极化数据效果较差,另外 4 种方案的 F_1 分数均能达到 85% 以上。(3)采用卷积神经网络,总体精度均在 85% 以上, κ 系数均在 0.8 以上,分类效果均

高于随机森林的分类效果。

4.2 讨论

本研究充分利用 GEE 云平台强大的计算能力,通过建立研究区遥感地物分类模型,分析采用 CNN 算法使用 Sentinel-1/2 数据对华北平原玉米种植面积进行提取的效果。从方法上看,采用卷积神经网络,使用光学遥感数据或 SAR 后向散射系数数据,均能有效提取玉米种植面积,相较于随机森林算法而言,卷积神经网络能更好地提取作物时序信息,5 种数据组合分类效果均高于随机森林的分类效果,总体分类精度均高于 85%, κ 系数均高于 80%,说明使用卷积神经网络能够有效地应用于典型地物识别。玉米提取的 F_1 分数除采用 VV 数据的分类方式外,均能达到 85% 以上,说明卷积神经网络能有效提取玉米种植面积,满足精细农业种植面积提取的需求,证明卷积神经网络在作物分类等应用中具有巨大的潜力。从数据上看,光学遥感数据分类效果比 SAR 后向散射系数数据分类效果更好。光学遥感数据对作物提取更有效,水体的后向散射系数与其他地物后向散射系数差异明显,故雷达影像数据对水体的提取效果更好。另外,在光学数据缺失的情况下,采用雷达后向散射时序数据也能达到较好的识别精度,对玉米的提取效果满足精细农业种植面积提取的需求,说明雷达影像在玉米面积提取上具备可分性。融合数据分类精度较单一数据分类精度更高,NDVI 数据分类结果有一定的错分现象,含有后向散射系数数据的分类结果椒盐现象较明显,这是因为雷达影像存在斑点噪声,单个像元受噪声影响较大,会影响时序曲线的构建。但是,SAR 数据的地物后向散射特征不同于光学影像,其穿透性不仅能获取植被的表面信息,对植被的茎、枝、叶等信息也有一定的反应,结合光学影像的光谱特性,能增加作物种植区的识别精度。结合 NDVI 时序数据以及后向散射系数数据椒盐现象及错分现象都得到明显的改善,降低将建筑错分为玉米的现象,有利于玉米的遥感提取与制图。综上,数据融合在一定程度上可以提高作物可分性,对提高作物分类精度有一定的效果。

参考文献:

- [1] 李 俐,孔庆玲,王鹏新,等. 基于时间序列 Sentinel-1A 数据的玉米种植面积监测研究[J]. 资源科学,2018,40(8):1608-1621.
- [2] Sonobe R, Yamaya Y, Tani H, et al. Assessing the suitability of data

from Sentinel-1A and 2A for crop classification[J]. *GIScience & Remote Sensing*,2017,54(6):918-938.

- [4] 潘 力,夏浩铭,王瑞萌,等. 基于 Google Earth Engine 的淮河流域越冬作物种植面积制图[J]. 农业工程学报,2021,37(18):211-218.
- [4] 魏鹏飞,徐新刚,杨贵军,等. 基于多时相影像植被指数变化特征的作物遥感分类[J]. 中国农业科技导报,2019,21(2):54-61.
- [5] 黄启厅,曾志康,谢国雪,等. 基于高时空分辨率遥感数据协同的作物种植结构调查[J]. 南方农业学报,2017,48(3):552-560.
- [6] 罗 明,陆 洲,徐飞飞,等. 基于快速设定决策阈值的大范围作物种植分布的遥感监测研究[J]. 中国农业资源与区划,2019,40(6):27-33.
- [7] 王庚泽,靳海亮,顾晓鹤,等. 基于改进分离阈值特征优选的秋季作物遥感分类[J]. 农业机械学报,2021,52(2):199-210.
- [8] Zou Q, Ni L H, Zhang T, et al. Deep learning based feature selection for remote sensing scene classification[J]. *IEEE Geoscience and Remote Sensing Letters*,2015,12(11):2321-2325.
- [9] Kussul N, Lavreniuk M, Skakun S, et al. Deep learning classification of land cover and crop types using remote sensing data[J]. *IEEE Geoscience and Remote Sensing Letters*,2017,14(5):778-782.
- [10] Garnot V S F, Landrieu L, Giordano S, et al. Time-space tradeoff in deep learning models for crop classification on satellite multi-spectral image time series [C]//2019 IEEE International Geoscience and Remote Sensing Symposium. Yokohama, 2019:6247-6250.
- [11] 朱凤敏,吴 迪,杨佳琪. 基于 Sentinel-1B SAR 数据的农作物分类方法研究[J]. 测绘与空间地理信息,2020,43(5):105-108.
- [12] 谢新乔,杨继周,邓邵文,等. 多时相 Sentinel-1 影像反演玉溪典型烟区烤烟种植分布的方法[J]. 农业资源与环境学报,2023,40(1):188-195.
- [13] Ndikumana E, Minh D H T, Baghdadi N, et al. Deep recurrent neural network for agricultural classification using multitemporal SAR sentinel-1 for Camargue, France[J]. *Remote Sensing*,2018,10(8):1217.
- [14] 郭 交,朱 琳,靳 标. 基于 Sentinel-1 和 Sentinel-2 数据融合的农作物分类[J]. 农业机械学报,2018,49(4):192-198.
- [15] 成科扬,荣 兰,蒋森林,等. 基于深度学习的遥感图像超分辨率重建方法综述[J]. 郑州大学学报(工学版),2022,43(5):8-16.
- [16] 姜伊兰,陈保旺,黄玉芳,等. 基于 Google Earth Engine 和 NDVI 时序差异指数的作物种植区提取[J]. 地球信息科学学报,2021,23(5):938-947.
- [17] 张 森,吴炳方,于名召,等. 未种植耕地动态变化遥感识别——以阿根廷为例[J]. 遥感学报,2015,19(4):550-559.
- [18] 杨泽航,王 文,鲍健雄. 融合多源遥感数据的黑河中游地区生长季早期作物识别[J]. 地球信息科学学报,2022,24(5):996-1008.
- [19] 陈思宁,赵艳霞,申双和. 基于波谱分析技术的遥感作物分类方法[J]. 农业工程学报,2012,28(5):154-160.

龙 阳,肖小玲. 基于多注意力机制的苹果叶部病害检测方法[J]. 江苏农业科学,2023,51(23):178-186.
doi:10.15889/j.issn.1002-1302.2023.23.027

基于多注意力机制的苹果叶部病害检测方法

龙 阳,肖小玲

(长江大学计算机科学学院,湖北荆州 434100)

摘要:针对苹果叶部病害在复杂环境下出现的多尺度病斑识别和定位不准确等难题,并提高对小目标病害特征的识别能力,提出了一种基于改进 YOLO v5s 的苹果叶部病害检测算法。首先,在骨干网络中采用 ConvNeXtBlock 模块替换传统的 CSP 模块,以增强对病害特征的提取能力。同时,为了更准确地检测小目标病害特征,引入自注意力和卷积集成的 ACmix 注意力机制,使得改进后的骨干网络在复杂环境下也能更准确地检测到苹果叶部病害。在颈部网络中,采用通道和空间混合的 CBAM 注意力机制以及 SK 注意力机制,以进一步提高模型对多尺度病害特征的精确定位和提取能力,并赋予了模型更强的语义信息理解能力,使其能够更好地适应不同尺寸和复杂度的病害状况。试验结果表明,与传统的 YOLO v5s 相比,改进后的算法精确率提高了 12.2 百分点,召回率提高了 12.0 百分点, $mAP@0.5$ 提高了 12.72 百分点,平均精度达到 94.03%,具有较好的识别精度,误检和漏检概率显著减少。本研究的改进算法在苹果叶部病害检测方面取得了显著成果。该算法不仅提高了对多尺度病害特征的识别和定位能力,在复杂环境下也能更准确地检测到小目标病害特征。这一算法的应用潜力不仅限于苹果叶部病害,还可推广至其他农作物病害检测领域,对于提高农业病害防控水平具有重要意义。

关键词:苹果叶部病害;YOLO v5s;ConvNeXtBlock;ACmix;CBAM;SK;注意力机制

中图分类号:TP391.41 **文献标志码:**A **文章编号:**1002-1302(2023)23-0178-09

当今社会,农业产业的发展对于粮食供应和经济增长至关重要。然而,农作物病害是农业生产的主要障碍之一,尤其是水果作物,其中苹果叶部病害对苹果产量和质量的影响尤为显著。因此,快速、准确地检测和识别苹果叶部病害对于农业生产具有重要意义。

目前,深度学习方法在图像处理和分类任务中取得了显著的突破。其中,卷积神经网络被广泛应用于目标检测和图像分类任务中。针对苹果叶部病害检测,研究者们也尝试利用深度学习方法来提高检测的准确性。例如,王云露等提出了基于改进 Faster R-CNN 的苹果叶片病害识别方法,能够在实际场景中进行无损测量识别,平均精度均值(mAP)达到 86.2%^[1]。Jiang 等提出了一种新的使用 Deep-CNN 的苹果叶病检测模型来检测 5 种常见的苹果叶病。结果表明,INAR-SSD 模型在 ALDD 上达到了 78.80% 的 mAP 检测效果^[2]。Sun 等提出了一种可以实时检测苹果叶片病害的轻量级 CNN 模型,其试验结果表明,该模型能够达到

收稿日期:2023-06-18

基金项目:国家自然科学基金(编号:61771354)。

作者简介:龙 阳(1999—),男,贵州铜仁人,硕士研究生,研究方向为计算机视觉与目标检测。E-mail:2022710657@yangtzeu.edu.cn。

通信作者:肖小玲,博士,教授,研究方向为智能信息处理与网络安全。E-mail:xxl@yangtzeu.edu.cn。

[20] 赵子娟,刘 东,杭中桥. 作物遥感识别方法研究现状及展望[J]. 江苏农业科学,2019,47(16):45-51.

[21] 罗荣辉,袁 航,钟发海,等. 基于卷积神经网络的道路拥堵识别研究[J]. 郑州大学学报(工学版),2019,40(2):18-22.

[22] Kampffmeyer M,Salberg A B,Jenssen R. Semantic segmentation of small objects and modeling of uncertainty in urban remote sensing images using deep convolutional neural networks [C]//2016 IEEE Conference on Computer Vision and Pattern Recognition Workshops. Las Vegas,2016:680-688.

[23] 屈 炀,袁占良,赵文智,等. 基于多时序特征和卷积神经网络的农作物分类[J]. 遥感技术与应用,2021,36(2):304-313.

[24] Zhong L H,Hu L N,Zhou H. Deep learning based multi-temporal crop classification[J]. Remote Sensing of Environment,2019,221:430-443.

[25] 王小慧,姜雨林,傅漫琪,等. 海河低平原典型县种植制度与农田景观格局变化遥感监测[J]. 农业工程学报,2022,38(1):297-304.